TAL TECH

DOCTORAL THESIS

Forecasting of Electrical Network Corona Losses Based on Machine Learning Methods

Pradeep Kumar Gupta

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2026

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
24/2026

Forecasting of Electrical Network Corona Losses Based on Machine Learning Methods

PRADEEP KUMAR GUPTA

TAL TECH PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Engineering
Department of Electrical Power Engineering and Mechatronics
The dissertation was accepted for the defence of the degree of Doctor of Philosophy on March 24th, 2026.

Supervisor: Professor Jako Kilter,
Department of Electrical Power Engineering and Mechatronics, School of Engineering,
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Dr. Kaur Tuttelberg,
Department of Electrical Power Engineering and Mechatronics, School of Engineering,
Tallinn University of Technology
Tallinn, Estonia

Opponents: Professor Irina Oleinikova,
Department of Electric Energy, Faculty of Information Technology and Electrical Engineering,
Norwegian University of Science and Technology (NTNU)
Trondheim, Norway

Professor Janne Seppänen,
Department of Electrical Engineering and Automation,
Aalto University
Espoo, Finland

Defence of the thesis: May 12th, 2026, Tallinn, Estonia

Declaration:
Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Pradeep Kumar Gupta _____
signature

Financial support was provided by the Estonian Ministry of Education and Science (The Erasmus+ Scholarship for study period abroad scholarship), and by the European Union and co-funded by the Estonian Ministry of Education and Science (Project No TemTa-134). Reasearch data was provided by Estonian Transmission System Operator Elering and by the Estonian Environmental Agency.

ISSN 2585-6898 (publication)
ISBN 978-9916-80-483-4 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-484-1 (PDF)
DOI https://doi.org/10.23658/taltech.24/2026

Gupta, P. K. (2026). *Forecasting of Electrical Network Corona Losses Based on Machine Learning Methods* [TalTech Press]. https://doi.org/10.23658/taltech.24/2026

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
24/2026

Masinõppe meetoditel põhinev elektrivõrgu koroonakadude prognoosimine

PRADEEP KUMAR GUPTA

TAL TECH
KIRJASTUS

Contents

List of Publications 8

Author's Contributions to the Publications 9

Introduction 11

1 Corona Phenomenon and Corona Losses 15
- 1.1 Literature Review 15
- 1.2 Empirical Approaches for Corona Loss Estimation 17
- 1.3 Experimental Studies for Corona Discharge Estimation 18
- 1.4 Estimation of Corona Losses using Phasor Measurement Units 22
- 1.5 Error Correction Method 23
- 1.6 Intermediate Summary 24

2 Weather Dependency on Corona Losses 25
- 2.1 Introduction 25
- 2.2 Methodology and Data Collection 26
- 2.3 Description of Overhead Transmission Line Configuration 27
- 2.4 Influence of Weather Parameters on Corona Losses in Estonian 330 kV Transmission Line 29
  - 2.4.1 Air Temperature 31
  - 2.4.2 Air Pressure 31
  - 2.4.3 Humidity 32
  - 2.4.4 Air Density 34
  - 2.4.5 Precipitation 34
  - 2.4.6 Wind Speed 35
  - 2.4.7 Wind Direction 36
  - 2.4.8 Visibility 36
  - 2.4.9 Conductor Temperature 37
  - 2.4.10 Voltage 38
  - 2.4.11 Seasonal 39
- 2.5 Analysis of Weather Parameter Rankings Based on Feature Importance Techniques 39
- 2.6 Comparison of 10 Minute and Hourly Data 43
- 2.7 Intermediate Summary 44

3 Performance of Corona Loss on Ageing Transmission Lines 45
- 3.1 Background 45
- 3.2 Description of Overhead Transmission Line Configuration and Design Parameters 46
  - 3.2.1 Tower Configuration: 46
  - 3.2.2 GMD, GMR and Bundled Conductor: 46
  - 3.2.3 Conductor Spacing: 48
- 3.3 Seasonal Behaviour of Corona Loss Performance 48
- 3.4 Weekly Comparison of Corona Losses 50
- 3.5 Influence of Conductor Temperature on Corona Losses 51
- 3.6 Influence of Weather Parameters on Ageing Transmission Lines 53
  - 3.6.1 Conductor Temperature 53

3.6.2 Air Temperature 54
3.6.3 Air Pressure 54
3.6.4 Wind Speed 57
3.7 Intermediate Summary 57

4 Machine Learning Framework For Corona Loss Forecasting 59
4.1 Background 59
4.2 Methodology 60
4.2.1 Data Collection and Preprocessing 62
4.2.2 Corona Loss Estimation Method Using PMU Measurement 63
4.3 Machine Learning: Regression Model 63
4.3.1 Multiple Linear Regression Model 64
4.3.2 Ensemble Random Forest Regression 64
4.3.3 Extreme Gradient Boost Regression (XGBR) 65
4.3.4 Error Metrics for Evaluations 66
4.4 Results and Analysis 66
4.4.1 Varying Time Step Length 67
4.4.2 Forecasting Model Seasonal Performance 71
4.4.3 Combined Transmission Line Corona Loss Forecasting 74
4.4.4 Forecasting Performance With Reduced Features 75
4.5 Incorporating Corona Loss Forecasting Scenarios into Operational Planning Activities of Transmission System Operators 78
4.6 Intermediate Summary 78

Conclusions and Further Work 81

List of Figures 84

List of Tables 87

References 89

Acknowledgements 97

Abstract 99

Kokkuvõte 100

Appendix 1 101

Appendix 2 111

Appendix 3 117

Appendix 4 129

Appendix 5 137

Appendix 6 145

Curriculum Vitae 157

Elulookirjeldus 159

List of Publications

The present Ph.D. thesis is based on the following publications that are referred to in the text by Roman numbers.

I P. K. Gupta, J.-R. R. Ruiz, K. Tuttelberg, and J. Kilter, "Quantification of corona discharge intensity applied to sphere-plane configurations," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2025

II P. K. Gupta, J.-R. R. Ruiz, P. Casals-Torrens, J. T. i Garvín, K. Tuttelberg, and J. Kilter, "Sensitivity analysis of corona discharges measuring instruments using different electrodes and high voltage supplies," *IEEE Sensors Letters*, vol. 8, no. 12, 2024

III P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power and Energy Systems*, vol. 155, p. 109537, 2024

IV P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023

V P. K. Gupta, K. Tuttelberg, and J. Kilter, "The weather impact on corona losses of 330 kV aging transmission lines," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, *Dubrovnik*, *Croatia*, pp. 1–5, 2024

VI P. K. Gupta, K. Tuttelberg, and J. Kilter, "Forecasting corona losses on high voltage transmission lines using machine learning," *IEEE Transactions on Power Delivery*, vol. 40, no. 5, pp. 2696–2705, 2025

Author's Contributions to the Publications

I Pradeep Kumar Gupta was the main author, who conducted the experiment for corona discharge detection. Furthermore, developed an algorithm for corona discharge detection using various methods for the sphere-plane configuration, analyzed the results, and wrote the initial manuscript.

II Pradeep Kumar Gupta was the main author, who conducted the experiment for corona discharge detection and analyzed the sensitivity of corona discharge, prepared the figures, and wrote the initial manuscript.

III Pradeep Kumar Gupta was the main author, developed the algorithm for estimation of corona losses using PMUs data, the concept of weather dependency on corona losses, developed the program to test and analyze the feature importance under various conditions to understand their influence, prepared the figures, wrote the initial manuscript, and published the work in IJEPES.

IV Pradeep Kumar Gupta was the main author, developed the algorithm, conducted simulation, prepared the software implementation of the algorithm and analyzed the mass testing results, as well as wrote the manuscript.

V Pradeep Kumar Gupta was the main author, developed the algorithm, conducted simulation, prepared the software implementation of the algorithm and analyzed the mass testing results, as well as wrote the manuscript.

VI Pradeep Kumar Gupta was the main author, developed the algorithm, conducted simulation, prepared the software implementation of the algorithm and analyzed the mass testing results, as well as wrote the manuscript.

Introduction

This thesis presents the research and development work carried out by the author during PhD studies at Tallinn University of Technology (TalTech). Most of this work was conducted at TalTech. However, some of the experimental work to understand the corona loss phenomenon was performed through the Erasmus+ program in Barcelona, Spain, at AMBER laboratory Universitat Politècnica de Catalunya.

Motivation and Background

For more than 115 years, since the development of AC transmission lines, the phenomenon of corona losses has been studied. It is of relevance to determine the intensity of corona and the corresponding losses to more adequately design the transmission network and consequently plan its operation. Over the years, various approaches have been implemented to determine the level of corona losses. However, these methods have been of limited accuracy, or are difficult to implement in real power networks. With the advancement of modern measurement and communication possibilities, more sophisticated algorithms and approaches can be developed, leading to higher accuracy.

Transmission system operators (TSOs) face the challenging task of estimating network losses during network operation under various conditions. Network losses in general can be divided into two categories: losses that are directly dependent on network elements and their operation, which can be easily calculated, and losses that are dependent on other factors, including weather, and which are difficult to determine, as direct calculation methods are not available, are termed as corona losses. In modern power systems, it is of relevance to estimate these types of losses as accurately as possible. This is mainly because the TSOs need to purchase the network losses from the energy markets, and therefore the more accurate estimation of losses is of relevance due to direct influence on TSOs operational costs.

In modern power systems, the network operation can be monitored by synchronized phasor measurement based wide-area monitoring system (WAMS). This system enables monitoring the network operation in real-time and makes it possible to develop various approaches and algorithms to enable more sophisticated network monitoring and operation assessment. This system can also be used to estimate and monitor the network corona losses and this can be done without the need to install additional dedicated equipment. It is possible to perform long-term monitoring of corona losses over the whole network, leading to a better understanding of the phenomenon. The system-wide long-term measurement results can be combined with weather monitoring data to gain a deeper understanding of corona losses. In addition, using machine learning techniques, the complex relation can be modelled and monitored using suitable forecasting methods.

Wide-area monitoring systems have been used for various purposes, including system inertia monitoring, voltage stability monitoring, etc. Its use for corona loss estimation however cannot be straightforwardly determined as the suitable reference cannot be found in publicly available literature. Currently, there is no standard strategy for forecasting corona losses. It is challenging to monitor and estimate this type of loss due to limited data availability over significantly longer periods of time. In addition, corona losses have a complex relation with weather conditions, and it is difficult to measure or monitor corona loss under atmospheric conditions, especially for short duration. Therefore, to overcome the limitations an approach of wide-area based measurements together with machine learning has been proposed in this thesis.

The machine learning techniques together with wide-area based measurements provide an alternative to solve the corona loss estimation challenge in a more efficient and re-

liable manner. If there is sufficient data available and alternative modelling techniques do not provide effective solutions, then machine learning tools may be appropriate. From examining weather dependence to predicting corona losses based on weather projections, machine learning models are easily accessible with their feasibility and highly effective functions. There is potential to anticipate the corona losses using easily available weather data. The main purpose of this research is to provide corona loss forecasting tools that significantly outperform analytical and empirical solutions in terms of reliable functionality and usability in real transmission networks. It is seen that these kind of solutions would offer TSOs with significant information to improve the accuracy of their network loss forecasts and consequently to reduce the respective costs.

Main Objectives and Tasks of the Thesis

The key research question asked in this thesis is "*How can we reliably monitor and forecast corona losses in future electric power systems?*". The answer to this question should form an understanding on how the currently used and new developed methodologies cope with changing weather conditions to obtain reliable corona loss results for overhead transmission lines.

To answer the key question formulated for the thesis, several hypotheses were devised:

- Corona loss monitoring data can be analysed to understand how corona losses depend on different weather phenomena and related measured quantities.
- Corona losses increase with the ageing of a transmission line.
- Corona loss and weather monitoring data can be used to train machine learning models for the forecasting of corona losses based on weather forecasts.

Based on the hypotheses listed above, the following tasks were formulated:

- Experimental studies to visualise corona discharge phenomenon.
- Use of synchronized phasor measurement data to estimate corona losses in an overhead transmission lines.
- Assessment of the effects of changing weather parameters on corona losses.
- Utilisation of the feature importance method to select the most influential weather parameters.
- Assessment of the influence of weather parameters on ageing transmission lines
- Development of a forecasting model using machine learning to forecast corona loss efficiently. Assessment and comparison of the performance of different developed methods.

Contribution of the Thesis and Dissemination

The main objective of this thesis is to develop a fundamental knowledge of the forecasting scenario for corona losses in overhead transmission lines using phasor measurements and corresponding weather data. This work will facilitate the transmission system operators by reducing power loss purchases, enabling more detailed monitoring of losses, and increasing system reliability through more effective network operational planning.

Theoretical Contribution:

- The results of the research provide valuable information and a strategy to overcome existing challenges of power systems to monitor transmission network corona losses under complex weather conditions.
- The results of the research provide understanding of the corona loss variation considering the dependence of weather conditions and ageing transmission line infrastructure.
- The developed machine learning based approach will enable more accurate online monitoring of corona losses and thereby enabling more accurate system control and increased operational reliability.

Practical Contribution of the Dissertation:

- The methods and tools developed enable the TSOs to decrease the financial costs related to purchasing network losses from the market.
- The results of the thesis enable the TSOs to continuously monitor the corona losses on operating transmission lines without the need of additional dedicated equipment.
- The developed machine learning based algorithm is useful in predicting historic corona losses based on inputs of historic weather data and also estimating the future corona losses given the inputs from a weather forecast.

Dissemination of the Research Work

The results of this research have been published in the form of papers in various internationally recognized journals and conferences. The experimental research work to visualize corona discharge has been published in peer-reviewed journals (publication [**I**] and [**II**]). The impact of weather dependency on corona losses using phasor measurement data together with feature importance analysis has been published in a peer-reviewed journal [**III**]. The assessment of the dependence of corona losses on weather parameters and their relation to ageing transmission line infrastructure has been published in a journal [**IV**] and also presented at a conference [**V**]. The developed machine learning forecasting model for real-time transmission network corona losses estimation has been published in a peer-reviewed journal [**VI**].

The connection between publications, the factors and domain of corona phenomenon and corona losses, and the research questions of the thesis are depicted in Fig. 1. Publication I and II focuses on how can corona intensity be quantified under controlled conditions and how reliable are corona measurement instruments. Publication III addresses the different weather parameters influence on corona losses in high voltage transmission lines. Moreover, the ranking of weather parameters has been established to gain insights into which weather factors exert the greatest influence. Publication IV and V examined the design configuration and weather parameter influence on aging transmission lines. Publication VI addresses the monitoring and forecasting of corona losses using weather considering various scenarios and measure the accuracy level, which can be used by TSO for monitoring.

Outline of the Thesis

The thesis is divided into four main parts to form a comprehensive overview and discussion on the experimental studies, weather dependency statistical analysis, impact of ageing transmission lines, feature importance to select important weather parameters, as

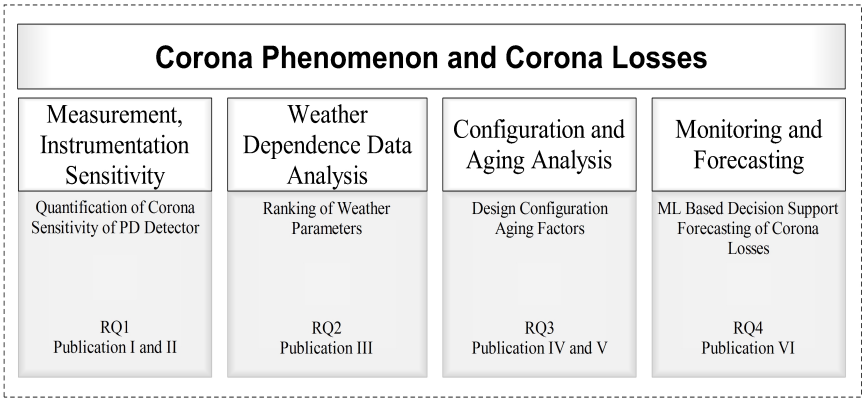


Figure 1: Connection between main domain of the corona phenomenon and corona losses monitoring, forecasting and decision making, research questions and publication used in the thesis.

well as developing a forecasting model using machine learning to forecast corona losses under different circumstances. In addition, the thesis includes an introduction and ends with conclusions and recommendations for future work. Overview of Chapters is as follows:

- Chapter 1 focuses on corona phenomena and corona losses in general. This includes, an overview of corona discharge assessment focusing on the use of various high voltage electrodes, together with discussion on visualizing the intensity of a corona using an image processing method with DSLR camera images, and overview of various other methods to detect a corona using a shielded laboratory experimental setup.
- Chapter 2 presents a corona loss estimation approach using phasor measurement data and determines the weather dependence on corona losses for 330 kV transmission line using statistical and feature importance methods. In addition, a time resolution analysis is presented, enabling influence of weather parameters to be assessed based on the use of data with various time resolutions.
- Chapter 3 discusses the corona loss performance in case of ageing transmission lines infrastructure.
- Chapter 4 presents an application of the machine learning algorithm to forecast corona losses in transmission networks using weather parameters.

1 Corona Phenomenon and Corona Losses

In this Chapter, a brief overview and discussion of the phenomenon of corona discharge resulting in the occurrence of corona losses in overhead transmission lines is presented. Additionally, this section includes discussion on empirical approaches for corona loss estimation and presents results of experimental analysis of corona discharge including different case scenarios conducted on spherical electrodes, needle to plane, and new conductors using DSLR cameras and partial discharge detectors. Further discussion is provided on estimation method for corona losses based on phasor measurements and as well as on systematic error correction method.

Analysis results and discussion presented in this Chapter are published in I and II.

1.1 Literature Review

Since 1911, various researchers have proposed different techniques to calculate corona losses in transmission lines. Some of the early works were proposed by Peek [7] and G. Faccioli [8], where the study of corona phenomena was investigated in laboratories under different conditions. Furthermore, Peek introduced a series of experimental results conducted for various factors influencing corona loss curves to understand the quadratic law above visual critical voltage [9, 10, 11] and the influence of harmonics [12]. In 1933, Peterson proposed the development of a theoretical formula that has modified an empirical equation with a correction in response to calculating both lower and higher levels of corona loss, which should be of interest to engineers working on corona discharges [13]. Another empirical method for calculating corona loss from high-voltage transmission lines was proposed by Carroll and Rockwell in 1937 [14]. More profound discussion on empirical approaches is presented in Section 1.2. In case of conventional methods, there is not much detailed information available on the influence of various weather conditions. Therefore, some of the researchers studied the behaviour of corona loss in desert conditions [15], snow [16] and rain conditions [17], etc.

It is well known that corona loss performance is mainly influenced by the type of voltage (AC or DC), the type of conductor (which includes different dimensions and sizes), surface irregularity of the conductor, rusting or other forms of surface deterioration, and climatic changes such as snow, rain, icing or desert conditions, etc., that involve changes in temperature, pressure, humidity, precipitation, etc. Evaluation of corona losses under complex weather dependence has been studied for many years using different experimental settings and environmental conditions [18, 19, 20]. In order to develop measures and understand the behaviour of corona discharge under hoar frost conditions, a coaxial measurement arrangement for corona loss detection was performed with different types of conductors and number of bundles [21]. The observation discussed that the two-conductor bundle had corona losses about 2-5 times higher than three-conductor bundle. Similarly, In [22], an indoor experimental setup was carried out with simulated rainfall on DC-energized overhead conductors, where rain intensities (3.6-36.5 mm/h and dry conductor) and different types of conductor ($E_{\max}$ (ideal, strands and real)), were investigated. Further, the positive corona generated by the rime ice-covered conductor and audible noise characteristics were examined to analyse the effect of severity of icing in [23]. In addition, to understand the performance of corona loss, the influence of wet weather transitions and under varying humidity were studied in [24]. Table 1 and Table 2 present an overview of different literature that discusses the impact of various climatic weather conditions and other factors that influence corona loss in overhead transmission lines and high voltage equipment. A more detailed analysis of the effects of different

weather parameters on corona losses is presented in Chapter 2.

Table 1: Weather factors influencing corona discharge in an overhead transmission line

Weather Condition	Corona Losses	Literatures	Results
Sunny/Fair	Low	[25]	During fair weather conditions an increase in the number of pulses of higher amplitude was seen with reduced humidity. The radio interference voltage increases when the corona current decreases, whereas the audible noise of their HVDC line was seen reduced during rain and rises during dry weather.
Wet/Rain	Medium	[22] [23] [26] [27]	During rainy conditions corona current behaved in a linear manner until 100 kV and started exponentially magnify with increasing voltage. It was observed that for dry conductors, even at the highest applied voltages, corona currents are negligible as compared to rain conditions. The rate of corona losses is high, especially during light rainfall but tends to saturate as the amount of rainfall increases.
Snow	High	[20] [25] [27]	Corona losses in wet snowfall conditions are very much similar to rainfall increasing in nature and it becomes half during dry snowfall.
Hoar Frost	High	[10] [19] [21]	It was found that the highest corona losses were caused by hoarfrost compared to sunny or rainy weather.
Icing	Very High	[28]	The audible noises keep increasing with the severity of icing conditions at a given nominal electric field intensity. During heavy icing severity it was observed that the audible noises saturated at a certain point.

Table 2: Factors influencing corona discharge in an overhead transmission line

Weather Condition	Corona Losses	Literatures	Results
Altitude	Low	[25] [29]	With the increase in altitude it was observed that corona inception voltage significantly decreased with the decrease of air pressure. The negative/positive corona inception electric field of the conductors decreased but in the case of experiments it remains high as the altitude rises.
Conductor Type	Medium	[21] [22] [25]	The untreated standard 22-mm diameter conductor with round strand profiles (Type A) leads to the highest corona currents whereas the sandblasted 32-mm diameter Z-profile conductor (Type G) leads to the lowest values.
Humidity	High	[24] [25] [27]	Due to an increase in humidity corona losses increase for both +ve and -ve corona.The relative humidity plays a key factor when it comes to corona during fair weather conditions. An increase in the number of pulses of higher amplitude was seen with reduced humidity.
Air Density	Low	[27]	In the case of increasing air density, it decreases.
Harmonics	High	[11] [30]	The magnitude of harmonics increases when the corona losses increases with increasing voltage

Physical phenomena such as sound hissing noise, light violet glow, strong electric field, magnetism, etc., are all part of the corona discharge phenomenon process. Therefore, by detecting these physical signals, the condition of high voltage equipment may be assessed.

The most widely used methods for detecting corona discharge include the PD measuring instrument (PD Base II, Techimp-Altanova Group, Zola Predosa, Italy) with an HFCT pulse current approach, the ultra-high frequency (UHF) method [31], the acoustic NL Camera [32], the Corona camera, acoustic sensors, and the ultrasonic method [33], etc. However, because of the complex interference, difficulty in the localization of corona discharge and inadequate capabilities of the instrument, sometimes a false alarm may occur. In this case, the optical detection method is very useful as they are highly sensitive and offer robust resistance to acoustic and electromagnetic vibration interference [32, 34]. It has drawn a lot of interest in the corona detection area where the optical methods are widely used for their simplicity and low cost. This alternative approach for corona detection, especially in the case of pulseless glow, involves the sensitivity and application of DSLR cameras for corona detection under atmospheric conditions for various high-voltage equipment. A few studies that have examined the specific implications of cameras for discharge detection under various environmental conditions with different setting are presented in [35, 36, 37] . The visual corona extinction voltage corresponding to various spherical radii for different supply sources was studied in [38]. To detect and locate the high-voltage corona discharges using DSLR and Sonic Cameras was proposed in [32]. In [39], corona extinction voltage is determined experimentally using a gas-filled tube solar-blind ultraviolet (UV) sensor and a CMOS imaging sensor, with electric field strength calculated by finite element method (FEM) simulations for different frequencies that range with various pressure levels. Under +DC, -DC, and AC voltages, the authors propose the application of a digital single-lens reflex (DSLR) camera to analyse corona discharge light patterns captured to detect corona initiation for different electrodes [34].

1.2 Empirical Approaches for Corona Loss Estimation

Corona losses refer to power losses caused by corona discharge on transmission line conductors. The phenomenon was first observed during the early stages of the development of high-voltage AC power transmission lines and has a significant impact on power transmission efficiency. When a corona discharge appears on the conductors, the high voltage source attached to the conductor provides the energy needed to produce the high electric field conditions required for corona discharges to occur at the conductor surface. Corona losses are mostly caused by the formation and motion of positive and negative ions.

Corona discharges will appear as voltage increases in non-uniform electric fields. It is shown that atmospheric factors influence the corona critical voltage [7]. Mathematically, (1) is used to calculate the corona's critical electric field.

$$E_C = m * E_0 * \delta_0 * \left(1 + \frac{K}{\sqrt{\delta_0 * r_0}}\right) \tag{1}$$

where, E_C is the corona critical electric field in kV/mm, E_0 is the critical electric field strength measured at standard atmospheric conditions in kV/cm, K is fixed constant, m is the surface irregularity factor with conductor radius of r_0 and δ_0 is the relative air density factor. The estimation of the relative air density factor is made according to (2).

$$\delta_0 = \left(\frac{273 + t_0}{273 + t} + \frac{P}{P_0}\right) \tag{2}$$

where t is the ambient temp in degrees Celsius (with initial t_0 = 25 degrees Celsius) and P is the pressure of the ambient air (P_0 = 101.3 kPa).

The inception voltage for negative corona for the HVDC conditions has been simplified by using the Peeks empirical constants (E_0 and K) with their average value [7]. It is

represented by (3) and the updated air density factor is given by (4).

$$E_C = 30 * m * \delta_0 * \left(1 + \frac{0.308}{\sqrt{\delta_0 * r_0}}\right) \text{ kV/cm} \tag{3}$$

$$\delta_0 = \left(\frac{293p}{760 * (273 + t)}\right) \tag{4}$$

where, p is the prevailing atmospheric pressure in mm and t is the ambient temperature in degrees Celsius.

Based on the Peek formula for calculating the corona loss mathematically, it is expressed by using (5).

$$P_c = \frac{K}{\delta_0}(f + 25)\sqrt{\frac{r}{d}}(V_{ph} - V_0)^2 \times 10^{-5} \text{ kW/km/phase} \tag{5}$$

where, f is supply frequency Hz, K is constant value, r and d are radius and spacing between conductors in meters, V_0 is disruptive critical voltage (rms) per phase in kV, and V_{ph} is rms phase voltage in kV.

However, in the case of low corona losses, the Peek formula produces an unsuitable result. Therefore, In 1933, Peterson introduced a modified empirical equation that was applicable to both higher and lower losses [13]. It is expressed as (6).

$$P_c = \frac{0.0000337 f * e_n^2}{\left(\log_{10}\frac{s}{r}\right)^2}\phi \tag{6}$$

where, P_c is the corona power loss in kW, f is supply frequency Hz, s is the spacing between conductors in inches, r is the conductor radius in inches, e_n line to neutral voltage kV, ϕ is an empirically derived coefficient that is a function of the ratio $\frac{e_n}{e_0}$ (e_0 is the disruptive voltage in kV).

1.3 Experimental Studies for Corona Discharge Estimation

In this section, the results from the experimental study to estimate the corona discharge in various conditions are presented. In order to detect corona discharge, different methodologies have been proposed, with the most widely used method being the PD Detector, i.e., TechIMP PD Base II [40]. However, recent studies show that the corona discharge intensity may be detected using a digital camera instead of a PD detector measurement device [32]. This approach, using a DSLR camera, offers higher sensitivity. It has the features of a bulb mode that provides control over both the aperture and shutter speed. Photographs were taken for a long exposure of 60 sec with f/5.6 aperture, and ISO-800 sensitivity. Using a DSLR camera, long exposure time based photographs allow for sensitivity to be increased because they are based on a larger number of photons. The photographs allow for not only the detection but also the localization of the corona discharges. It should be operated in complete darkness in order to optimize its sensitivity. To visualize corona discharges using a high-voltage setup three different scenarios were used. Initially, the sensitivity of the DSLR camera and TechIMP PD detector were analysed for different supplies using needle-to-plane and sphere-to-plane configurations. In the second scenario, the intensity of discharges is calculated based on the image processing method for the sphere-to-plane configuration. In the last, the visual corona is detected and compared using the camera for old and new conductors.

The basic setup using the camera and PD detector is shown in Fig. 2, and the detection of corona using the HFCT via an oscilloscope and a camera, PD detector is shown in Fig. 3. It consists of three different supply voltages AC, -DC and +DC for the analysis. More information on the experimental setup and results are discussed in [1, 2].

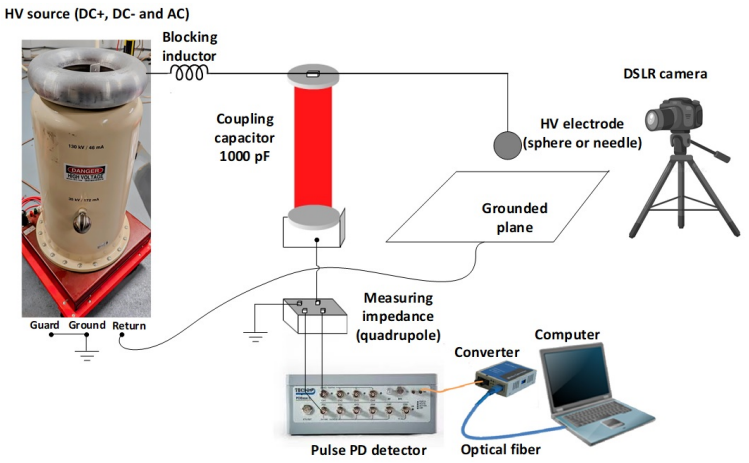


Figure 2: Experimental setup to test the sensitivity of the instruments [2].

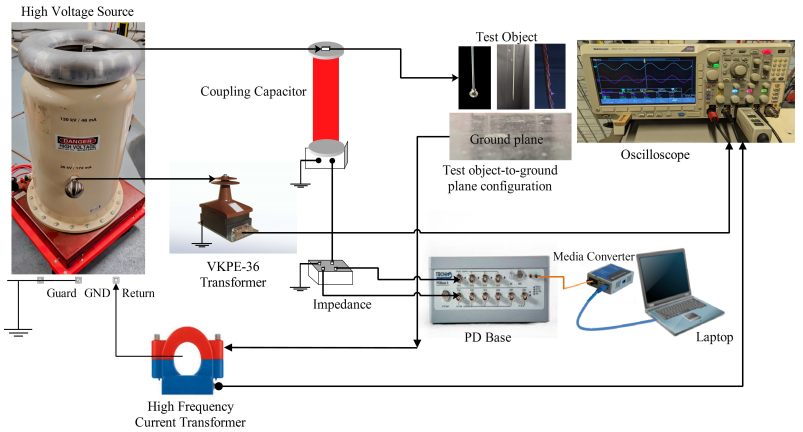


Figure 3: Experimental set-up to visualize corona discharge.

First case study, the analysis shows that the CIV determined by using a DSLR camera observed occurs earlier than the PD detector, as shown in the Table 3. In addition, note that the digital camera clearly detects the corona glow activity, while the PD detector is unable to detect any traces of PD activity under DC+ supply and corona glow mode, since the discharge is pulseless. For more information about the experimental results, please refer to [2].

In the second case study, the quantification of corona discharge for various spherical electrode test objects is discussed using the DSLR camera, PD Detector, and HFCT via oscilloscope reading. Here, the optic sensors are impervious to the background electromagnetic noise characteristic of power networks, offering a greater advantage over other sensing techniques. In addition, in the case of the glow phenomenon, the PD Detector

Table 3: CIV Observed with Electrical PD Detector Method and the Camera [1].

Electrodes	Source	CIV (kV peak)		
		Electrical PD detector	DSLR Camera	Difference
Needle-to-Plane (air gap = 11.5cm)	AC	10.2 kV	9.6 kV	-6.3 %
	DC-	10.5 kV	9.5 kV	-10.5 %
	DC+	11.0 kV	11.0 kV	0 %
Needle-to-Plane (air gap = 17.5cm)	AC	11.3 kV	10.2 kV	-10.8 %
	DC-	10.1 kV	9.5 kV	-6.3 %
	DC+	11.5 kV	11.3 kV	-1.8 %
Needle-to-Plane (air gap = 21.5)	AC	12.3 kV	10 kV	-23.1 %
	DC-	11.6 kV	7.2 kV	-61.1 %
	DC+	14.4 kV	11.7 kV	-23.0 %
Sphere-to-Plane (air gap = 8cm)	AC	38.2 kV	37.6 kV	-1.6 %
	DC-	38.0 kV	32.0 kV	-18.8 %
	DC+	45.0 kV	42.0 kV	-7.1 %
Sphere-to-Plane (air gap = 19.5cm)	AC	46.2 kV	45.7 kV	-1.1 %
	DC-	63.0 kV	42.3 kV	-48.9 %
	DC+	Pulseless glow	65.0 kV	-

and HFCT via the oscilloscope couldn't trace any discharge current reading. Therefore, using camera images, the image processing method is used to detect the corona depicted in Fig. 4.

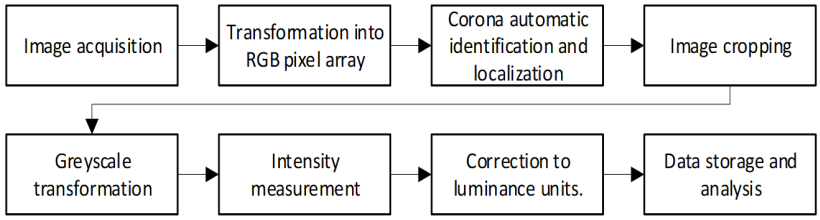


Figure 4: Flowchart of image processing method to detect corona discharge using high voltage setup.

The corona detected is bounded with limit m and n matrix that is cropped and stored. Further the image is converted into grayscale using (7) [41].

$$I_{Grayscale}(i,j) = 0.299 * R(i,j) + 0.587 * G(i,j) + 0.114 * B(i,j) \tag{7}$$

To calculate the intensity of the corona region using the image processing method with gray value pixel is expressed as (8).

$$Intensity\left(I_{Grayscale}\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} I_{Grayscale}(i,j) \tag{8}$$

The intensity is plotted against the peak voltage for different voltage supplies, as shown in Fig. 5. The intensity measured at CIV for AC, -DC, and +DC is nearly identical, showing

little difference. However, comparing the overall scenarios, the intensity is higher for DC-supply, while it has the lowest value under +DC supply. In the case of different sphere-plane geometries under AC supply the intensity is shown in Fig. 6.

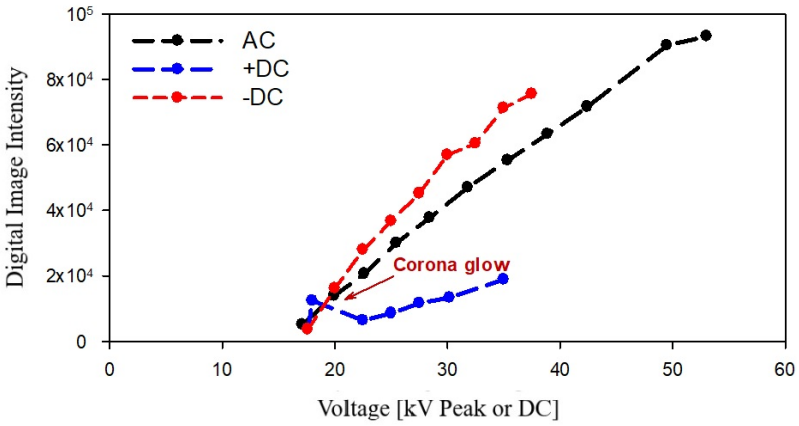


Figure 5: Intensity of the corona discharges calculated from the images taken with the DSLR camera for the sphere-plane geometry with the sphere placed 8 cm above the ground plane for different 3 mm spheres for various supplies [1].

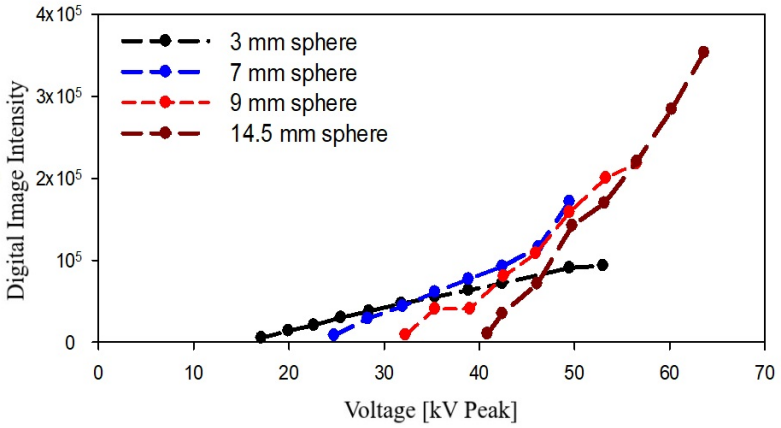


Figure 6: Intensity of the corona discharges calculated from the images taken with the DSLR camera for the sphere-plane geometry with the sphere placed 8 cm above the ground plane for different spheres under AC supply [1].

Third case study, an electric power system is ageing and needs upgrading; therefore, the comparison of visual corona is presented for aged rusted conductor and new conductor of strand diameter = 2.53,1.99 mm and number of strands = 10 using the same experimental setup as shown in Fig. 3. From the figures Fig. 7 and Fig. 8, it is clearly visible that the corona discharge is more visible in the case of the old conductor compared to the new conductor with a higher number of discharges at the same voltage level under the same environmental conditions.

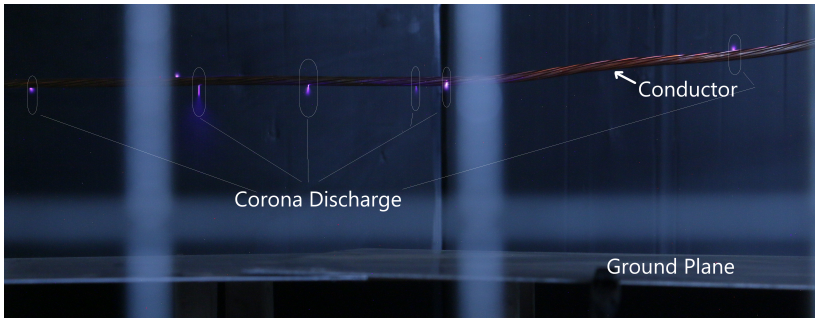


Figure 7: Aged conductor corona discharge at 50kV AC supply

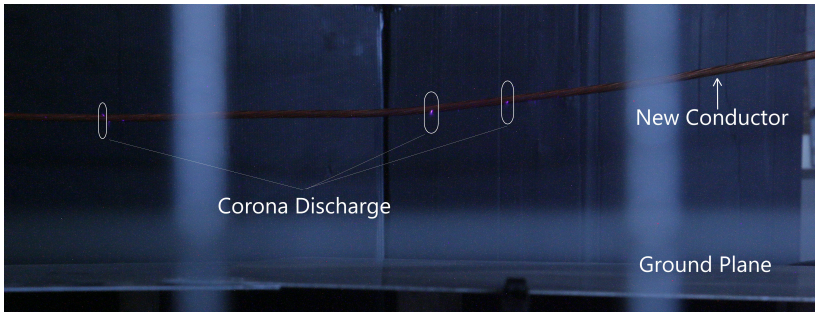


Figure 8: New conductor corona discharge at 50kV AC supply

1.4 Estimation of Corona Losses using Phasor Measurement Units

In recent years, the application of PMU and line measurement data has been widely used for various applications and to estimate corona losses in overhead transmission lines. In [42], the estimation of line parameters is discussed, which can be used for different loss calculations in an overhead transmission line. The literature [20] discusses the algorithm that calculates dependency between the corona losses and voltage based on PMU data, where the Corona losses per unit length under good weather, rain, light rain, and snow are discussed for real 500 kV lines.

The current focus has been on the development of modern techniques for measuring corona losses. In this system, use of PMUs play an important role to estimate transmission losses in the power network. The availability of PMU measurement data from both ends of the overheads transmission line is essential. Here, the PMU's data is used to calculate different losses by computing the differences between sending and receiving power. This data is further being used for the measurement of line parameters. A clear understanding of different types of losses such as Joule losses, corona losses, inductive losses, and capacitive losses can be established from adequate knowledge of line parameters.

In this research, the estimation of corona losses is based on the approach according to [43] and the systematic error is reduced using the correction method from [44]. In order to estimate corona losses for different weather conditions, PMU data from both ends of the overhead transmission line are needed. The difference between sending and receiving end power (i.e. losses) can be expressed as (9).

$$\underline{V}(L)\underline{I}^{*}(L) - \underline{V}(0)\underline{I}^{*}(0) = \Delta P + j\Delta Q \tag{9}$$

where $\underline{V}(L)$ and $\underline{I}(L)$ are the sending end voltage and current phasors, $\underline{V}(0)$ and $\underline{I}(0)$ are

the phasors quantities at receiving end, ΔP are total active losses, ΔQ are total reactive losses (or generation), and * denotes the complex conjugate.

The active transmission loss and reactive power balance can be represented by (10) and (11), integrating over the line length as follows:

$$\Delta P = r\int_0^L |\underline{I}(l)|^2\,dl + g\int_0^L |\underline{V}(l)|^2\,dl \tag{10}$$

$$\Delta Q = x\int_0^L |\underline{I}(l)|^2\,dl - b\int_0^L |\underline{V}(l)|^2\,dl \tag{11}$$

where $\underline{V}(l)$ and $\underline{I}(l)$ are the voltage and current phasor values at distance l along the line. Here, the losses dissipated in the line are resistive loss through resistance (r), reactive power consumed in reactance (x), losses in conductance (g), which is mainly due to corona or leakage current, and reactive power generation in susceptance (b) — all per unit distance values.

For the estimation itself, the line parameters above can be expressed from the distributed parameter transmission line equations and computed using (12) and (13).

$$\underline{\gamma} = \frac{1}{L} arcosh\left(\frac{\underline{V}(L)\,\underline{I}(L) + \underline{V}(0)\,\underline{I}(0)}{\underline{V}(0)\,\underline{I}(L) - \underline{V}(L)\,\underline{I}(0)}\right) \tag{12}$$

$$\underline{z}_c = \frac{\underline{V}(L) - \underline{V}(0)\cosh(\underline{\gamma}L)}{\underline{I}(0)\sinh(\underline{\gamma}L)} \tag{13}$$

where L is the length of the line (km) and per unit impedance and admittance are given by $\underline{z} = r + jx = \underline{\gamma}\underline{z}_c$ [Ω/km] and $\underline{y} = g + jb = \underline{\gamma}/\underline{z}_c$ [S/km]. For the purpose of this study, it is assumed that leakage losses in the g term are negligible and it represents corona losses.

In this study the data from synchronized phasor measurements was analysed. It included assessment of the propagation of various uncertainties to interpret the confidence interval for accuracy. The analysis also demonstrated how systematic measurement errors influence the fluctuation range of the monitored quantities. These results are statistically represented in a two-dimensional complex format that is valid for both magnitude and phase values [45].

1.5 Error Correction Method

The measurement error correction model was implemented using (14) to reduce the systematic error in corona loss measurement based on the proportionality linear fit of shunt parameters to the power flow [44]. At a time instant t, the measurement of corona loss based on measured values $y(t) = \left[\underline{V}_+(0), \underline{I}_+(0), \ldots\right]$ to establish a model is described by using (14).

$$T_C(t, y(t)) = P_C(t) + B(t, x(t)) + R(t) \tag{14}$$

where $P_C(t)$ is the actual power consumed by the corona discharge at t time instant, $R(t)$ is the measured power based on random measurement errors and $B(t, x(t))$ is defined as the measured power which depends on measured quantity $x(t)$ due to systematic measurement errors.

The expected value of measured corona loss is represented by (15) at time instant t, assuming the random error mean value is zero.

$$E\left[T_C(t, y(t))\right] = T_C(t, y(t)) = P_C(t) + B(x(t)) \tag{15}$$

Previous work on estimating components of transmission losses from synchronized phasor measurements has provided a simplified correction method to reduce some systematic errors in the estimates of corona losses [44]. In this method, corona losses from fair weather conditions are analysed in order to determine the dependence of the systematic errors on the power flow on the line. The correction is further supplemented by the average fair weather corona loss value calculated from the design parameters of the line [44].

The corrected corona losses value is determined using (16).

$$P_C(t) = T_C(t, y(t)) - B(P(t)) \tag{16}$$

where the systematic error was approximated by $B(P(t))$, which is defined based on a linear approximation model, is estimated by using (17).

$$B(P(t)) \cong D\,P(t) + C \tag{17}$$

where $P(t)$ is the power flow at time t and the constant values D and C are found for the given transmission line.

1.6 Intermediate Summary

The Chapter presents a brief introduction to the corona phenomenon and provides an overview on different methodologies to calculate corona losses, to measure the intensity of corona discharge in high-voltage equipment, and empirical methods. Results from experimental studies, which address corona discharge for various electrodes using a PD detector and a camera, are also presented. These studies analyse the sensitivity of the instruments and quantify the corona intensity. Additionally, the proposed methodology, which utilizes synchronized phasor measurement data to estimate corona losses, along with an error correction method, is discussed in detail.

2 Weather Dependency on Corona Losses

Corona losses are dependent on weather conditions. In this Chapter the influence of various weather parameters on corona losses is discussed. It is shown that corona losses are influenced by several parameters, including voltage, various weather conditions, air conductivity, conductor type, and design. The possible correlation is observed to better understand the impact of these factors at different temperature levels.

The results and discussion presented in this Chapter is published in III.

2.1 Introduction

Corona loss is a discharge phenomenon that occurs due to rise in ionisation of air across the conductors as a result of a strong electric field due to free electrons. Corona loss is correlated with the length of the path that the electricity is routed, the design of overhead transmission lines and complex weather conditions. To ensure optimal operation and minimize losses, a study on corona losses is very useful for the future to allow the maximum possible transmission capacity of electricity over long distances.

Due to drastic changes in climate, it is important to understand these weather dependencies better. However, It is very difficult to define a mathematical formula for corona losses where the accuracy rate is very high with the observed data given under various weather conditions. Numerous papers have discussed the effect of weather conditions on corona losses. Some of the earlier work for corona losses observing the weather conditions from 1911 onward are done by Peek's and Peterson's equation [7, 13, 14]. The dependence of corona loss on weather parameters such as temperature, pressure, and the air density correction factor (a function of air pressure and ambient temperature) was studied.

Various researchers have published many research results based on the influences of different weather conditions on corona losses, such as rainfall, icing, humidity and fair/sunny, etc. During the 1980s, the authors presented a conductor designs based reduction of corona losses with varying rain rates carried out in a test cage inside the anechoic room where the rain generator was installed [46]. In [18], the authors proposed the computation of corona losses using measured corona losses in a laboratory with a corona cage in case of heavy rain and sandy-dusty weather with varying altitude, also an annual sum of corona losses was determined for rainy, snowy, foggy, and fair weather conditions. The influences of instantaneous rainfall rate on corona losses, audible noise, and radio noise were expressed in [47], where the measured data were obtained from AEP-ASEA UHV Station, Indiana, which also shows the relationship between conductor configuration and age, where aged conductors seem more sensitive as compared to new conductors. In [48] and [49], a practical laboratory experiment was conducted, where the corona cage was used in a climate room to classify the impact of weather conditions, such as fair weather, rain, rime ice and glaze ice on corona losses. Another experiment analysed corona losses in rime ice and glaze ice conditions. The highest corona loss was seen in the case of glaze ice, the lowest in rime ice, and it was almost negligible in fair weather.

Despite the fact that these studies have shed light on the ways in which different weather variables affect the severity of corona losses, it is challenging to measure the correlations and variations among them. In [50], the authors gathered a larger and more varied set of measurement data from operational transmission lines over a longer period of time to analyse the impact of various weather parameters on corona losses such as conductor temperature, relative atmospheric humidity, dewpoint, wind speed, clouds, visibility and precipitation. However, it poses a new question, even if it offers more measurable

outcomes. The main influencing weather parameters are unclear, as are the strength and consistency of the dependencies that are shown.

In this Chapter, the effect of corona losses on each reported weather parameter is graphically analysed, facilitating the independent correlation between changes in corona losses and each of the weather parameters. The important scores and ranks of the various weather parameters are then determined using feature importance analysis. The following ranking strategies are proposed based on feature selection methods: Pearson, Kendall, Spearman, SelectKBest, boruta method, XGBR, MI technique and tree based methodology that offers valuable information about the influence of weather parameters on corona losses for designing future generation overhead transmission lines under complex weather conditions. This improves a comprehensive understanding of which weather factors have the greatest impact on corona losses. The efficacy of the system can be determined by monitoring the operational availability, reliability, assessment, and maintainability of the overhead transmission line.

Synchronized phasor measurements have been widely used to estimate corona loss in recent years [51, 52, 43] and various other applications. Knowledge of PMU data makes it easier to monitor corona losses over an extended period of time without requiring any new experimental equipment or sophisticated instruments. The present study leverages the increased deployment of synchronized phasor measurements across diverse regions of overhead transmission lines in Estonia, facilitated by the TSOs. This initiative aims to improve the estimation and monitoring of corona losses under varying and complex weather conditions. The methodology used to estimate corona loss and the error correction method are the same as those discussed in Chapter 1, using synchronized phasor measurement data. For the purpose of this analysis, two-year synchronized phasor measurement data and weather data are collected for monitoring purposes. Both hourly and 10-minute weather monitoring data were used. In order to ascertain whether more widely accessible hourly data are adequate for these kinds of study, the various time-steps in weather monitoring have been evaluated.

2.2 Methodology and Data Collection

The assessment of corona losses and weather dependency is made using synchronized phasor measurement based wide-area data from the Estonian power system. It uses the synchronized phasor measurement logged data, which are stored more frequently with the help of a Phasor Data Concentrator (PDC). In the study, the collected voltage and current phasor data was used. For the analysis, the voltage and current phasor data were resampled at 1 sample per second. This data is used for the calculation of line parameters using the two-port network and the line parameter equations referred from [43]. Further, using the line parameter values, the corona loss is estimated, where the total loss is equivalent to Joule losses, corona losses, inductive losses, and capacitive generation, respectively. To reduce systematic errors in measured corona loss, the correction method from synchronized phasor measurements units is applied [44]. Weather parameter data was obtained from the nearest weather station to evaluate its impact on corona loss, which was logged every ten minutes except for humidity data. To match the synchronized phasor measurement data and weather data rows, PMU data was converted into a 10-minute analysis considering the mean value, and the data set was merged, arranged by date and time. The data set time span is two years (2018-19).

The dataset contains a collection of weather parameter features considering the nearest weather station to observe the influence of weather parameters on corona losses. For this task, the Mutual Information (MI) [53], Extreme Gradient Boosting Regression

(XGBR) method [54], Tree-based classifier [55, 56], and Boruta package [57] are used to select the important features with their rank. Several researchers have utilized these feature-importance methods to filter out the most important variables or parameters [58, 59, 60, 61].

To test the Mutual Information (MI), two random variables were introduced, where one random variable obtained the information given to the other one. From [62], which defines the Mutual Information using the Venn diagram, it gives more information on how the two variables are related to each other. The levels of individual relevance are defined in [60] in terms of MI.

The XGBoost package was built to provide weighted classification among diverse variables and ranking functions, as well as user-defined goal functions. [63]. In [61], the authors proposed a method that incorporated a systematic feature selection approach to facilitate a more suitable feature subset to characterize the collected data. To rank the top important features based on importance score, different regression models were used, such as LR, XGBR, RFR, DTR, and SVR. XGBoost methodology was used to calculate the feature importance score in the UNSW-NB15 data to generate a reduced and optimal feature vector [64]. The ranking of weather parameters is evaluated based on gain and score values which range between zero and one.

The feature importance function in a Tree-Based Classifier is an internal class that assigns a score between 0 and 1 to each feature. Based on the score obtained, the most significant input weather parameter features are chosen; the higher the score, the more pertinent the feature is to influence the corona loss. This helps remove the less important features from the datasets. In our study, we have classified the corona loss as low ($P_c <$ 0.16 MW), medium (0.16 MW $\leq P_c \leq$ 0.19 MW) and high ($P_c >$ 0.19 MW) based on maximum and minimum percentile values. A value that is below the 25 percentile is considered as lower corona loss, above the 75 percentile is higher corona loss, and between the 25 and 75 percentile values is marked as medium corona loss. The unit of corona loss Pc is megawatt (MW). The Tree-Based Classifiers are used from the Scikit Learn libraries for feature selection techniques for different applications [55, 56].

The Boruta algorithm is a technique that ranks the significance of features or parameters by constructing an ensemble of equivalent artificially added "shadow" variables that are randomly selected from the dataset [65]. The computation of Boruta algorithm is done iteratively, where the Z-score is compared with each variable to the shadow variable and then feature importance is assessed [66]. In [67], the Boruta method and RF-RFE method were used to assess the feature importance of power system transient stability.

The feature importance scores are used to determine which input weather parameters influence corona loss in the overhead transmission lines. In addition, linear (Pearson's, SelectKbest) and non-linear (Spearman, Kendall) correlation methods are discussed. Fig. 9, shows the methodological framework for corona loss estimation and the feature selection process to select the important weather parameters.

2.3 Description of Overhead Transmission Line Configuration

The observed overhead transmission line was built in 1959 and runs from Narva to Tartu in eastern Estonia, primarily north to south. Fig. 10, depicts the approximately 170-kilometer-long line corridor and neighbouring weather stations along the line. Weather parameter data from four stations along the transmission line, including those near both ends, were studied to investigate weather changes along the line. The Narva station is around 5 kilometres from the transmission tower, Tartu is 20 kilometres away, and the remaining stations are within 15 kilometres. The transmission line covers relatively flat terrain with an

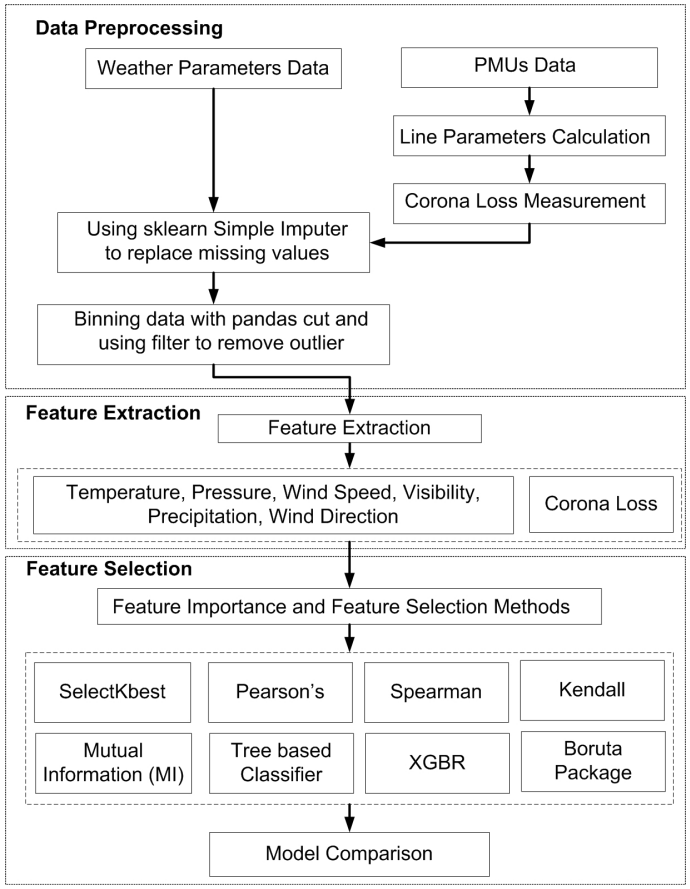


Figure 9: The framework of estimating corona loss and weather parameters selection using feature selection.

elevation difference of around 55 meters. Fig. 10 depicts the mapped 330 kV transmission line in blue, with weather station sites denoted by dots.

Fig. 11 depicts the probability density distributions of the chosen weather parameters such as—air pressure, air temperature, wind speed, and air density, using data from the Narva station, Tartu station, and along the transmission line (averaged weather data over many sites). The distributions are computed using the Kernel Density Estimate (KDE) from Python's Seaborn module [68], where no large proportional changes were observed. The most noticeable variation is in wind speed, which is most likely caused by coastal wind conditions at the Narva end vs the more inland Tartu end. In general, the overall variances between individual stations and the average are minor, with no significant large differences over a given period of time, indicating that weather conditions along the line length are typically uniform and consistent. As a result, the data used in the following analyses are from a single weather station. There are several ways to merge weather data from different sites, including linear and exponential combinations, complicated weighting, and geometric mean approaches, etc [69]. However, selecting the best representative approach for an entire area can be challenging. This research compares the simple averaging method for collected data from numerous stations against the data from a single station.

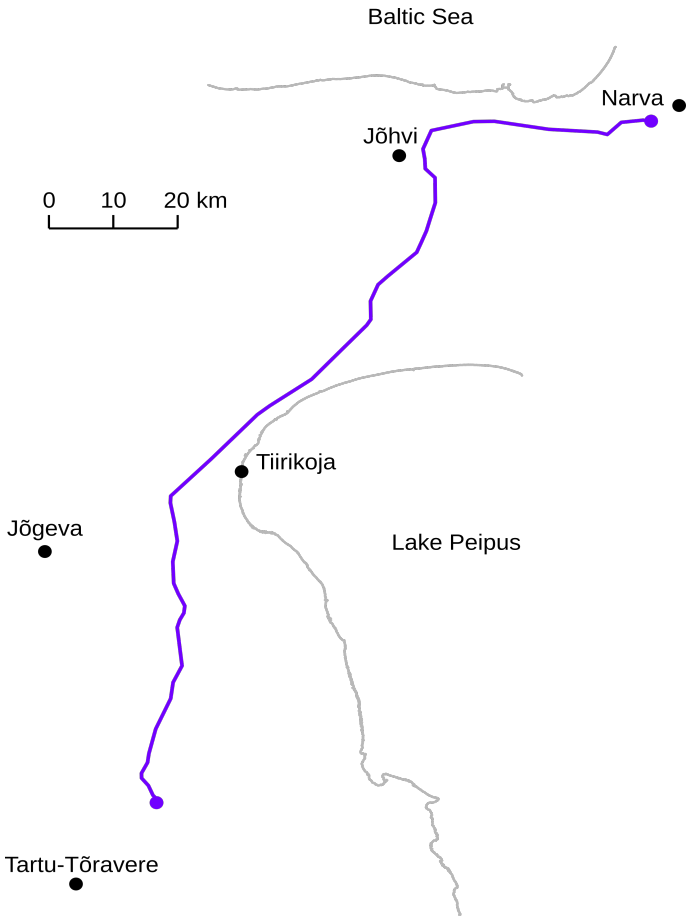


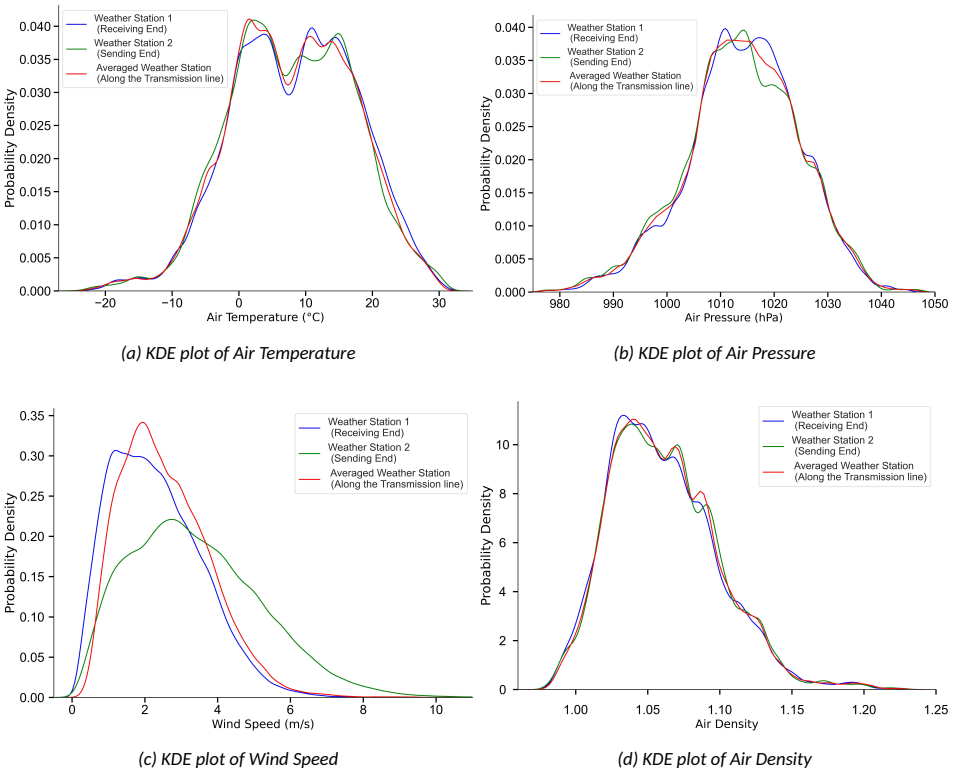
Figure 10: Map of North-Eastern Estonia depicting transmission line (solid blue line) and respective weather stations (black dots) [3].

Fig. 12 depicts the transmission line's typical tower configuration, where two bundle conductors are shown. For more information about the design specifications, the conductor type and transmission line's measuring impedance are shown in Table 4.

Over long periods of time, environmental conditions are likely to influence how corona losses occur on the line due to gradual changes in their surface properties. In this work, the effect of degradation is considered to be minimal, almost negligible and insignificant throughout the two-year monitoring period. However, the effect may need to be taken into account in other situations.

2.4 Influence of Weather Parameters on Corona Losses in Estonian 330 kV Transmission Line

The main objective of this research was to determine which weather factors have a greater impact on corona losses in high-voltage transmission lines. There are no ranking-based approaches for different weather characteristics in the literature that highlights how important it is to construct transmission and conductor lines. A 330 kV transmission line in Estonia was selected to monitor corona losses in the event of unfavourable weather



(a) KDE plot of Air Temperature

(b) KDE plot of Air Pressure

(c) KDE plot of Wind Speed

(d) KDE plot of Air Density

Figure 11: Probability density distribution of weather parameters from single stations and the average of multiple stations [3].

Table 4: Overhead transmission line design parameters and measured value of the line impedance [3].

Description	**Value**
Line Length	170 km
Conductor Type	GOST 839-59
Bundle Spacing	400 mm
Conductor Radius	14 mm
Conductors Per Bundle	2
Bundle Conductor GMR	66 mm
Positive Sequence Impedance	55.0∠83.3°Ω

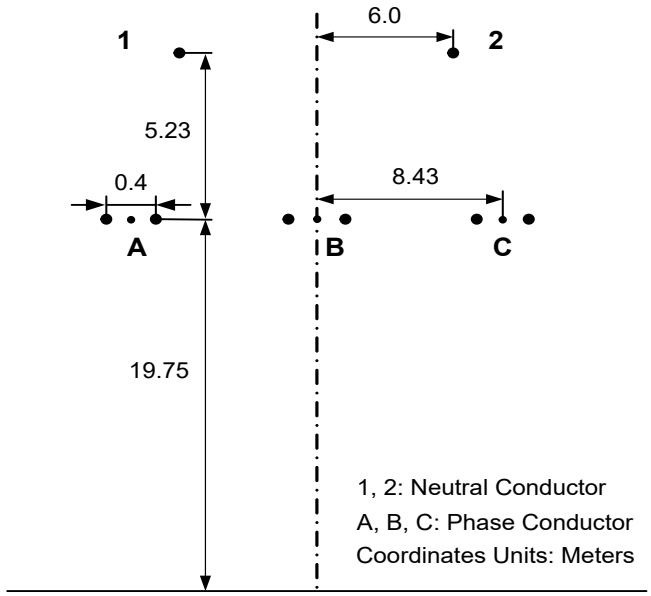


Figure 12: Tower configuration of the transmission line [3].

conditions.

The research shows that air temperature has a significant impact on corona losses. Therefore, all weather data are grouped together in the figures to demonstrate their combined effect with respect to different levels of air temperature. For the analysis purpose, four distinct air temperature levels, i.e., low ($T \leq -10.0\ ^{\circ}C$), medium-low ($-10.0\ ^{\circ}C < T \leq 2.8\ ^{\circ}C$), medium-high ($2.8\ ^{\circ}C < T \leq 15.6\ ^{\circ}C$) and high ($T > 15.6\ ^{\circ}C$), with roughly equal division taking into account the minimum and maximum values are discussed.

2.4.1 Air Temperature

A visualisation of the increasing or decreasing pattern of corona losses for varying air temperature is shown in Fig. 13. The maximum levels of corona losses occur mainly during the winter as a result of ice crystals and heavy snowfall on the conductor surface [70]. In order to comprehend the distribution of corona loss, we also looked at percentile levels. This study used the 95th percentile and 75th percentile as percentile levels to examine the impacts of air temperature at the normal distribution of corona loss. The corona loss intensity is mostly high and often initiated at a lower temperature, according to percentile data plots. It is evident that a decrease in air temperatures results in increasing corona losses. This is consistent with the widely held hypothesis that higher corona losses happen during the winter i.e., below ($T < -5\ ^{\circ}C$). However, if the load current maintains the conductor temperature higher, it may have an impact on preventing ice or snow accumulation on the conductor surface, which will result in lower corona loss.

2.4.2 Air Pressure

The dependence between atmospheric pressure and corona loss severity is depicted in Fig. 14, demonstrating a complex interdependence. It is evident that with increasing air pressure, the corona loss increases, reaching its peak between 1010-1020 hPa due to the decreasing temperature. Typically, when the sky is clear and there is high air pressure accompanied by low winter temperatures, it is reasonable to assume that higher air pres-

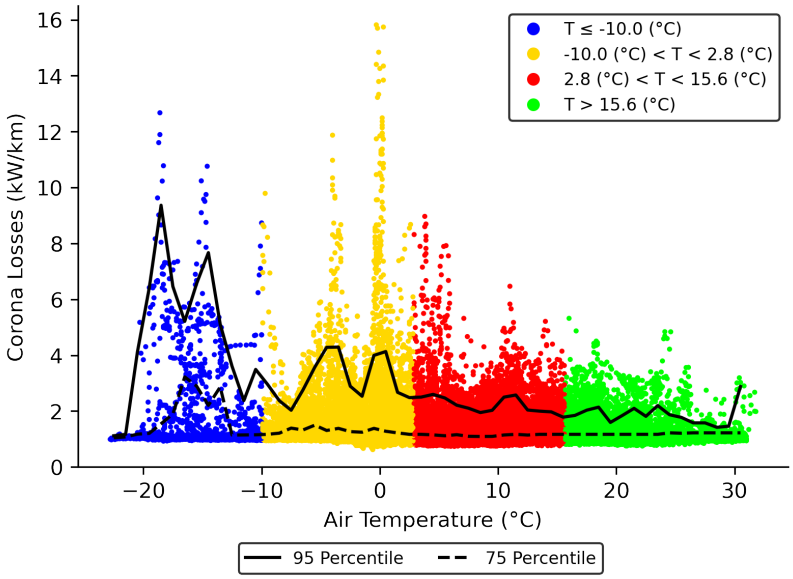


Figure 13: Dependence between varying air temperature and corona losses [3].

sure values will result in higher corona losses. Conversely, high air pressure in summer, results in smaller corona losses at warmer air temperatures. However, this is not always the case and it is difficult to say that it is a good indicator as the difference in lower and higher pressure values is not very large.

This suggests that corona losses can occur during times of fluctuating air pressure, which typically signals shifting weather patterns. In high-altitude regions, the air pressure significantly drops, which leads to an increase in audible noises that affect the corona loss. In our case, at lower pressures of 980-1000 hPa, an unpredicted increasing pattern of corona loss was seen, where it increases slowly and reaches a peak at 995 hPa and decreases further with a medium level of temperature following the Peek's hypothesis. The data's peaks are largely followed by the percentile lines. One incident of larger losses and a small number of data points are mostly responsible for the biggest peak in the low pressure values. A greater concentration of data points with low temperature values is correlated with the biggest peak in high air pressure values.

2.4.3 Humidity

Humidity refers to the amount of water vapor in the air. Numerous variables, including geographic location, air temperature, air pressure, precipitation and wind speed, etc., influence it. It is mainly categorized in two ways: absolute humidity and relative humidity. The total amount of water vapor in a specific volume is known as absolute humidity, and it is commonly stated as g/m3. The ratio of the absolute vapor pressure (AVP) to the saturation vapor pressure (SVP) in a specific volume is known as relative humidity [71]. The SVP refers to the point at which, due to ambient temperature, air becomes saturated with water vapor and condensation occurs.

The study of humidity shows a strong dependence positive correlation with corona losses. In our case, the hourly available humidity data is taken into account, where a linearly increasing pattern is observed. Under higher humid conditions at 75% or higher,

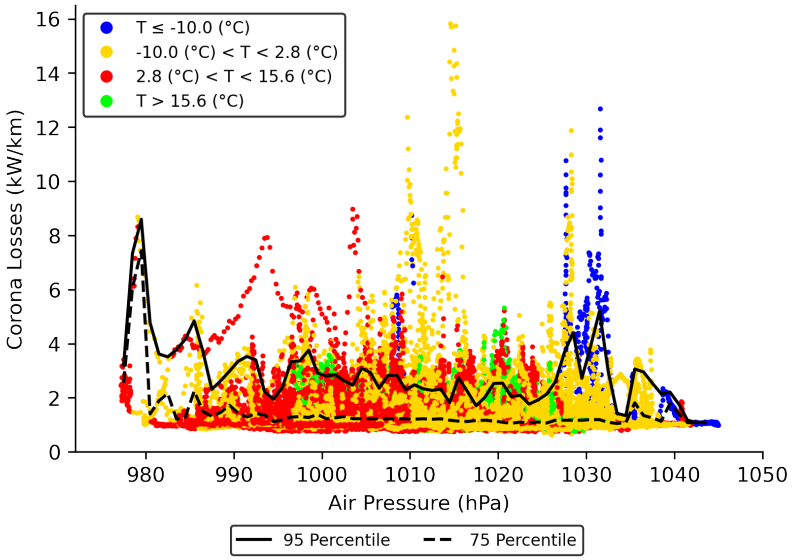


Figure 14: Dependence between varying air pressure and corona losses [3].

corona losses are at their highest, and the air temperature is found to be low, as shown in Fig. 15. The main hypothesis is that for lower temperatures, the correlations are high as expected due to inverse relation with humidity and hydration effects of ion and electron [72]. The corona current is directly related to the mobility of ions which is affected by hydrated ions at higher humidity. However, at lower humidity, the possibility of a higher corona occurrence is reduced.

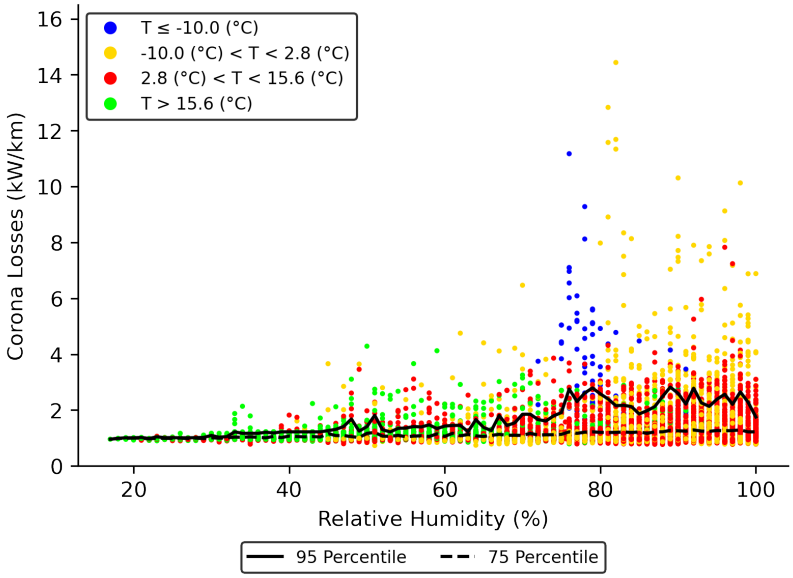


Figure 15: Dependence between varying relative humidity and corona losses [3].

2.4.4 Air Density

Fig. 16 presents the expected result of air density (δ) analysis to determine the influence on corona loss. The presented study will serve as a simplified view of how corona loss is affected by the air density where $\delta = \frac{3.92p}{273+T}$ (p is air pressure in cmHg and T temperature in $^{\circ}$C) [7].

As seen in Fig. 16, the dependency is highly complex, erratic, and rapidly varies with weather conditions. According to [7], δ is inversely proportional to corona loss, i.e., a lower value of δ causes a higher corona loss. In general, air density tends to decrease during inclement weather, such as rain, snow, and hailstorms. This increases corona loss during transmission due to lower critical disruptive voltage values.

However, in our case, it was seen that as the air density increases, corona losses increase due to snow or icing conditions at lower temperatures. This is due to the fact that the accumulation of snow on the conductor surface forms sharp edges of icing, which increases the electric field. The percentile values provide additional insight into the dominating character of the greater corona loss data division, despite the fact that this study demonstrates a strong association between them at higher air densities.

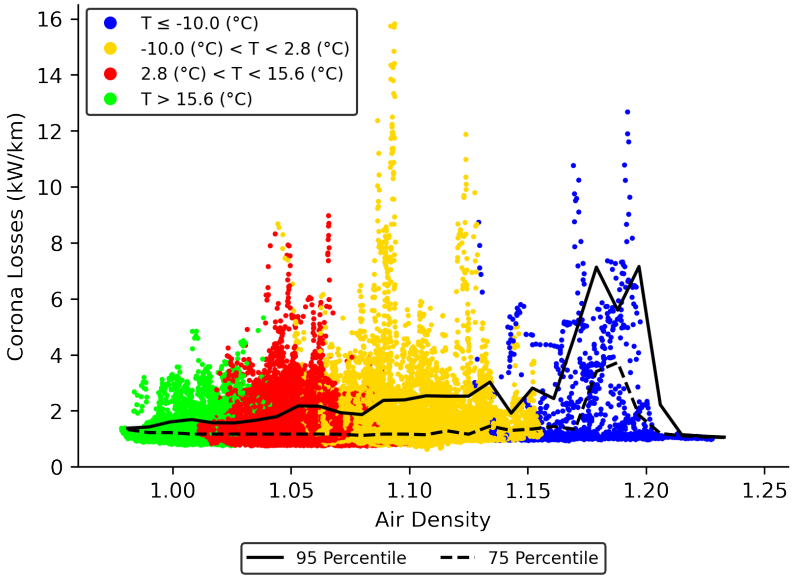


Figure 16: Dependence between varying air density and corona losses [3].

2.4.5 Precipitation

According to a preliminary investigation of the data as shown in the Fig. 17, the quantity and amplitude of the corona loss increase with decreasing precipitation values. Corona loss responds differently on the basis of climate changes and specific site weather conditions. Precipitation considering weak/strong rain or snow should also be discussed to understand the behavioural correlation pattern. However, in general, when it rains or there is high humidity, the water droplets accumulating on the line will form a cone structure due to the electrostatic effect; as a result, it causes a higher electric field gradient, leading to high corona loss.

The effects of precipitation on corona losses have been presented in Fig. 17. To illus-

trate the lower value of precipitation with the higher corona loss, where the majority of data points fall, a few data points with extremely high precipitation values have been left out. Although it seems that the highest corona losses happen when there is no precipitation, a considerable corona loss occurrence was observed for lower values. This is also consistent with visibility statistics, which showed that little precipitation would result in low visibility, and high visibility would not need precipitation. It should be mentioned that precipitation values are likewise low in low-temperature ranges. The peak at zero is comparable to the visibility peaks, meaning that while there are many data points with low corona losses, there are also many data points with larger corona losses. To illustrate, the percentile is drawn, which gives more information about how most of the high corona loss data points lie with low precipitation levels indicated by the 95th percentile line. Due to their reduced dependency, the percentile line does not show larger variations in corona losses.

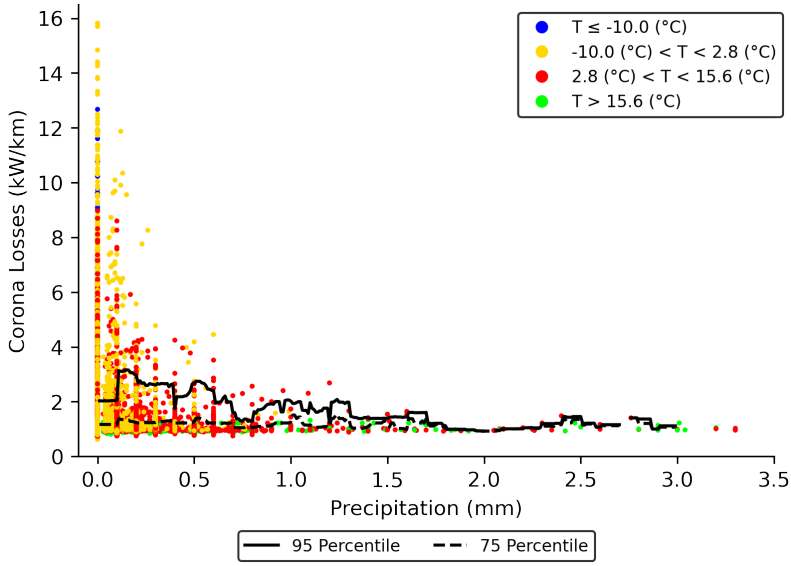


Figure 17: Dependence between precipitation and corona losses under varying air temperature [3].

2.4.6 Wind Speed

The correlation between corona loss and wind speed is depicted in Fig. 18. In general, a spark starts the conducting channel at lower wind speeds because of the instantaneous ionization of the air, which is sustained; at higher wind speeds – the corona discharge phenomena are disturbed because of ion mobility, resulting in reduced corona loss values. The other reason for a decreasing pattern of corona loss values at higher wind speeds is that the conductor surface is cleared.

It is also interesting to see that the lower air temperature corresponds to the low wind speed, indicating larger corona losses. In [73], authors discuss the ranges of wind speed and their impact in determining the current density profile that influences corona losses in the transmission line. In our case, we see that corona loss begins to decrease linearly as wind speed increases. This may also be seen as a decrease in peak current density. The most frequent explanation is that when the temperature rises, the snow melts or the conductor surface clears, which results in little corona losses.

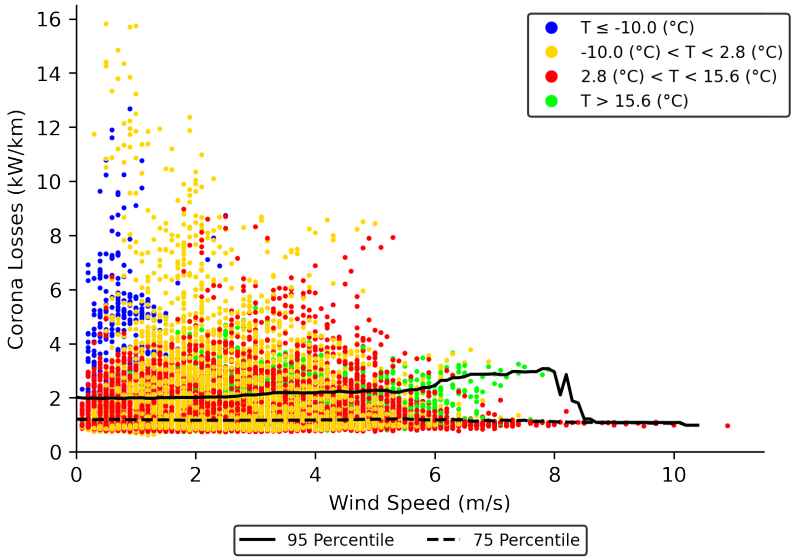


Figure 18: Dependence between varying wind speed and corona losses [3].

2.4.7 Wind Direction

The relationship between wind direction and corona losses is illustrated in Fig. 19. The broad variations between the effect of wind direction angle from the 0-360 degree is studied. The analysis shows that the correlation is weak, as a stable and uniform distribution of corona loss with varied angles was seen.

However, in [74], the authors demonstrate that the corona losses increase in proportion to the strength of the wind's perpendicular component in the case of bipolar DC. This means that there are several other weather factors that need to be considered while investigating the influence. To understand the distribution of corona loss, visual inspection of the resulting scatter plot with coordinates is plotted as shown in Fig. 19, where the distribution of data points lying within different quadrants is seen. Most of the higher corona loss lies in between 180-290 degree angle which means perpendicular or parallel to the line depending on line direction that corresponds to lower temperature values. In conclusion, the wind direction has a weak correlation, but depending on the other environmental conditions, the impact on corona loss will increase. It will help industries to enhance transmission line capabilities by better monitoring the measurement while taking into account a number of other factors.

2.4.8 Visibility

Fig. 20 illustrates the notable positive correlation between corona loss and visibility at different temperature levels. The study exhibits peak values of corona losses that occur mainly at lower visibility and with decreasing temperature. In winter, low visibility causes moisture and humidity to accumulate on the conductor, leading to increased corona losses.

Higher corona loss values below 20 kilometres are shown in the research as the maximum value point for a certain time period; however, this value was later changed to 34 kilometres, resulting in the peak seen in Fig. 20.

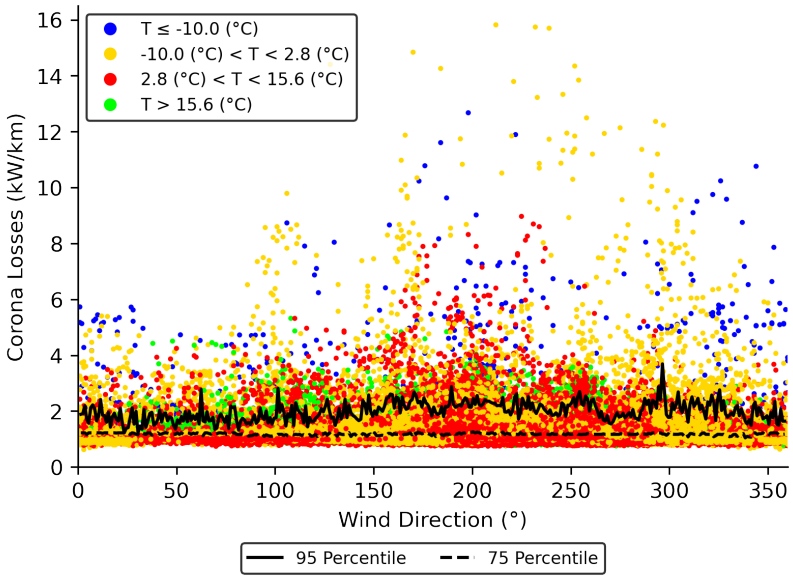


Figure 19: Dependence between varying wind direction and corona losses [3].

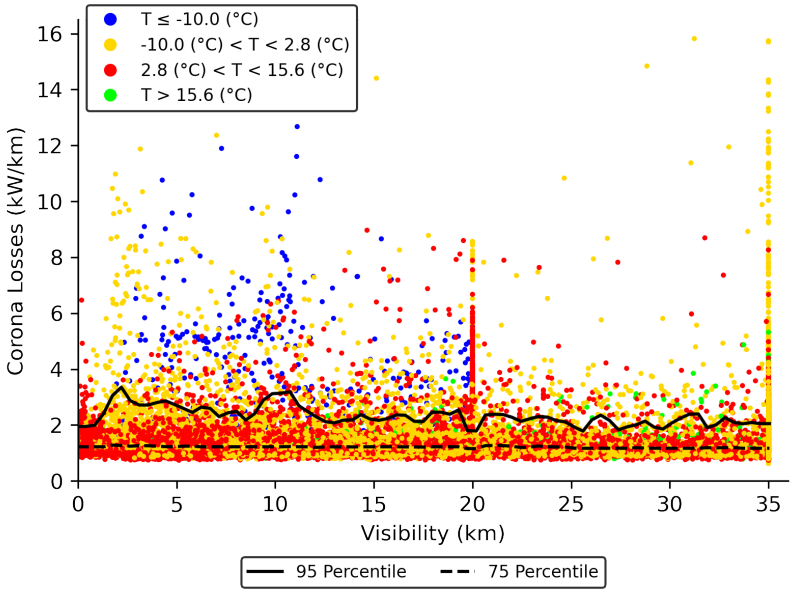


Figure 20: Dependence between visibility and corona losses [3].

2.4.9 Conductor Temperature

Here, a nonlinear heat-balance equation (IEEE Std 738-2012) that depends on a number of other factors is used to approximate the conductor temperature; such as air temperature, heat gain, resistance, wind speed, azimuth, electrical current amplitude, etc. (IEEE Standard, 2012).

To understand the effect of conductor temperature on corona loss, varying air temperature conditions are shown in Fig. 21. Overall, the findings are rather comparable, meaning that although conductor temperatures are higher in values when comparing air temperature as a result of loading, more corona losses are still occurring at lower temperatures. The percentile data lines for conductor temperature show behaviour with the biggest proportion of higher corona losses occurring in the lowest conductor temperature range and the majority of high corona loss values appearing during the winter. Nevertheless, lower temperatures have not necessarily resulted in higher corona losses, due to their complex relation and the loading effects.

The characteristics of the conductor temperature of a practical set-up using a corona cage and the experimentally measured results indicate a direct proportional relationship with the intensity of the loss of the corona, which was discussed in [75] where increasing the conductor temperature at a higher excitation voltage increases the intensity of the corona. However, In actual climate conditions several other factors are involved to understand the impact.

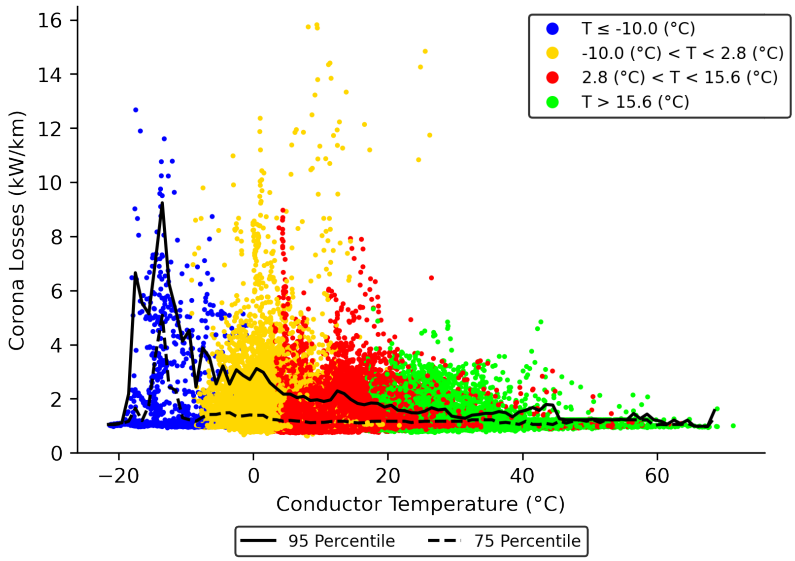


Figure 21: Dependence between conductor temperature and corona losses under varying air temperature [3].

2.4.10 Voltage

It is challenging to correlate the measured corona losses with the phase voltage derived from the PMU data. The distribution of voltage profile and corona losses is displayed in Fig. 22, which has monitored every 10 minutes of logged data for two years. In [70], the typical linearly increasing voltage corona loss dependency is shown validated with laboratory experiment results with gradual voltage variations.

Here, the impact of the altered phase voltage is far lower on corona losses, as the rate of change of the overall voltage variation is so small that the correlation pattern indicates that no abrupt relational alterations.

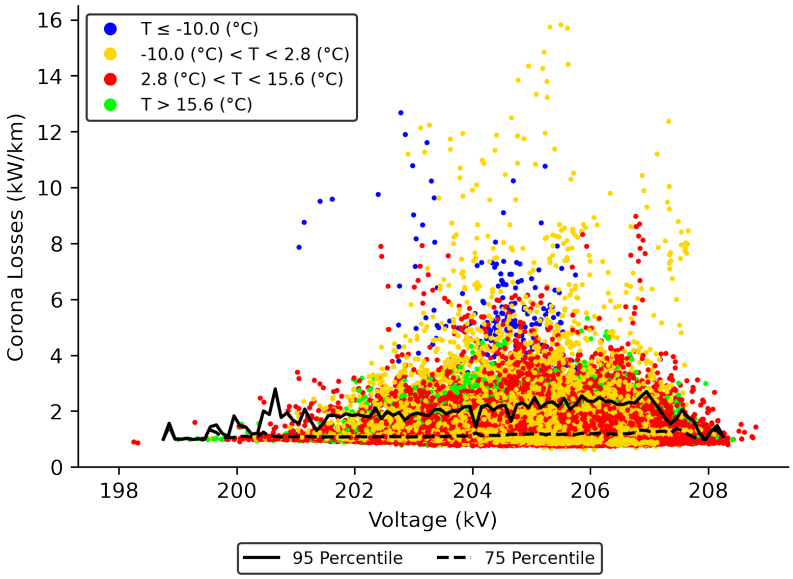


Figure 22: Dependence between voltage and corona losses [3].

2.4.11 Seasonal

Additionally, apart from the weather parameters that have been observed, the categorical weather conditions obtained from the automated sensors have also been studied. We examined the effects of seasonal weather on corona losses considering different atmospheric phenomena such as rain, snowfall, ice pellets, rain causing ice glaze, drizzle, mist/fog, snow showers, etc., where all are categorized into weak, moderate and strong phenomena. Fig. 23 shows the four major weather conditions: fog, rain, snow, and fair weather are considered, which signify atmospheric phenomena when dividing the corona losses under seasonal conditions.

It is evident that in the coldest temperature range, snowy conditions result in the higher corona losses; yet, both fair and snowy conditions may have medium losses. At a moderately low temperature, Conversely, fair weather tends to produce the most corona losses, whereas foggy, rainy, or snowy situations may result in moderate losses. Rainy weather seems to be the primary cause of increased corona loss levels in the medium-to-high temperature range. Overall, the outcomes align with the findings from the aforementioned examination of the various weather parameters.

2.5 Analysis of Weather Parameter Rankings Based on Feature Importance Techniques

To summarize the overall ranking of the weather parameters is challenging. For linear correlation mapping analysis, we used SelectKbest Method from scikit-learn library [59] and Pearson correlation [58] to determine how the correlation mapping fit with the weather parameters and corona losses. Initially, a heat map was used to observed if there is a patterned relationship between the weather parameters and corona losses. This was shown in Fig. 24. The relation depending on correlation coefficient and P-value based Pearsons method has been sorted as strong (close to +1 or -1) or weak association (close to 0). It was

seen that when using the Pearson correlation, the most important parameters affecting losses are pressure, temperature, air density, and conductor temperature.

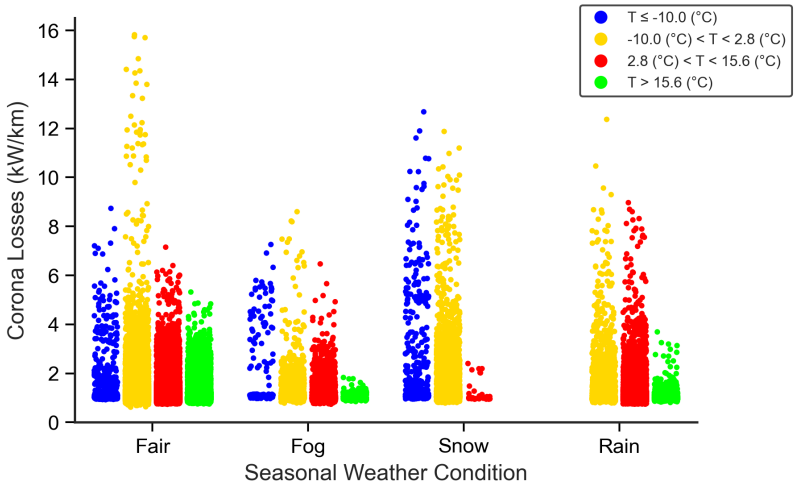


Figure 23: Dependence between different seasons and corona losses [3].

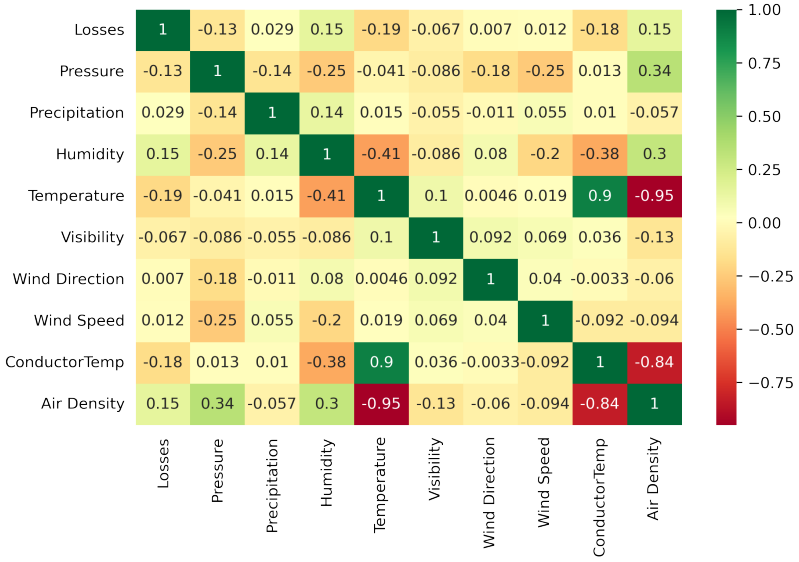


Figure 24: Dependence between weather parameters and corona losses using Pearson correlation.

Similarly, in case of the SelectKbest regression model, the most important weather parameter features are selected by varying the K value. It seems that air pressure, air temperature, visibility, and conductor temperature have a more correlated pattern. The solution obtained from SelectKbest and Pearsons may not be the optimal strategy here because our dataset should ideally reflect the real-life scenario.

In view of the importance of nonlinear correlation, the Spearman [76] and Kendall method [77] is used. In the case of both the Kendall and Spearman correlation the tem-

perature, visibility, wind direction, pressure and conductor temperature were highly negatively correlated whereas precipitation, wind speed, and air density seem to be positively correlated as shown in Fig. 25 and Fig. 26. Here, with the overall data visualization, we can figure out the major parameters that are influencing corona losses in the overhead transmission lines.

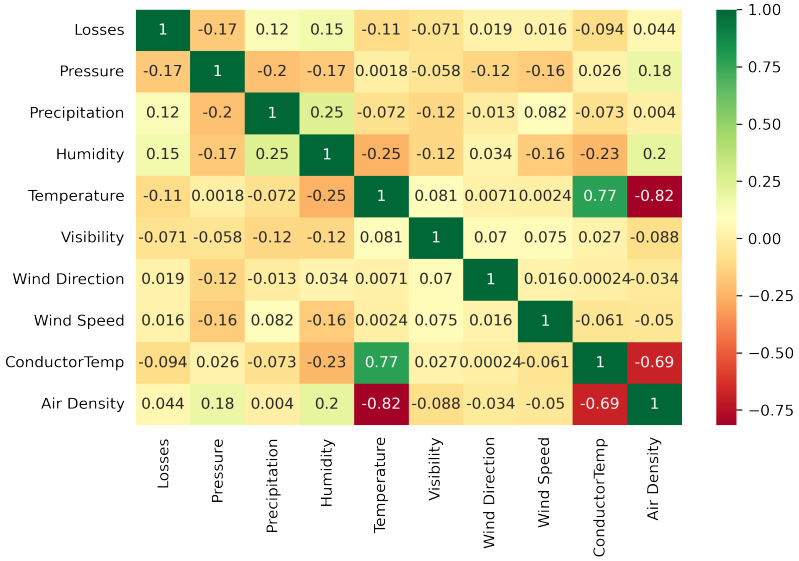


Figure 25: Dependence between weather parameters and corona losses using Kendall correlation.

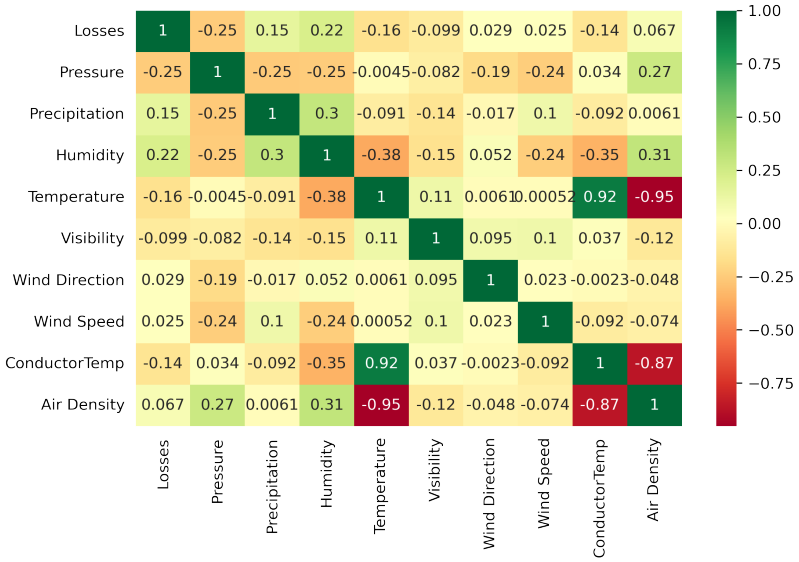


Figure 26: Dependence between weather parameters and corona losses using Spearman correlation.

To better identify the influence of the parameters, these are further evaluated based on ranking the importance score using different feature selection methods such as the

Boruta method, XGBR, MI, and the tree-based classifier as shown in Fig. 27. It can yield a wide range of results that depend on a variety of other variables, including data volume, types of data, importance weight, dependent variable, different machine learning algorithm types, etc.

Each weather parameter was given a ranking score; for the Tree-based classifier, XGBR (based on gain and weight of parameters), and the vector of all mutual information scores has been normalized so that the overall score is equal to 1, whereas Boruta displays the ranking of the parameters. Higher ranking scores are strongly correlated with the correlation status between weather parameters and corona losses. The analysis was done on both hourly and 10-minute data. For comparison, the association between corona losses and different weather parameters, linear regression analysis is also performed. The linear equation Y = MX + C, where X is the weather parameter and Y is corona loss, describes their relationship. While evaluating results, the R-squared value is utilized, and a greater number indicates a better correlation.

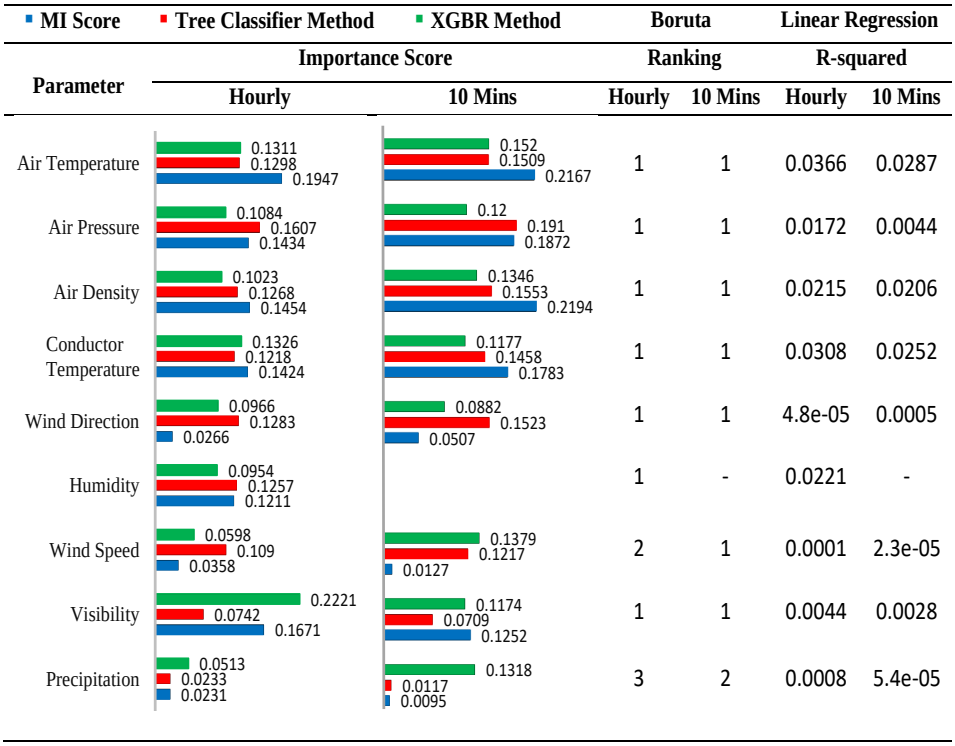


Figure 27: Weather parameters importance score and ranking using feature importance methods for 10 mins and hourly data 2018-19 [3].

The results show that air temperature, air pressure, and humidity are the measured weather characteristics that have the biggest effects. Similar strong influences are also seen in air density and conductor temperature, but these parameters are derived from measured quantities rather than being directly measured; for example, air density combines temperature and pressure, whereas conductor temperature is estimated from a

greater number of measured quantities following the IEEE standard 738-2012 [78]. While wind direction, speed, and visibility exhibit differing degrees of significance depending on the method and time-step in the data, precipitation seems to have the least impact. It can be inferred from the R-squared value indicates that simple linear dependencies are extremely weak.

The purpose of this analysis was to evaluate the relative significance of weather parameters by evaluating the importance scores. It can be inferred from the linear regression analysis of the 10-minute and hourly data that, in spite of varying R-squared values, the weather parameters are ranked similarly in terms of importance. According to the investigation, there is a complex relationship between corona losses and weather parameters, with air temperature, air pressure, and relative humidity having the greatest influence.

2.6 Comparison of 10 Minute and Hourly Data

The study collected weather data at both 10-minute and hour-long time intervals to see if greater resolution in weather data revealed additional specifics relationship between corona losses and weather parameters. Weather data collected in hourly time intervals is typically more easily available, making comparable analysis easier to do. To demonstrate the discrepancies, Fig. 28 shows hourly data for air temperature.

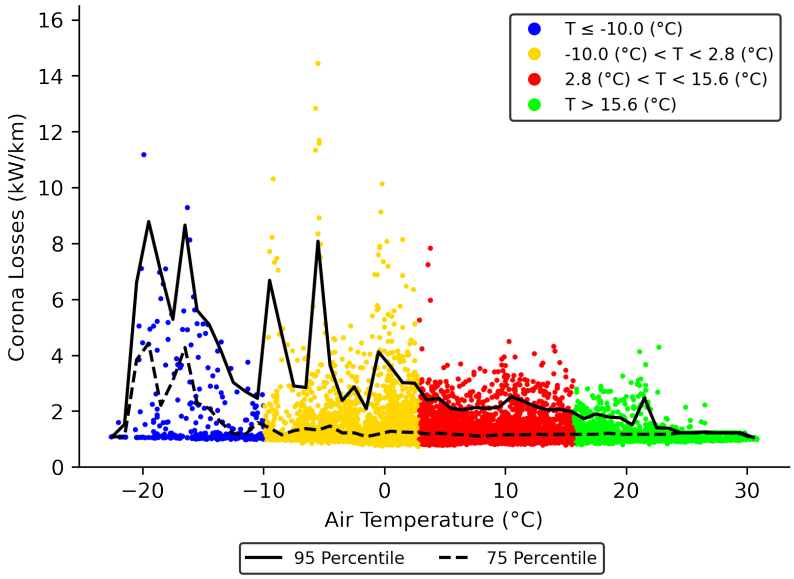


Figure 28: Dependence between air temperature and corona losses, hourly data [3].

It can be seen that in the case of findings from data with varying temporal resolutions, the statistical graphs are greatly comparable for hourly and 10-minute data with smaller differences. Although there are more data points in 10-minute data, the temperature ranges with the highest corona losses are almost the same and consistent. In the case of larger corona losses, the finer resolution data displays some additional detail, but hourly data is adequate for long-term investigations; although further detail can be useful for more thorough assessments, the overall reliance remains constant and unchanged given that weather conditions rarely alter several times within an hour. In addition, the analysis of features importance reveals that the 10-minute and hourly data yielded similar results

in most of the cases.

Using hourly data has the primary advantage of being more accessible. However, using the 10-minute data set, more data points identified dependencies between weather conditions and corona loss more precisely, demonstrating irregularities with sudden higher peaks. These peaks occur within very short time periods and are missed by hourly data. Nonetheless, the results analysis shows that using hourly data is sufficient to comprehend the dependence. Also, in some regions, due to the scarcity of finer-resolution data, the hourly data would be very beneficial with less computational effort.

2.7 Intermediate Summary

Understanding network loss percentage is important for present-day power systems due to financial implications. It is important to have an in-depth understanding of the weather parameters associated with corona losses, and how they are influenced. The Chapter presents statistical data analysis of collected weather and PMU data that has shown the real-life unpredicted correlation scenario in the analysis. Weather monitoring data collected over a two-year period and real network PMU data from the Estonian 330 kV transmission network were utilized in the investigation. The pair-wise Statistical graph plots, feature importance and linear regression are used to highlight the correlation between different weather parameters with corona losses under various temperature conditions.

This research provides a more thorough understanding of how weather conditions affect corona loss in high voltage overhead transmission lines. The investigation indicates that air temperature, pressure, and relative humidity are highly correlated with corona loss, however, the underlying dependency is complex. Various methodologies yielded comparable rankings for the most significant weather parameters.

To gain a deeper understanding of the consequences and variations, additional research on ageing transmission lines would be beneficial, which will be discussed in Chapter 3. The findings in this work will be applied to future studies in order to enhance the accuracy of corona loss predictions.

3 Performance of Corona Loss on Ageing Transmission Lines

It is of interest to understand the corona losses on aged transmission lines. This item is discussed in this Chapter with an example of two transmission lines in the Estonian power system with an age difference of 50 years. The analysis includes a comparison of tower design parameters and monthly corona loss patterns, as well as the dependence of corona on temperature over a short time period, and the dependence of weather parameters on corona loss for both lines using two years of data collected from the Estonian TSO and the Estonian Environmental Agency.

The findings and discussions presented in this Chapter are published in IV and V.

3.1 Background

The corona effect occurs on all types of overhead transmission lines, although it becomes more noticeable at high voltages. Many high-voltage transmission lines are constructed with sophisticated arrangements that increase the effectiveness of power transfer in response to increasing power demand. In today's world, losses in power system components are mainly caused by a number of things, including changes in the operating environment and conditions, specified useful life ageing, preventative and corrective maintenance procedures, etc. To reduce corona losses in overhead transmission lines, various researchers have proposed several studies and methodologies using bundled conductors, corona rings, increasing conductor diameter, removing sharp edges, etc. It is of interest to understand the main factors that have an impact on overhead line performance and the degree to which age should be taken into account, because operational losses in transmission lines are influenced by the fluctuating rate of deterioration.

The difficulty of measuring corona loss performance varies with weather conditions, and their detailed relationship is complex and should be verified over a longer period of time. The durability of insulating materials on overhead transmission lines is dependent on their ability to tolerate the damaging effects of environmental factors, including high wind, accumulation of dust, corrosion, fast temperature changes, humidity, and the emission of hazardous gases, etc [79]. The main factors that influence corona loss are complex weather conditions, conductor surface conditions, high voltage, conductor radius, and distance between conductors, which are some of the significant variables that alter the strength of corona discharge.

Joule (or resistive) and corona losses are the two primary categories of active power losses. Although corona losses are complex to measure and heavily influenced by weather, resistive losses are predictable and may be determined based on the line loading. Numerous earlier studies have examined the impact of weather variables or environmental conditions on corona loss using experiments based on corona cages and artificial climate chambers due to the limitations of the empirical technique. Some of the experiments are performed in an indoor and outdoor fields setting with a high voltage experimental setup and a climate chamber with a corona cage for different weather conditions such as icing [48], hoar frost [21], dusty and sandy [80], fog [81], and rain [82]. Another factor that influences corona loss is the state of the conductor surface and its roughness, which deteriorates and varies over time. Peek introduces the roughness factor formula, represented by m, to estimate the conductor's degree of roughness [83]. There have been discussions of changes in the roughness factor for new and old transmission lines in [84] and the corona discharge influenced by the conductor cross-sectional difference in [85].

Research on weather factors, corona loss assessment has already been done to examine, comprehend, and explain the pertinent relationships [50, 86]. However, It fails to rank

their importance taking into account the ageing issue. Some recent studies have used statistical analysis and prioritisation to infer the influence of different weather parameters on corona losses [3]. Nevertheless, in order to comprehend the intricate relationship, further research is required to examine the weather dependence of ageing transmission lines. Some of the studies in which corona loss performance was evaluated in the context of used and unused conductors were addressed in [87]. In [88], the case of aged sand-cast substation connectors was investigated and presented for corona loss assessment.

In this Chapter, using actual network PMU data, the effects of various design parameters, tower configurations, and weather parameters on the corona loss for aged overhead transmission lines are discussed. As of now, no studies have used actual network operational measurements to thoroughly compare and assess the corona loss performance of transmission lines of varying ages under actual operating settings. In this work, we examine seasonal corona loss performance, the weekly influence of air temperature and conductor temperature (estimated using a nonlinear heat balance equation from IEEE Std 738-2012) in a different season of the year, and the possible relation between corona loss and weather parameters such as air temperature, pressure, wind speed, and conductor temperature, etc, using percentile distribution and statistical graphical analysis for two years of data. The findings enable a deeper understanding of how corona losses are influenced in aged transmission lines. Moreover, the author hopes that this study will assist TSOs in monitoring and predicting corona losses, particularly in categorizing estimated corona losses for different seasons.

3.2 Description of Overhead Transmission Line Configuration and Design Parameters

The design and configuration of the overhead transmission lines are essential and should be checked and improved constantly in order to comply with the new standard. Climatic changes can deteriorate transmission lines, and various parameter constraints may not meet the required standards, which can lead to increased losses in the transmission lines. In this section, we present the results of two studies conducted to evaluate the performance of corona loss. The case study is taken from two different overhead transmission lines in Estonia. The new overhead transmission line (new conductor) was built in 2006 and the old transmission line (old conductor) was constructed in 1959 with a different configuration. Both the line location starting points are from north-east and are directed toward the west and south. The new transmission line, which lies in the northern region of Estonia, is spread over east to west with 210 km of line length whereas the old overhead transmission line bridges the north-south region of Estonia with a line length of 170 km.

3.2.1 Tower Configuration:

The tower configuration for both transmission lines is shown in Fig. 29 and Fig. 30, where the units of coordinates are measured in meters (m). To reduce corona losses, bundle conductors are used; the new transmission line consists of three conductors per bundle, whereas the old line has two conductors per bundle. The primary purpose of having more conductors in a bundle is to reduce corona losses. The height of the new transmission tower is higher compared to the old transmission line. GOST 839-59 and ACSR-402 are the conductor types for the old and new lines, respectively.

3.2.2 GMD, GMR and Bundled Conductor:

High-voltage transmission lines corona discharge is significantly influenced by the geometric mean distance (GMD), geometric mean radius (GMR), and bundled conductors. To

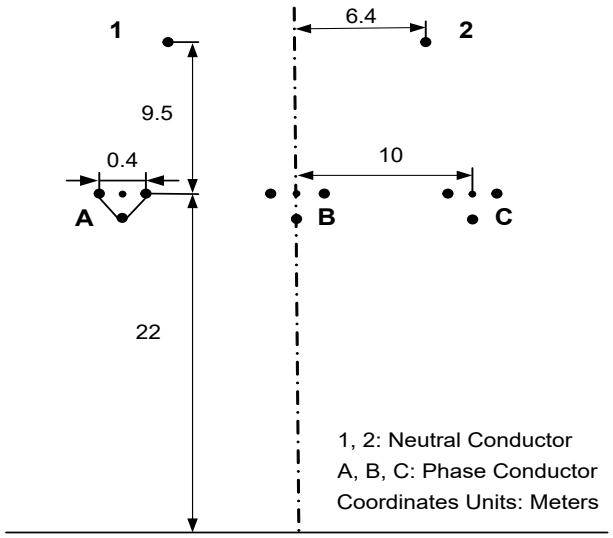


Figure 29: Tower configuration of the new line [4].

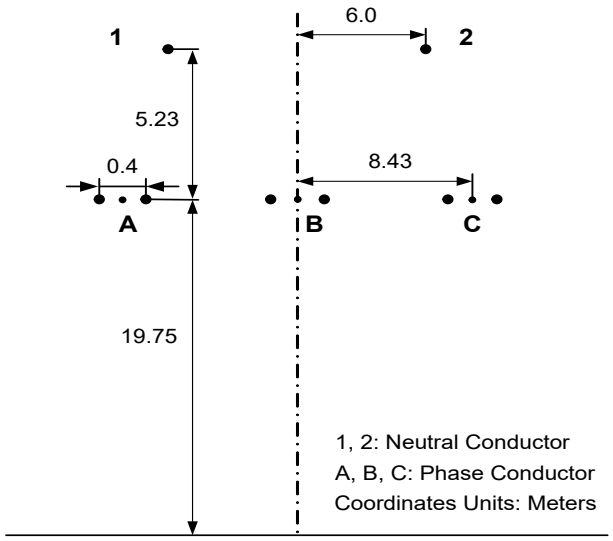


Figure 30: Tower configuration of the old line [4].

analyse corona loss, different values of GMD, GMR, and conductor count per bundle were evaluated in [89]. Generally speaking, when conductors are bundled, their own GMD rises, i.e. the effective distance between conductors, which reduces the electric field intensity because of an increase in surface area and raises the critical disruptive voltage, thus lowering the likelihood of corona discharge. Therefore, employing bundled configurations rather than isolated configurations reduces corona loss.

In the case of a conductor effective radius, or GMR, the electric field distribution around is affected; higher GMR values result in a more uniform field, which lessens corona effects. Therefore, a large conductor radius reduces the intensity of the corona loss, and a larger surface area results in a lower intensity of the surface field. The effective radius is greater in the case of bundled conductors, which results in lower corona losses. Table 5, shows the radius of the conductor, conductor type, bundle spacing, and number of conductors per bundle for both transmission lines.

Table 5: Design Parameters of Transmission line [4].

Description	New Transmission Line	Old Transmission Line
Line Length	210 (km)	170 (km)
Bundle Conductor GMR	12 (cm)	6.60 (cm)
Conductor Type	ACSR-402	GOST 839-59
Conductor Radius	13.8 (mm)	14.0 (mm)
Bundle Spacing	40 (cm)	40 (cm)
Conductors Per Bundle	3	2

3.2.3 Conductor Spacing:

Conductor spacing is one of the major factors that influences corona discharge in high-voltage transmission lines. Ionization of air is likely to happen when the conductor is spaced closely as it increases the electric field strength. To reduce the electric field intensity, the spacing between conductors is increased, or bundled conductors are used to prevent corona discharge, improving the power transmission efficiency. As seen in Fig. 29 and Fig. 30, the sub-conductor spacing of both transmission lines is the same in our situation; however, the new transmission line horizontal conductor spacing is 10 m, while the old transmission line is 8.43 m.

3.3 Seasonal Behaviour of Corona Loss Performance

The purpose of the present study is to identify and monitor corona loss for ageing transmission lines. To allow for more insightful interpretations of the corona loss pattern, it may be useful to understand how the monthly trend of corona loss magnitude transitions vary from month to month in order to assess the conditions. In our case, the estimation of the corona loss for both transmission lines using PMU data is referred from [43] and, in addition, the system correction method was used to reduce systematic errors in the measurement [44]. Further details on the estimation method are provided in Chapter 1. Given that the two transmission lines are located in somewhat different directions and have seen slightly different weather changes, it is concluded that both transmission lines were exposed to similar weather conditions based on data from the nearest weather station and geographical latitudes and longitudes.

In this work, the corona loss magnitude analysis is examined for over almost two years

(2018–19) using PMU data from new and old transmission lines. It is categorized into monthly patterns under varying temperature ranges, as illustrated in Fig. 31 and Fig. 32. It is evident that greater corona loss occurs throughout the winter months, particularly in December, February, and March, except January, where there is not enough data available, as the analysis was conducted from February 2018 onward. Here, the old transmission line experiences higher corona loss values due to environmental changes and a rapidly changing rate of deterioration and design configuration factors. It is clear from the figures that lower temperatures reflect higher corona losses, and as summer approaches, the corona losses are significantly reduced. This helps in determining the trend shift in the magnitude of monthly corona losses.

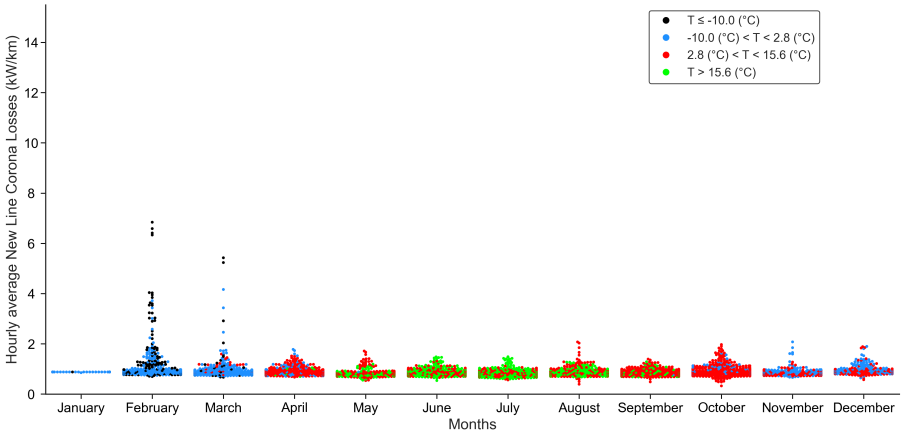


Figure 31: Monthly comparison of hourly corona loss for New line [4].

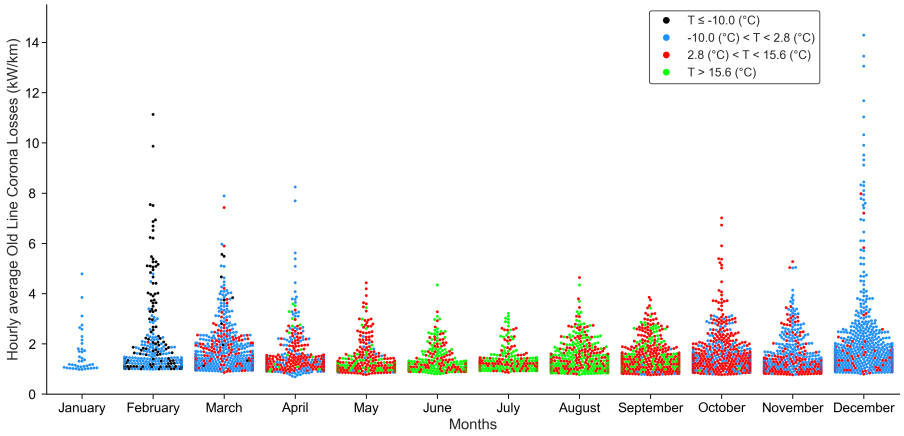


Figure 32: Monthly comparison of hourly corona loss for Old line [4].

The findings of the study, shown in Fig. 31 and Fig. 32, provide evidence that older transmission lines have higher corona losses than new ones throughout a range of time cycles in a year. This is because of the ability of the power system network's overhead transmission lines to manage growing loads under inconsistent weather conditions, which are impacted by the ageing process and changing surface conditions. The graph plot shows

the maximum and minimum corona loss magnitude changes over a given period of time.

3.4 Weekly Comparison of Corona Losses

In this section, the weekly trend of the magnitude of corona losses for the months of February and November are presented as shown in Fig. 33 and Fig. 35, respectively, to provide a better understanding of corona losses at a smaller time step using 10-minute time-step PMU data for both the transmission lines. From the figures, it is clearly visible that no matter whether it is winter, summer, or any other season, the corona loss magnitude is always higher for older transmission lines compared to new transmission lines. Nevertheless, it should be considered that dependence on various factors might vary and the specifics depend on the line loading and environmental conditions.

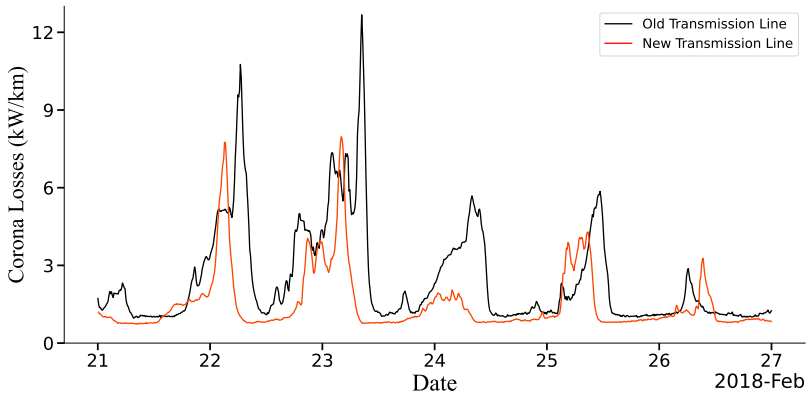


Figure 33: Comparison of corona loss over the span of one week [4].

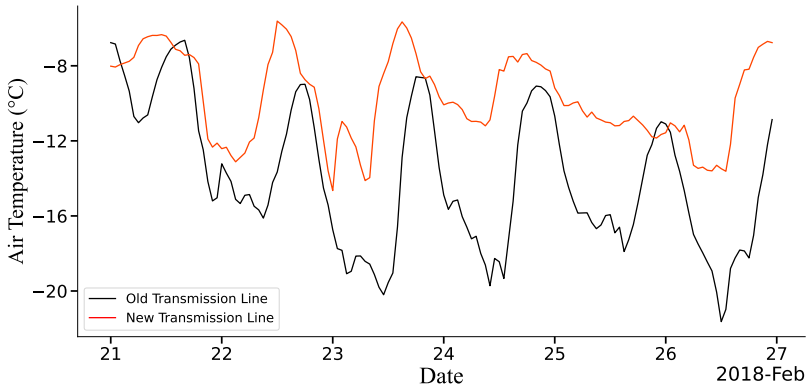


Figure 34: Comparison of air temperature over the span of one week [4].

The air temperature variations from the same time period, i.e, February and November, are shown in Fig. 34 and Fig. 36, respectively, and are included for reference in order to show how temperature and the magnitude of corona losses are related. In conventional methods such as Peek, the empirical formula for corona loss calculation is inversely related to the air density correction factor, which is a function of temperature and pressure. The temperature data shown in our case study lies mostly below -5 °C and has plummeted

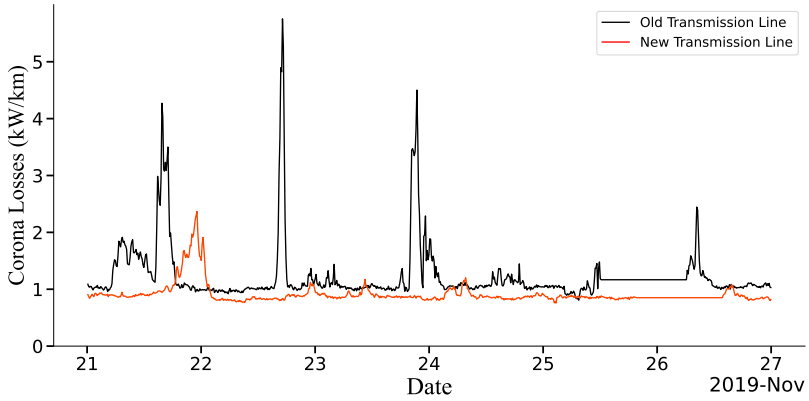


Figure 35: Comparison of corona loss over the span of one week [4].

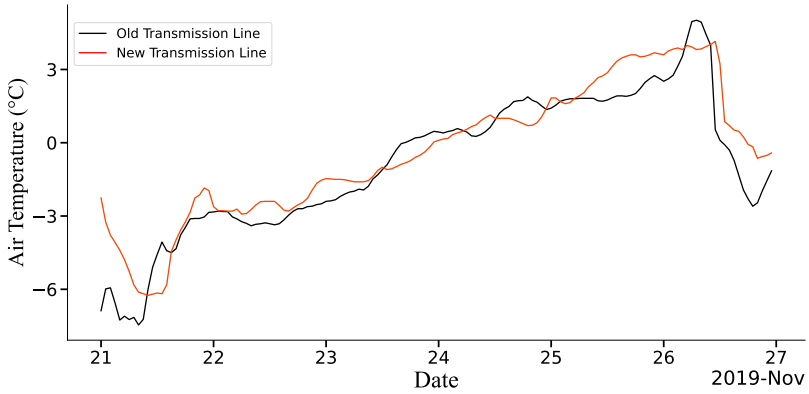


Figure 36: Comparison of air temperature over the span of one week [4].

as low as -22.5 °C, which is in line with the noticeably higher magnitude of corona loss in the month of February, i.e., winter. In November, a similar finding was made, i.e., where high corona loss is less likely to occur when temperatures rise. However, the focus of the present study was to examine whether the magnitude of corona loss is higher for older lines or new lines, keeping the fact that the temperature relationship shows that even though it has a similar pattern with almost the same values in November, the old line has a larger corona loss. This shows how ageing and other factors also influence corona loss. Similarly, the influence of the conductor temperature is presented in more detail in the next subsection.

3.5 Influence of Conductor Temperature on Corona Losses

According to earlier research, the magnitude of corona loss or corona discharge is significantly influenced by the geometry, material, size, and surface conditions of the conductor [84]. The temperature of the conductor is the major factor that affects its physical properties and surface conditions, i.e., the resistivity of the material and its thermal properties, which change the geometry, directly influencing the electric field distribution and surface roughness conditions. Fig. 37 and Fig. 38 illustrate the relationship and dependency be-

tween conductor temperature and corona loss, showing the increasing corona loss pattern as the conductor temperature decreases and vice versa. Here, a non-linear heat-balance equation (IEEE Std 738-2012) is used to approximate the conductor temperature, which depends on a number of other factors, including wind speed, azimuth, air temperature, heat gain, resistance, and electrical current magnitude, among others [78].

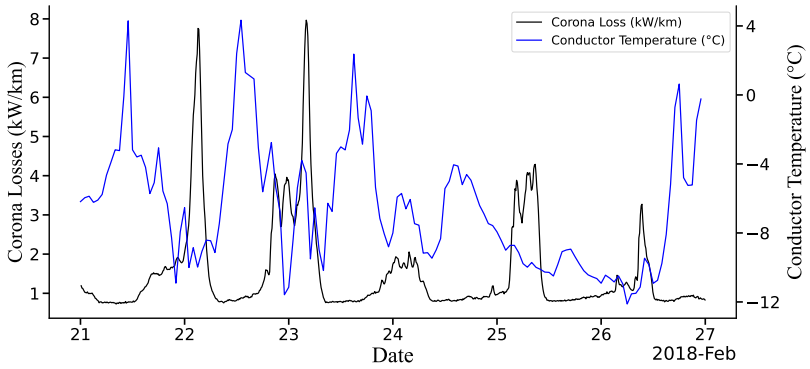


Figure 37: Dependence of corona loss and conductor temperature for New line [4].

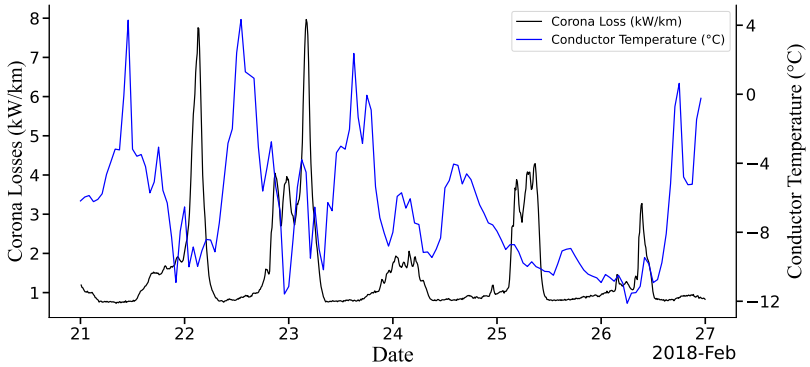


Figure 38: Dependence of corona loss and conductor temperature for Old line [4].

To gain insight into the effect of conductor temperature on corona loss, which depends greatly on various factors, indicates that a number of variables correlation should be assessed. Nevertheless, here, lower conductor temperatures, particularly in winter, are associated with a greater magnitude of corona loss. The arguments are supported by Fig. 37 and Fig. 38, which illustrate that the higher corona loss value for new transmission lines and old transmission lines, respectively, are exhibited mostly when the conductor temperature falls below -10°C and -15°C. In this case, the loading factor also has a significant impact on the conductor temperature, which in turn affects the formation of sharp icing edges or snow on the conductor surface, lowering the amount of corona loss. More details on conductor temperature influence are studied in the next subsection, where two years of data are evaluated.

3.6 Influence of Weather Parameters on Ageing Transmission Lines

In this section, some of the other major weather parameter correlations are examined. The data were collected from a nearby weather station. A combination of data from four weather stations refers to a uniform distribution of weather conditions for each line. The time-step for collected weather data is 10 minutes and estimated corona loss from PMU data are merged accordingly using two years (2018-2019) of data collected from the Estonian TSO.

3.6.1 Conductor Temperature

To understand the dependence of conductor temperature (estimated using nonlinear heat-balance equation from IEEE Std 738-2012 [78]) and corona losses, as illustrated in Fig. 39 and Fig. 40, statistical analysis is carried out using scatter plots, which plot the distribution of corona loss magnitude for different air temperature ranges, to comprehend the dependence. In the previous section, the weekly trend of conductor temperature impact was discussed, and in this section, two years of data are considered to analyse the correlation. From the figures using the percentile line, it is clearly seen that the higher magnitude of corona loss occurs mainly during winter. Comparing both lines, similar results were predicted, i.e., higher corona loss for lower conductor temperature where the older transmission lines exhibit complex relations and a new line shows direct relation with a value of three times higher, and as the conductor temperature increases, the corona loss is significantly reduced.

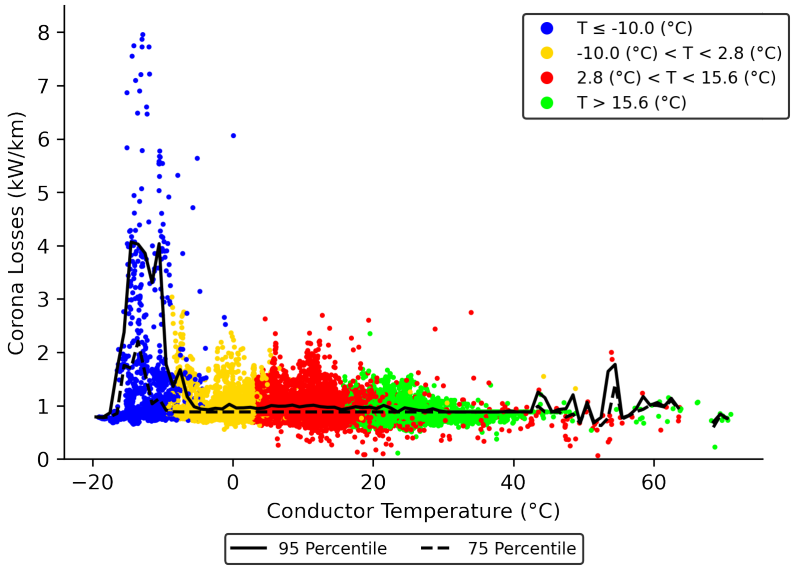


Figure 39: New line conductor temperature dependence on corona losses under varying air temperature [5].

One of the main reasons for the diminishing effect is clearing sharp icing surfaces, snow, and thermal expansion. Even using the percentile data plots the distribution of magnitude of corona loss values is high during winter compared to summer. The investigation acknowledged that corona loss is not necessarily high at lower temperatures based on the figures. In practice, corona loss decreases in winter due to the load current that

determines and affects the surface temperature of the conductor.

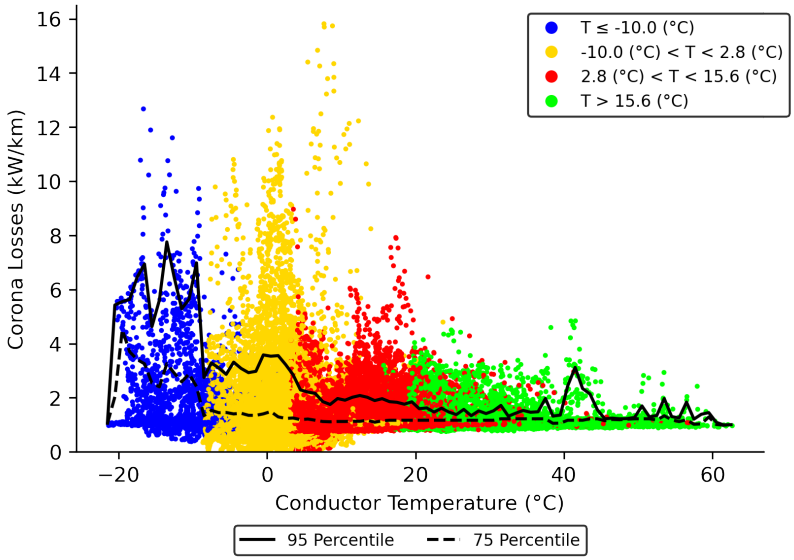


Figure 40: Old line conductor temperature dependence on corona losses under varying air temperature [5].

3.6.2 Air Temperature

The results shown in Fig. 41 and Fig. 42 have important implications for the dependence pattern, since they demonstrate the rate of change in corona loss magnitude with changing air temperature. The investigation demonstrates that for both lines, the corona loss reaches its maximum values during the winter months when temperatures drop primarily less than $-5\,^{\circ}C$. However, it is interesting to observe that corona loss for old transmission lines shows a complex relation with higher magnitude and fluctuating behaviour between $2\,^{\circ}C$ and $-20\,^{\circ}C$, whereas, in the case of a new transmission line, a clear relation is observed, i.e. higher corona loss in winter ($T < -8\,^{\circ}C$) almost four times compared to the summer season ($T > 5\,^{\circ}C$).

The formation of sharp ice crystals, sharp icing edges and the accumulation of snow on the conductor's surface are considered to be the cause of the increased corona loss during the winter [50]. In order to comprehend the distribution of corona loss, the percentile levels of 95th and 75th percentiles are evaluated.

3.6.3 Air Pressure

Fig. 43 and Fig. 44 illustrate the relationship between air pressure and corona loss at different temperatures for old and new transmission lines. Seasonality must also be taken into account when examining the correlation between air pressure and corona losses. However, in general, high air pressure that corresponds to winter with lower air temperature will exhibit higher corona losses, whereas during summer, due to higher temperature, it will have a complex relation with varying values. Low air pressure typically indicates milder weather in the winter and damper or rainier weather in the summer. However, the focus of the present study was to examine any correlation between air pressure and corona losses, where the new line effectively shows higher corona loss at higher pressure,

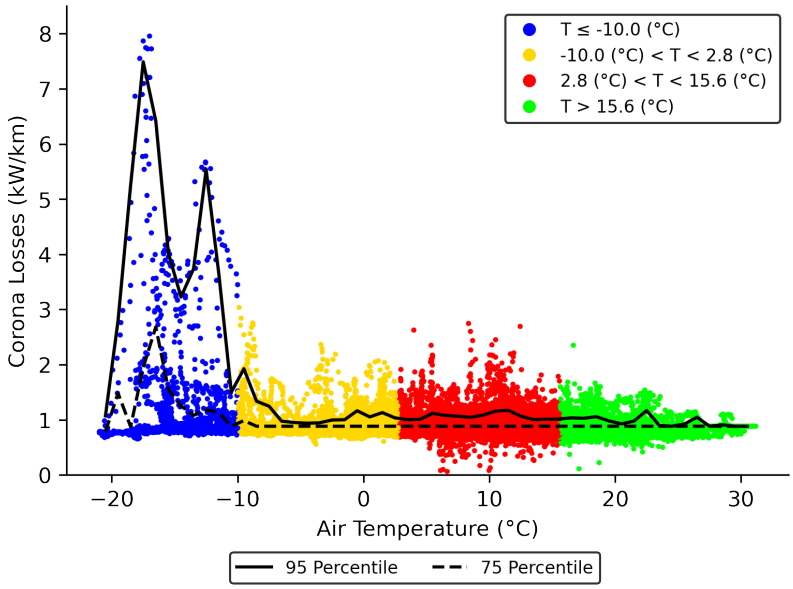


Figure 41: New line air temperature dependence on corona losses [5].

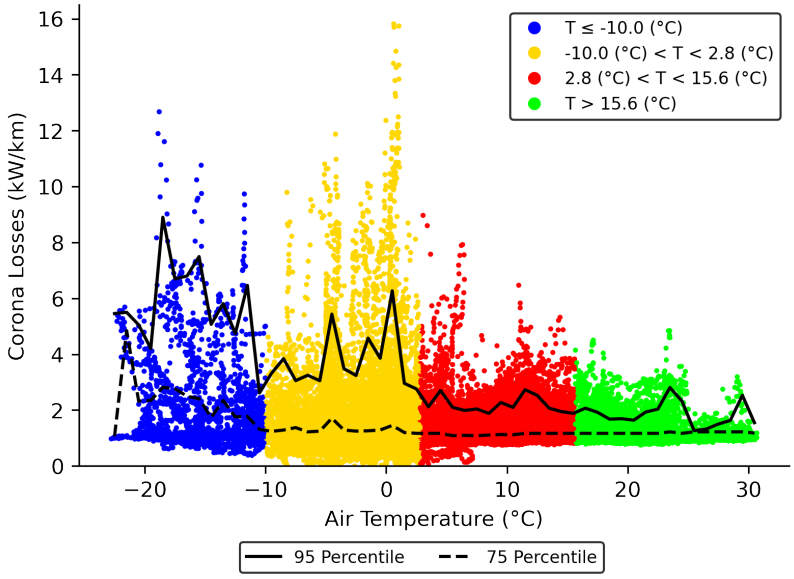


Figure 42: Old line air temperature dependence on corona losses [5].

but this may be the result of cold winter conditions indicated by blue dots. In the case of the old transmission line, a rather complex relation was observed but with similar characteristics. The impact is difficult to discern in older lines, and corona peaks appear at lower pressure following Peek's empirical equation.

According to Peek's empirical formula, the corona loss should be large effectively at low pressure, an example of such a case is the mountainous region where the atmospheric

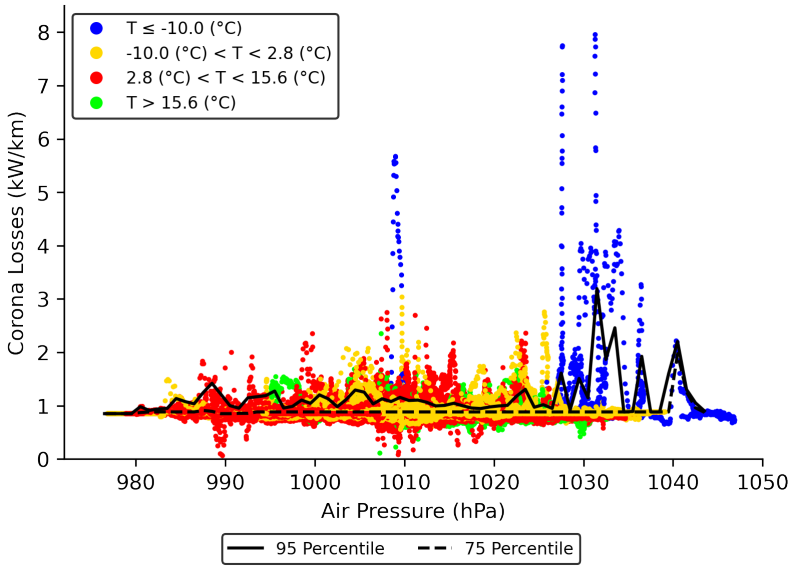


Figure 43: New line air pressure dependence on corona losses under varying air temperature [5].

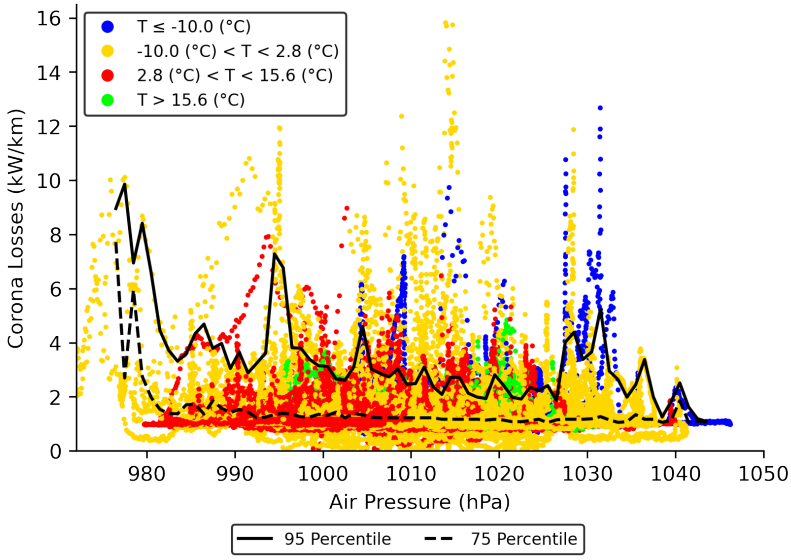


Figure 44: Old line air pressure dependence on corona losses under varying air temperature [5].

pressure is very low and the difference between the operating voltage and disruptive voltage is increased, which results in a higher corona loss value in the hilly region. Additionally, a high voltage supply at lower pressure, may exhibit higher corona losses when the dielectric strength of the air decreases. Nevertheless, in our case, higher corona loss was observed due to winter, i.e., lower temperatures and high pressures for both lines and a higher magnitude of corona loss for the older transmission lines showing complex correlation.

3.6.4 Wind Speed

Fig. 45 and Fig. 46 show the correlation between corona loss and wind speed for both transmission lines. Generally, at lower wind speeds, a spark commences the conducting channel due to the quick ionization of air, which is sustained, however at higher speeds, ion mobility disrupts the corona discharge phenomena, resulting in reduced corona loss values. The conducting channel at lower wind speeds heats up and maintains air ionization before forming an arc. The investigation revealed that when the wind speed increased, the temperature differences decreased, resulting in only high temperatures being observed, due to which a lower corona loss is observed. However, lower wind speeds coincide with lower air temperatures and results in higher losses for both lines, although the magnitude of the corona loss is higher for the old transmission line.

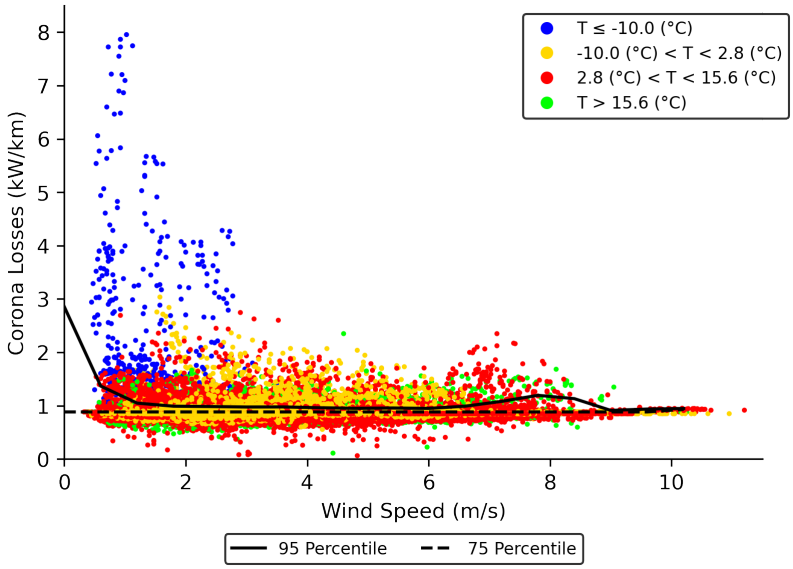


Figure 45: New line wind speed dependence on corona losses under varying air temperature [5].

In summary, a similar pattern was seen, with small variations among the percentile curves. The theory might be interpreted as peak current density decreasing or exerting less influence on conductor temperature. The most typical cause of decreased corona loss at increasing wind speeds is snow melting or conductor surface cleansing, which leads to greater ion mobility.

The overall results of the present study discussed in this section show the behaviour and dependence of different weather parameters on corona losses. Corona losses are strongest and exhibit higher magnitudes in low-temperature environments, and similar characteristics were seen for the conductor temperature. Although while considering other factors such as pressure or wind speed, it is very difficult to conclude as old lines show complex relationships and depend on various factors, but the new lines have a precisely proportional correlation.

3.7 Intermediate Summary

Many countries are experiencing a deterioration of overhead transmission lines with climate change. One approach to improving power transfer efficiency and reducing power

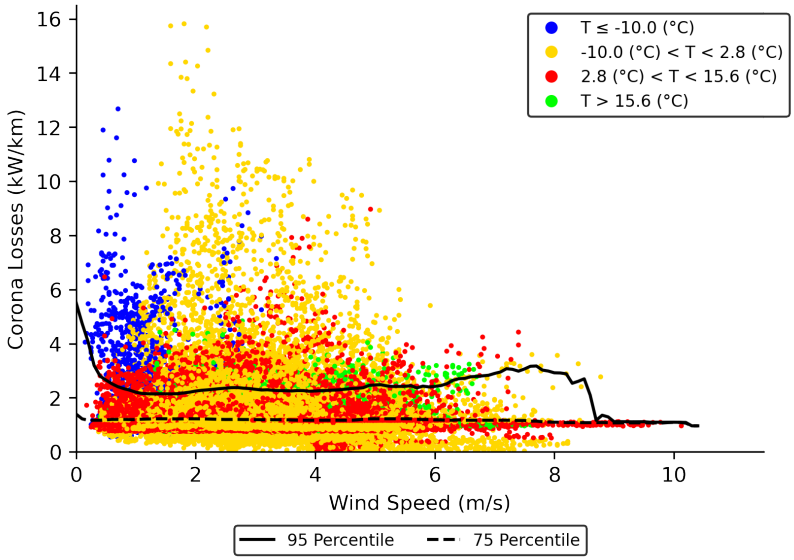


Figure 46: Old line wind speed dependence on corona losses under varying air temperature [5].

loss purchases in the energy market is to minimize the active power losses (joule loss and corona loss) in an overhead transmission line. This Chapter discusses the dependence between the design configuration, weather parameters, and corona losses for ageing transmission lines, presenting analysis results. Several parameters have been evaluated, such as tower configurations, GMD, GMR, conductor temperature, seasonal behaviour of corona loss, and different weather parameters, to investigate the effect of these parameters on the magnitude and trend of corona loss over a given period of time.

The study underlines the importance of several factors and demonstrates that aged transmission lines experiences higher corona loss and are sensitive compared to new transmission lines under various environmental conditions. This study contributes to the identification of important weather parameters and configuration setting-related resilience on corona loss for two different ageing transmission lines by assessing their influence. These results assist TSO in understanding the distribution of corona losses for various aged lines, selecting which lines to reconstruct and adapt to future weather circumstances. Furthermore, the data is particularly valuable for developing a prediction model for forecasting purposes, which is discussed more in detail in Chapter 4.

4 Machine Learning Framework For Corona Loss Forecasting

In this Chapter, the main aim of this study is to present framework for a corona loss forecasting model using machine learning. In the case study, the higher voltage transmission line is considered, where various weather station data along the transmission lines and respective PMU measurement data from both ends of the lines are collected. To understand the impact of different weather parameters on corona loss, different forecast scenarios are investigated using different machine learning algorithms such as XGBR regression and Ensemble random forest, linear regression, etc. The results are interesting and useful for TSO in reducing the purchase of power loss in power system networks.

The findings and discussions presented in this Chapter are published in VI.

4.1 Background

Losses in the electrical network can have a significant effect on network operation and on operational costs and therefore it is of importance to estimate these as accurately as possible. In literature [90], the author observes that the prediction model for grid losses was facing difficulties due to large fluctuations in corona loss for a short duration. Therefore, the realistic and precise forecasting of corona losses will help TSO to manage the increasing operational cost. While different methodologies have been developed for forecasting corona losses, arguably little research has been carried out, contributing to our understanding of complex weather conditions with different weather parameters, including analytical and empirical equations. The empirical methods have limitations considering complex weather conditions. Therefore, In recent years, statistical and machine learning approaches have been applied, with significant potential for forecasting corona losses with higher accuracy [91, 92, 93].

A machine learning system employing the CatBoost algorithm has been effectively deployed for commercial use in real-time day-ahead planning for expected grid loss for a Norwegian utility company at the distribution level, where the corona loss is ignored [94]. Similarly, the authors in [95] assertively employ a back propagation neural network model to compute line loss in the distribution system to improve the energy management system. Several other studies have conducted investigations on overhead transmission losses, neglecting corona losses, for example, in [96], the author proposes different machine learning algorithms such as ARIMAX, SARIMAX, Feedforward Neural Network, etc., to forecast day-ahead grid loss, which help to minimize the risk for both producers and consumers, facilitating a less volatile energy market by reducing the occurrence of high imbalance prices. However, in these approaches, the impact of corona losses, which influences the accuracy and purchase of overall losses in the energy market, is not adequately assessed.

Many additional studies have contributed to our understanding of corona loss forecasting using different methods. Among them, this work in the literature [92] offers valuable insights on a statistical approach tested to monitor hourly corona losses, to assist with the imbalance costs in the Bulgarian TSO. In [93], the approach utilizes more advanced techniques such as EEMD–LSTM–SVR prediction model based on ensemble empirical mode decomposition (EEMD) to decompose line losses. Using the heat balance equation to predict transmission line losses, the authors in [97] proposed a methodology using the numerical weather input and day ahead hourly forecast of the power flow, GIS mapping with line assets. In [98], the author proposes a method using power flow analysis and fuzzy based weight clustering to forecast short-term active power losses. Croatian Transmission System Operators (HOPS) studied a comparison of the support vector re-

gression (SVR) to the business-as-usual approach using weather input data to forecast short-term active power losses. Similarly, another machine learning-based linear regression and exchange flow modelling was applied to the Swedish National Grid for testing for transmission line loss prediction. Nevertheless, most of the analyses are done for overall transmission line losses, neglecting corona loss influence in an overhead transmission lines lacking information on multiple weather parameter influence.

The corona loss calculation model using a corona cage is appropriate, but it is difficult to measure under complex weather conditions [26]. Therefore, using the synchronized current and voltage phasor measurements obtained from the PMUs, the calculated transmission line parameters are more precise to estimate corona loss in an overhead transmission line [42, 43]. Consequently, this work uses real-time estimated corona loss using PMU measurements. Understanding the influence of different weather parameters on corona loss in high voltage lines was discussed in [3, 70]. However, in the previous studies, sufficient research has not been done to assess the corona loss forecasting scenario using network operational measurements and various weather parameters. Forecasting corona losses is not straightforward and depends on various interrelated factors. While there are various methodologies available in literature, they have limitations when it comes to considering complex weather conditions with different weather parameters.

In the absence of complete data availability, creating a forecasting model based on influencing factors such as weather parameters to tackle the forecasting problem of corona loss is beneficial for TSO and to balance the cost of transmission losses in energy market purchase. Monitoring large power networks or transmission regions can be greatly enhanced by utilizing this approach, especially in situations where only a limited number of line losses can be measured. This approach enhances the understanding of transmission network loss forecasting accuracy and efficiency. To the authors' knowledge, no research studies have comprehensively evaluated the practical use of PMU measurement data and different weather data for forecasting corona losses.

This Chapter discusses the corona loss forecasting model for the 330 kV overhead transmission line data available for various scenarios. Its performance was analysed using statistical residual plots and error metrics. The analysis consists of four sections where in the analysis, the following aspects were considered: i) influence of time-step between weather measurement data points, ii) seasonal trends to assess the behaviour of the forecast model during a particular season, iii) combined forecasting on two transmission lines to analyse forecasting on larger parts of a grid, and iv) training the forecast models with reduced input features to determine the most important weather parameters needed for forecasting. The main aim of the work is to look ahead in the time horizon where weather forecasting accuracy is relatively high – a few hours to a few days – and where forecasting corona losses is most critical. Synchronized phasor measurements can be used for online monitoring of losses and gathering historical data of losses over longer periods of time. Combining it with weather monitoring data allows us to train a prediction model. Once there is a model that can predict historic corona losses based on inputs of historic weather data, the model can also predict future corona losses given the inputs from a weather forecast. The study mainly uses different regression models for forecasting – Ensemble Random Forest, XGBoost regression and a simple linear regression model.

4.2 Methodology

In this section, the proposed methodology is described using a flow chart in which the overall forecasting process is shown in Fig. 47. The methodology includes two phases. Phase I discusses the estimation of corona losses using PMU data from both ends of the

high-voltage transmission lines, which were based on the ideas outlined in [43]. In addition, to reduce systematic errors, a correction mechanism introduced in [44] was used, which is discussed in more detail in Chapter 1. The estimated corona loss solution obtained from the PMU data can be applied while accounting for current weather conditions, using different weather parameters obtained from the Estonian Environmental Agency (seven different weather stations) over a period of two years. Further, the collected data are preprocessed in Phase II, where outliers are removed, and the missing data is filled in. The data set is divided on the basis of different time steps, seasonal, combined lined data, and reduced parameter data. In [3], using statistical graphs analysis and feature importance, the correlation of different weather parameters and corona loss were depicted. The analysis results show how different weather parameters influence corona loss. Therefore, it is important to identify how these important parameters affect the corona loss forecasting model and to improve the efficiency of the model. Forecasting models such as the multiple linear regression model, ensemble random forest regression (ERF) with different hyperparameters, and extreme gradient boost regression (XGBR) with three-fold cross-validation were tuned to improve the efficiency and model accuracy. The root mean square error (RMSE) and mean absolute error (MAE) are used to determine and evaluate the efficacy of the model in different case scenarios.

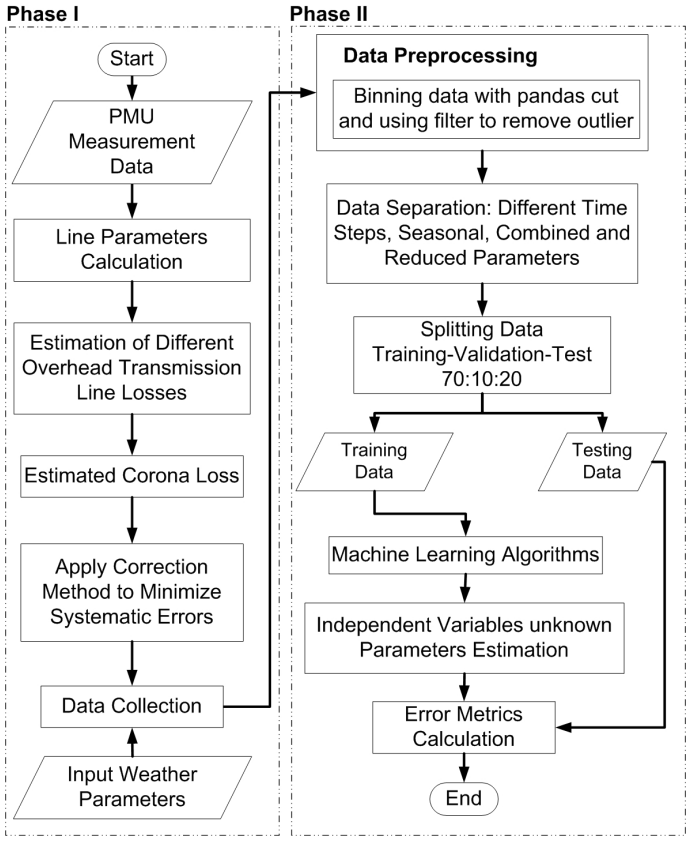


Figure 47: Flowchart of corona loss forecasting model using machine learning [6].

4.2.1 Data Collection and Preprocessing

The actual operational PMU data from both ends of the lines were collected from the Estonian TSO, Elering, to test the forecasting model. To monitor corona losses for Estonian 330 kV lines, data was collected for a period of two years, that is, 2018 and 2019 to assist in the analysis of the machine learning forecasting model. The availability of the PMU data will help to determine line parameters, which are then used for estimating corona loss and other types of losses, such as Joule, inductive, and capacitive losses.

In the case of weather parameters, the data are collected from the Estonian Environmental Agency; depending on the line directions, different weather station data are acquired, i.e., locations of four nearby weather stations for each line, as shown in Fig. 48 marked by a round dot. Various weather parameters, such as air temperature, humidity, air pressure, precipitation, visibility, wind speed, and wind direction, hourly data are collected for analysis [99]. Although Estonia has a flat terrain with an elevation difference of around 55 metres, it is necessary to combine weather station data to avoid the influence of changing weather conditions along the line, as previously discussed in more detail in Chapter 2, Section 2.3. The approximately transmission line length of 170 km (north-south) and 210 km (east-west) are illustrated in Fig. 48 represented by blue coloured lines and substations with square markers. More details about the transmission line tower configuration and other design parameters can be found in [4].

The collected weather station data are combined using simple averaging methods to observe the real condition of weather variations along the transmission line, although various methods are available to combine data in multiple ways, i.e., linear combination, exponential combination, MAPE-based combination, geometric mean combination, two-fold combination, GA-based combination, and sophisticated weighting techniques, etc. [69]. However, it is difficult to select which method would be suitable for combining the data and, therefore, a simple averaging method is selected.

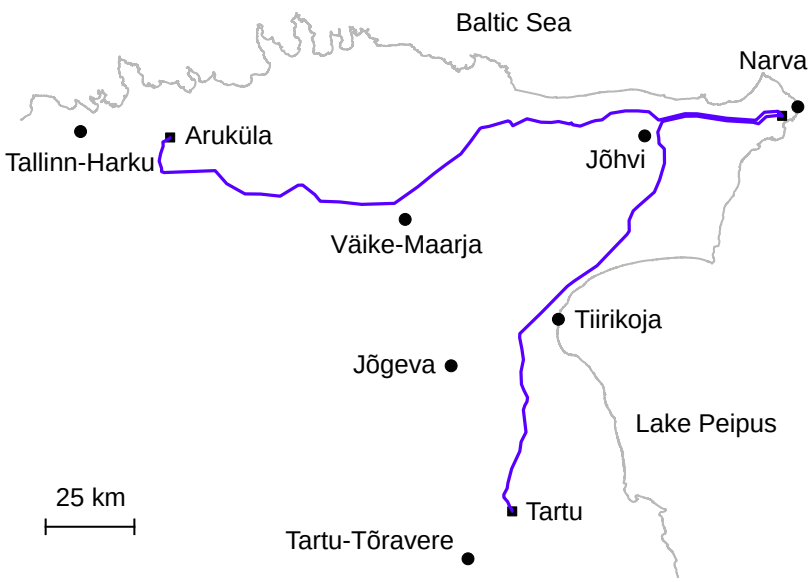


Figure 48: Map of monitored 330 kV transmission line corridors. Lines are coloured blue, substations are denoted by square markers and weather stations by round dots [6].

4.2.2 Corona Loss Estimation Method Using PMU Measurement

In the context of estimation algorithms referred from reference [43], the PMU measurement data from both ends of overhead transmission lines are collected to calculate the line parameters, which are further used for the estimation of different losses in an overhead transmission line such as resistive, corona and capacitive loss, etc. It is a useful way for real-time monitoring of corona losses in real-practice application domains. The distributed line equations are expressed according to (18) and (19).

$$\underline{\gamma} = \frac{1}{L} arcosh\left(\frac{\underline{V}(l_s)\underline{I}(l_s) + \underline{V}(l_r)\underline{I}(l_r)}{\underline{V}(l_r)\underline{I}(l_s) - \underline{V}(l_s)\underline{I}(l_r)}\right) \tag{18}$$

$$\underline{Z}_c = \frac{\underline{V}(l_s) - \underline{V}(l_r)\cosh(\underline{\gamma} l_s)}{\underline{I}(l_r)\sinh(\underline{\gamma} l_s)} \tag{19}$$

where L is the length of the transmission line (km), $\underline{V}(l_s), \underline{I}(l_s), \underline{V}(l_r)\, and\, \underline{I}(l_r)$ are the voltage and current phasor values at the sending and receiving end.

The per unit impedance and admittance are represented by $\underline{Z} = R + jX = \underline{\gamma}\underline{Z}_c$ [Ω/km] and $\underline{Y} = G + jB = \underline{\gamma}/\underline{Z}_c$ [S/km].

The active transmission loss and reactive power balance can be represented by integrating over the line length, as shown in (20) and (21):

$$\Delta P = R\int_0^L |\underline{I}(l)|^2\, dl + G\int_0^L |\underline{V}(l)|^2\, dl \tag{20}$$

$$\Delta Q = X\int_0^L |\underline{I}(l)|^2\, dl - B\int_0^L |\underline{V}(l)|^2\, dl \tag{21}$$

where $\underline{V}(l)$ and $\underline{I}(l)$ are the voltage and current phasor values at distance l along the line. The dependence of the dissipation losses across the transmission line network is mainly due to resistive (R) losses, consumption of reactive power in reactance (X), conductance (G) losses (due to the leakage current and corona phenomenon) and generation of reactive power in susceptance (B).

For the measurement accuracy problem, we applied the error correction method proposed in [44] to reduce the systematic PMU measurement error to improve the corona loss estimation, which is based on the linear proportionality fit of the shunt parameters to the power flow.

The corrected corona loss values is expressed using (22):

$$P_C(t) = T_C(t, z(t)) - B(x(t)) \tag{22}$$

where $B(x(t))$ is defined based on an approximation linear model and is given by (23).

$$B(x(t)) \cong D\,x(t) + G - F \tag{23}$$

where $x(t)$ is the assumed value of average power flow and the constant values D, G, and F are assumed for a given transmission line.

4.3 Machine Learning: Regression Model

In this analysis, multiple linear regression and ensemble random forest regression model along with the extreme gradient boost regression (XGBR) model were used. Regression models are generally a powerful tool in data analysis due to their simplicity, ease of interpretation, and ability to make predictions outcomes desirable for practical applications.

4.3.1 Multiple Linear Regression Model

The multiple linear regression model aims to analyse the impact of various weather parameters on corona losses. The objective is to explain as much variability as possible in the observed corona losses over a specific period. Using multiple input variables, it evaluates the strength of the relationship between the dependent variable (corona loss) and the independent variables (weather parameters). Mathematically, using a bias constant applied to the weighted sum of the input variables to estimate the predicted value of the target variable is expressed as:

$$Y = \alpha_0 + \omega_1 X_1 + \omega_2 X_2 + \ldots\ldots + \omega_n X_n \tag{24}$$

where Y is the corona loss (dependent or target variable), X_1,X_2,...,X_n are the selected weather parameters collected along the overhead transmission lines (independent variable), ω_1, ω_2,...,ω_n are the weighted slope coefficient and α_0 is the bias constant or intercept.

Multiple linear regression is chosen for comparison for two key reasons.

- To establish the linear relationship strength, it is crucial to understand the correlation between the dependent and independent variables. This allows for good predictions and informed decision-making based on the data.
- It is the most widely used method that is easily implemented to comprehend the rate of change.

4.3.2 Ensemble Random Forest Regression

The Random Forest Regressor decisively employs a collection of training multiple decision trees to make predictions and improve accuracy. Each decision tree is trained on a random subset of input features and training data, ensuring a robust model that effectively captures a variety of patterns in the data. To enhance performance, the concept of ensemble is introduced, which uses a series of decision trees that grow in randomly chosen data subspaces to form a constructive prediction ensemble, where each tree collection is created by selecting a group of inputs at each node with non-overlapping regions [100].

The ensemble random forest algorithm operates by sampling the original dataset numerous times, allowing for a robust construction of the random forest in parallel with different hyperparameters, and the number of variables in a tree node significantly influences the decision-making process of a decision tree. This process involves iterating through various subsets of the data, each time refining and enhancing the model's ability to capture essential user feature information with randomized subsets of predictors and uses averaging trees to improve accuracy and overfitting, making it a powerful tool in data analysis. It is widely used for various real-world applications, such as partial discharge localisation [101], power quality assessment [102] and in various other power systems applications. The network architecture of the ERF model is shown in Fig. 49.

The concept of using an ensemble system and random forest theory to forecast corona losses in overhead transmission lines is proposed. A voting regressor algorithm is used to predict the mean output value that effectively integrates multiple random forest algorithms to create the best predictor, enhancing predictive accuracy and providing a more robust model for forecasting to fit the final trained ERF regressor [103]. Following that, on the validation and testing set, the trained ERF final predictions were carried out. The main advantage of using ensemble systems is beneficial when complex multiple-input data are involved, i.e., forecasting corona loss, which is highly dependent on complex weather conditions. For real-time corona loss forecasting, the ERF model works efficiently using non-

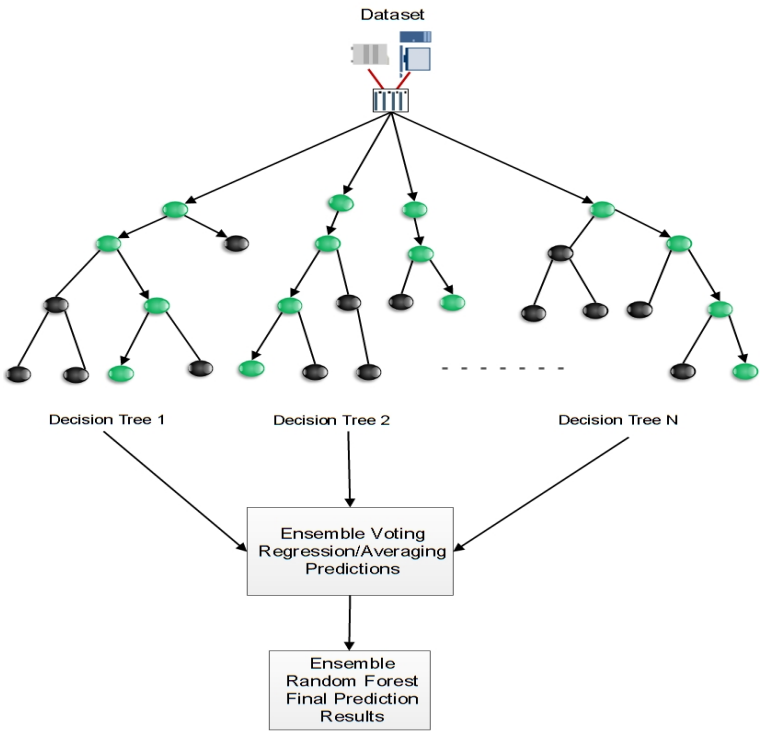


Figure 49: Network architecture of Ensemble Random Forest Regression [6].

linear functions and counteracts the flaws of equally performing random forest regression models with a majority vote system to combine the final forecasted value.

Hyperparameter tuning involves optimizing predefined settings of a machine learning model to improve its performance. The number of trees in the forest (forest size), the number of available predictor parameters (features) used, and the maximum depth of the tree nodes (leaf size) are parameter settings that significantly impact the model's accuracy and generalisation. Selecting the right hyperparameters is crucial for achieving optimal results. In our case, the selected combination of hyperparameters tuned for the ensemble random forest model are listed in Table 6.

Table 6: Hyperparameters tuned for ERF Regressor

Hyperparameters	**Experimental Chosen Values**
Number of Trees (n_estimators)	100, 500, 1000
Maximum Depth of Tree (max_depth)	20, 30, 50

4.3.3 Extreme Gradient Boost Regression (XGBR)

XGBoost Regressor is a supervised machine learning gradient boosting algorithm that employs an ensemble of decision trees. This method iteratively addresses the errors produced by the preceding trees, thereby enhancing overall accuracy and predictive perfor-

mance. It is an efficient ensemble learning approach that addresses complex and non-linear relationships between input and the target variables. It enables parallel processing, making it efficient for managing large datasets and reducing training time.

The analysis of the various case studies was conducted employing the Extreme Gradient Boosting Regression model, demonstrating its effectiveness, with three-fold cross-validation. The satisfactory performance was observed with the important default parameter while tuning, where n-estimators=100, learning rate (η = 0.1), max depth=6, min child weight=1, tree method='auto', n jobs=1, sample by level node and tree.

4.3.4 Error Metrics for Evaluations

Two sets of error metrics are chosen to identify and understand the behaviour of the forecasting model and how much error differs for an individual result. $RMSE$ is most widely used in the case of regression; accordingly, the model performance is evaluated. However, MAE is widely used in industries and is therefore considered to check the suitability of different metrics to recommend and evaluate how well various metrics capture error distributions.

- $RMSE$: The root mean square error ($RMSE$) serves as a standard statistical error metric for assessing model performance across various fields, including meteorology, energy markets, and climate research. The $RMSE$ assumes errors are unbiased and normally distributed, presenting a comprehensive view of the error distribution. While the $RMSE$ metric is a valuable tool for measuring errors, it's important to be aware that it can be particularly sensitive to outliers. This characteristic can influence the results, so it's beneficial to consider other metrics such as MAE to ensure a more robust analysis when outliers are present.

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y'_i - y_i)^2} \tag{25}$$

- MAE: The mean absolute error (MAE) is a valuable metric that is commonly used to evaluate models, providing insights into the accuracy of predictions. Mathematically, it is expressed as:

$$MAE = \frac{1}{m}\sum_{i=1}^{m}\left|y'_i - y_i\right| \tag{26}$$

  Here, m is the total number of non-missing data points, w_i is the weights multiplier, y' is the forecasted value and y is the actual true value of the dataset.

4.4 Results and Analysis

In this section, results using different data sets to understand the forecasting model for corona loss using machine learning, where various scenarios are evaluated as listed in the subsection, are presented. The performance of the machine learning forecasting model was examined using the PMU and weather data collected over two years, where the training, validation and testing sets were randomized split to approximately 70%, 10%, and 20% where validation data are needed. Yet, the validation data were not required in some cases of forecasting corona loss in seasonal weather such as autumn, spring, summer, and winter. The machine learning algorithms are different regression models with an input of weather parameters to the algorithm, i.e., independent variables and estimated corona

loss as dependent variable. The input weather data sets used for the evaluation include air temperature, air pressure, humidity, precipitation, visibility, wind direction, and wind speed, compiled from nearby weather stations.

4.4.1 Varying Time Step Length

In recent years, the vast and constantly growing amount of data and computing power, along with the development of better machine learning algorithms, have propelled a new wave of using machine learning algorithm applications in the practical power systems domain. The concept of using operational PMU data and weather data to apply for forecasting scenarios for overhead corona loss has not been observed yet. Therefore, in this section, different machine learning models are trained for various time steps to investigate the impact and role of time steps. The estimated corona loss is observed at hourly, 10 min, 5 min, and 2 min time steps, where the weather data are interpolated accordingly, as detailed below.

Usually, it is easy to gather the hourly resolution weather data, and the estimated corona loss data are downsampled to an hourly time step to establish a dataset for training the machine learning models. The data set contains over 15 000 data points after data pre-processing. Here, the training, validation and testing sets are divided into 70%, 10%, and 20%, where different regression models are trained – linear regression, ensemble random forest, and XGBoost regression model. The residuals of the predictions on the validation data are plotted as shown in Fig. 50 to evaluate the performance of the different regression models.

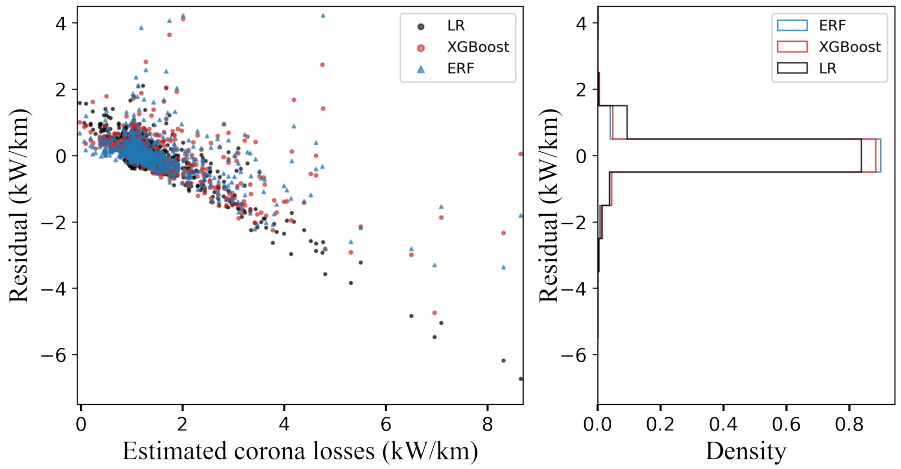


Figure 50: Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the hourly dataset [6].

Based on the analysis of the hourly dataset, we observe three main findings, which are listed below.

1. First, most estimated corona loss values are in the range of baseline corona losses, which are present under normal conditions, and the machine learning model ERF and XGBoost performance have relatively small deviations in the predictions compared to linear regression.

2. Secondly, to address the higher corona losses, as this is our primary area of interest and main focus, the prediction is not very good or may not be as adequate as the number of observations is limited, missing the spike at shorter duration.

3. Thirdly, the linear regression model tends to struggle with predicting losses in the higher range of corona losses. In contrast, more advanced regression models such as ERF and XGBoost, offer improved performance in the area of forecast corona losses. To enhance our predictive accuracy, it would be beneficial to increase the number of high corona loss observations. This would provide a stronger data foundation for our analysis. In addition, to enhance our analysis of model performance, it would be beneficial to focus on observations that exhibit higher corona losses. This approach can provide valuable insights and help us improve our overall understanding.

However, in the case of hourly estimated corona loss data, it has little effect on the distribution, implying that the short-duration spike is missing and that the variation is very small, resulting in the loss of greater detail regarding higher corona loss. Therefore, utilizing high-precision measurement data can significantly enhance the effectiveness of the machine learning forecasting model. Although the outcome suggests that the degree of improvement varies over different time steps, where a data-rich environment with a distinct larger sample size plays a very important role in acquiring suitable forecasting performance. To enrich the datasets and broaden the exposure of our models to a more extensive array of training data, we can explore innovative strategies that incorporate diverse sources. By systematically increasing the volume and variety of information, we can enhance the learning experience for our models, ultimately leading to improved performance and adaptability in a range of scenarios by proposing a hypothesis of interpolation of weather data with shorter time steps to expand and improve the performance to detect smaller spikes and higher corona losses.

Weather phenomena tend to evolve gradually, allowing us to observe consistent changes in physical quantities such as ambient temperature, air pressure, and humidity within an hour. This slow progression provides opportunities for better understanding and forecasting of weather patterns in a smooth manner sufficiently so that the values of weather parameters can be reasonably interpolated to downsample the data to shorter time steps. Therefore, we propose to interpolate the weather parameter data to different time steps such as 10, 5, and 2 minutes, using a linear interpolation, i.e., efficient to implement and involves filling the empty values based on the known values with a linear fit. By utilizing interpolated weather data alongside estimated corona loss values for each corresponding time point, we can create larger and more comprehensive datasets. This approach enhances our analysis and understanding of the underlying patterns.

In a similar manner to the approach carried out with hourly data, it is applied effectively to 10 min, 5 min, and 2 min time-step datasets using different regression models. Further, the models are trained and evaluated with the industry-standard error metric mean absolute error (MAE) for forecasting based on the validation dataset, where the best performance is observed by the ERF model as shown in Fig. 51. Although ERF models show promising results with reduced error rates, indicating their effectiveness in improving accuracy with the extended datasets, then, overall results are not conclusive about whether any improvements were made. However, in hourly data, it is observed that the analysis of models should take an important, more detailed look at corona losses in the range of higher values.

The residuals of predicting losses on the validation data and their histograms result obtained for different time-step using the ERF model is shown in Fig. 52, where the fore-

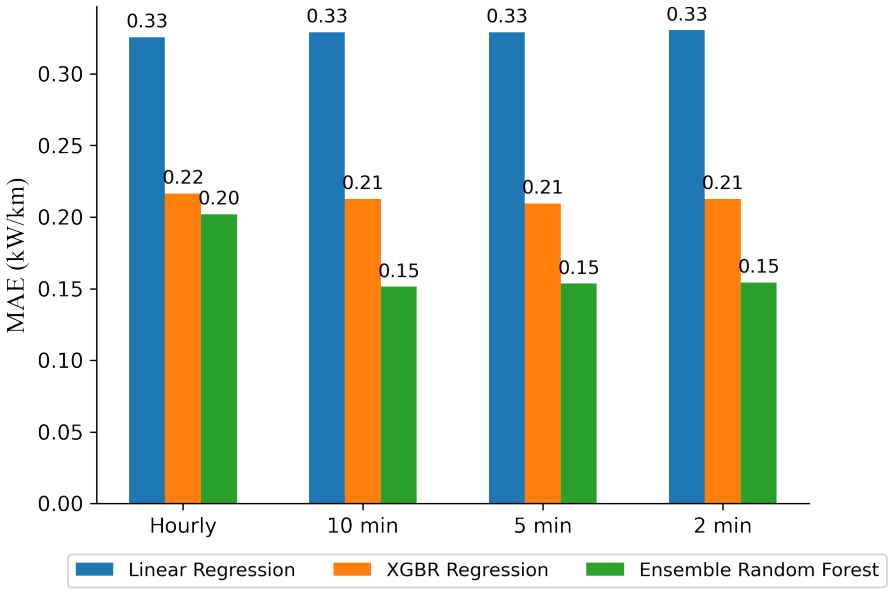


Figure 51: Evaluation of error metric MAE of forecasted corona loss values using machine learning models for different time steps [6].

casting is done taking into account the higher losses i.e. 90th percentile value of estimated corona losses, and looking at all data points above that value (i.e. top 10% of losses). This indicates that for a higher corona loss with interpolated weather data at a high resolution, the model performance is significantly improved. The distribution of residuals suggests that a significant portion of the predictions is closely aligned with the reference value of estimated corona losses. This alignment indicates a positive trend in our predictive accuracy. The residual plot serves as a valuable tool for understanding prediction errors, as it highlights the differences between the observed values of the target variable and the predicted values. By analysing this plot, we can gain insight into the model's performance and identify areas for improvement. In addition, the histogram shows that the error is mostly distributed around 0, indicating a better-fitted model for decreasing time-step models.

In order to be able to compare the behaviour of different machine learning algorithms, there is a need to investigate the accuracy of prediction losses trained based on the validation data on the 2 min dataset. The analysis includes all three regression models—linear regression, ERF, and XGBoost—along with their respective residuals and the histograms of predictions, as illustrated in Fig. 53 and Fig. 54. This comprehensive presentation allows for a clear comparison of model performance. Using the hourly dataset regression methods already demonstrated the result that the linear regression model finds it difficult to predict higher corona losses – the residuals show that the model predicts an almost constant value, whereas the XGBoost and ERF models show good-enough and satisfactory prediction of largely higher corona loss values, but the accuracy of the ERF model is clearly better as indicated by the histogram. On the test set with the full range of corona losses, the 2 min ERF model predicts results within ± 0.5 kW/km 93% of the time.

Although there is potential to decrease the time step further, it's important to recognize that there are diminishing returns to this approach. Exploring other methods alongside this adjustment could enhance overall effectiveness. Nevertheless, the results ob-

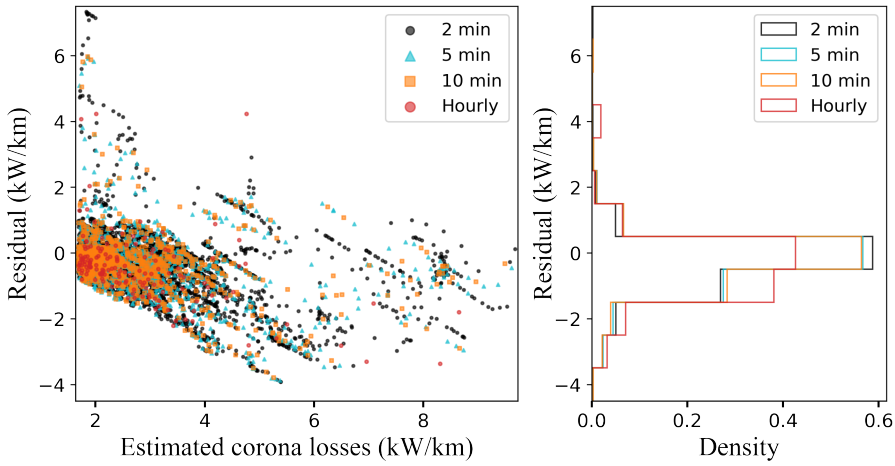


Figure 52: Residuals and their histograms comparing the performance of the ERF model's forecasting capability with different time steps [6].

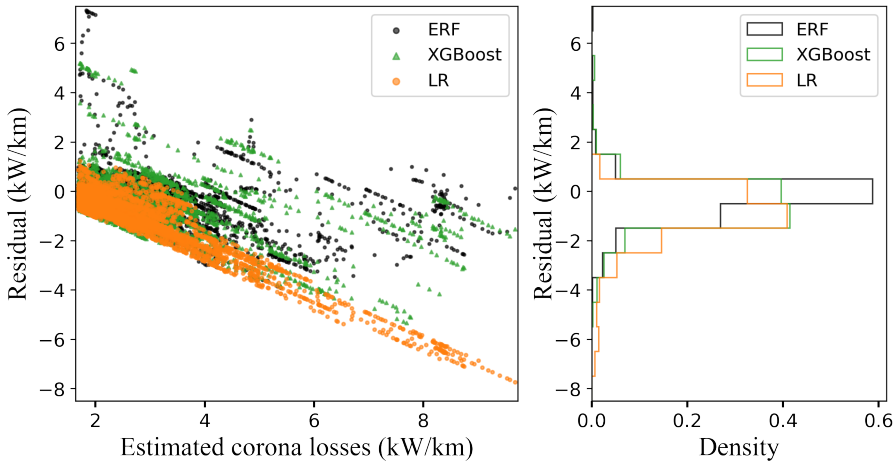


Figure 53: Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the 2 minutes data set [6].

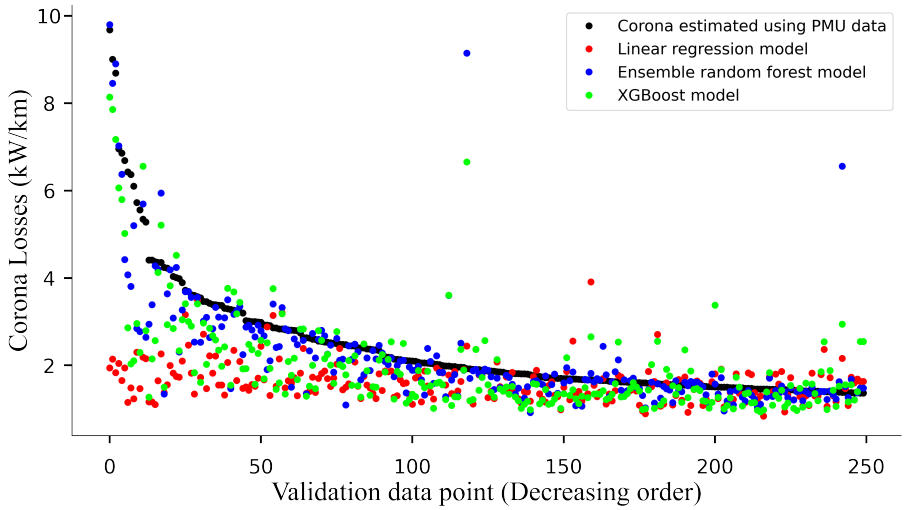


Figure 54: Comparison of different forecasting models and corona loss estimated using PMU data for 2 minutes data. Estimated measured values have been ordered descendingly from the highest recorded value.

tained from two-, five-, and ten-minute interpolated data show increasing accuracy becomes diminishing, increasing computational time for model training with large data volumes and decreasing time steps. This study identified that adopting a two-minute time step can enhance accuracy effectively with suitable computational time for training machine learning forecasting models. However, for smaller datasets, an even shorter time step might be more suitable.

4.4.2 Forecasting Model Seasonal Performance

Typically, classical seasonal patterns indicate that corona loss tends to increase during the winter months, while it decreases in the summer. This information could be valuable for the development of strategies to effectively manage and mitigate these losses throughout the year. Weather conditions are non-consistent, due to which the corona loss values are fluctuating with an increasing and decreasing pattern depending on various weather factors. Atmospheric changes are primarily the cause of corona loss fluctuation across various time scales, whether it is day-wise or seasonal divergences. To explore the correlation, one needs to understand the variability among seasonal trends. The analysis will help to investigate the behaviour and performance of the forecasting model, assessing the variability of seasonal atmospheric conditions. The input to the model data for training is the same input as used in the subsection 4.4.1.

To improve the analysis, it is essential that the monitoring data are divided into four seasons, namely winter, spring, summer, and autumn, according to the datasets timestamps. The ERF regression model is trained using a separate dataset created from each seasonal data, in order to observe and compare their differential behaviour. The analysis results for each seasonal data are depicted in Fig. 55, where an approximated probability density plot using the Seaborn library with a normalized histogram, which links the measured Corona loss observations and their probability, is plotted. The corona losses estimated from measurements are shown in the left subplot and the forecasted corona loss using ensemble random forests is depicted on the right side of the subplot for dif-

ferent seasons. It is clearly visible that during winter, higher corona loss data points are more frequent, and the distribution width is increased with higher values, increasing the probability, but a lower probability density function leads to less likelihood of finding similar values. During spring and summer, the forecast values are more stable, especially for low corona loss with higher probability density, leading to finding the values more likely, which enhances the forecasting performance as a majority of data points for corona losses represent low baseline losses. The prediction probability densities have higher peaks due to predicting closer to the average baseline corona loss more frequently. Using the ERF regression model, the seasonal patterns of the estimated corona losses align closely with the predictions for different seasons and combined annual data, suggesting a strong consistency that can be further explored for insight.

The results of seasonal corona loss forecasting are observed, and the performance of the ERF regression model is assessed using the highest 10% of corona losses depicted in Fig. 56 in the form of residuals and histogram, where the data points are evenly distributed around zero; however, at higher losses, it is complex and difficult to observe the same. The behaviour of the seasonal model trained is different compared to the datasets considered for a whole year. Table 7 displays the ERF regression model error metrics values for each season and the combined annual values, to help distinguish the findings. The results indicate that the ERF model during spring and summer gives outcomes of higher accuracy, whereas for higher corona loss i.e., in autumn and winter, forecast accuracy is reduced. The overall performance of the ERF model is better in every seasonal circumstance.

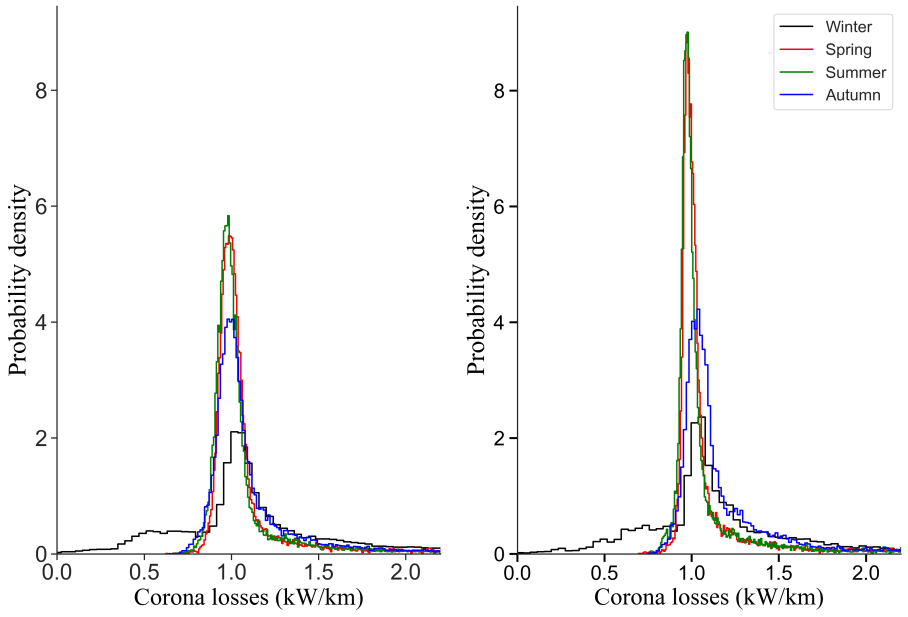


Figure 55: Measured corona loss (a) and forecasted corona loss (b) using the ERF model, for different seasons of the Estonian transmission network with a two-minute time step of interpolated weather parameters inputs [6].

When combining the overall results of all of the seasonal models, the prediction error on test data is within ± 0.5 kW/km 98% of the time, which is better than the single model

trained on the whole dataset. It can be seen that if the training data is collected over a period of several years, it can be beneficial to train seasonal models instead of or in addition to a model trained on the complete dataset. Predictions from such seasonal models can give more accurate results.

Table 7: RMSE and MAE comparison of corona loss forecasts for different seasons using a two-minute time step model [6].

Seasonal Period	ERF Model Error (kW/km)	
	RMSE	MAE
Winter	0.247	0.117
Spring	0.097	0.047
Summer	0.071	0.046
Autumn	0.163	0.090
Combined Annual	0.157	0.074

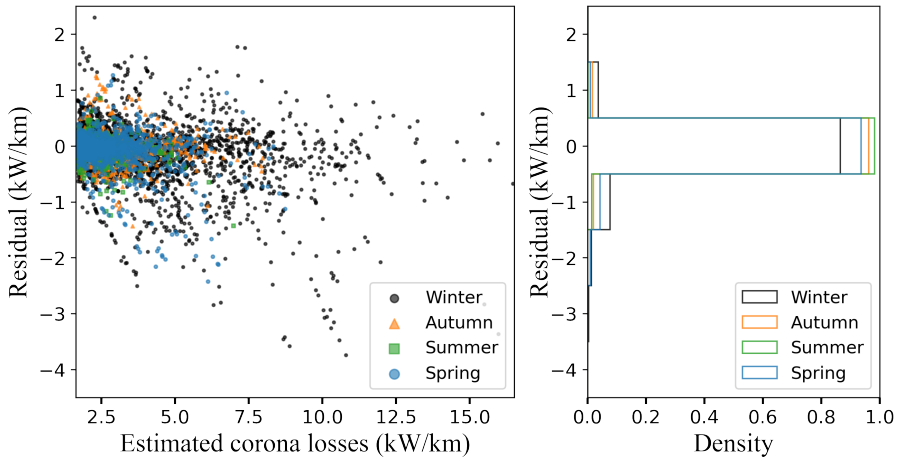


Figure 56: Residual plot using the ERF model, for different seasons of the Estonian transmission network with a two-minute time step of interpolated weather parameters inputs [6].

In summary, in order to analyse and interpret the learning behaviour for which data are collected based on the season, our model of learning behaviour targets the related patterns for each season – autumn, spring, summer, and winter. Here, computational time is significantly reduced as the data volume is reduced for training the model. The results show that machine learning is a powerful method that provides a fully automatic, unsupervised diagnostic behaviour analysis specifically tailored to seasonal behaviour-based analysis, thus it can be used as an approximated accurate tool for forecasting corona loss using multiple weather parameters.

As a result, we advise using more weather-related parameters in corona loss forecasting systems wherever possible, as these indicators can assist the model in generating better accurate forecasts. Overall, the suggested ERF model are more flexible in addressing the effects of seasonal behaviour.

4.4.3 Combined Transmission Line Corona Loss Forecasting

In this section, the forecasting of corona losses for the combined transmission line is presented. Generally, it is important in practical applications to monitor and forecast corona losses for the entire power system network or some part of the region, depending upon the transmission network. Therefore, using data from multiple transmission lines for different geographical regions with multiple weather stations would help to monitor and forecast corona losses with greater accuracy, providing a clearer overall picture of corona losses across the entire network.

In order to enhance the understanding and impact of this study, two overhead transmission line datasets are acquired to carry out the analysis effectively. The regression model is trained using the two-line estimated corona loss data and weather parameters data (collected across the line) to evaluate the forecasting model performance for a larger area or a two-grid network. The result obtained is for 2 minutes interpolated weather data and estimated corona loss, corresponding to a higher 10% of losses. This is depicted using the residuals of the predictions and the histograms for the combined and single lines as shown in Fig. 57. The plotted figures show the normal distribution of corona losses using the ERF regression model for combined and single lines, where the comparison shows a similar pattern corresponding to a higher corona loss.

The ERF regression model is trained for both transmission lines with error prediction rates of a range within ± 0.5 kW/km 95% of the time, compared to 93% when trained on the data for the single transmission line. The RMSE and MAE error metrics is shown in Table 8 for both the cases. The error difference is significantly small, and the limitation of data availability makes it a complex function with respect to increasing forecast accuracy with an increasing number of transmission lines data. Therefore, it is proposed that multiple lines be used to analyse forecasting behaviour. In our case, we observed a slight improvement when using data from multiple transmission lines, and this remains a potential technique for improving forecasting accuracy over a larger area.

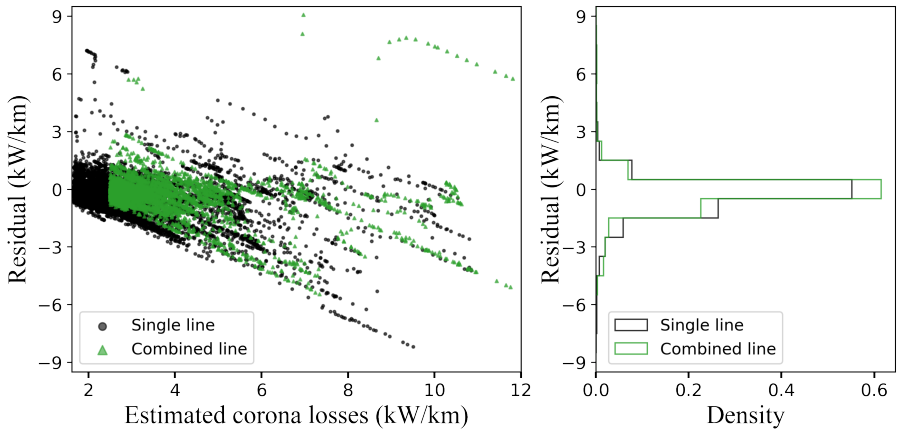


Figure 57: Residual plot of combined lines (i.e. two transmission lines) and a single line for the Estonian transmission network using the ERF model considering interpolated two-minute weather parameters considering 10 percent of higher losses [6].

Table 8: RMSE and MAE comparison of corona loss forecasts for a single line and larger network with combined data considering all weather parameters [6].

Observed Lines	ERF Model Error (kW/km)	
	RMSE	MAE
Single line	0.401	0.154
Combined lines	0.384	0.147

4.4.4 Forecasting Performance With Reduced Features

In this section, the reduced weather parameter feature is studied, where the input parameters are reduced and an analysis is carried out to understand their influence on the forecasted corona loss values using the ERF model. Using statistical graphical approaches and different feature importance techniques, the author examines the impact and their influence ranking of various weather parameters on corona losses in [3]. The aims is to build and enhance the existing research, acknowledging the valuable work that has been done in this field, and seeking to expand to different forecasting scenarios of corona losses using the reduced weather parameters, where the most important parameters were air pressure, air temperature and humidity. The minimum data requirement for weather parameters is investigated with different forecast scenarios of corona loss to eliminate unnecessary weather data, which may have very little influence. For example, in certain regions, obtaining some of the weather data can be quite challenging or might not be available in practice. Therefore, examining the performance of the corona loss forecasting model using limited weather parameter data would be interesting and valuable, providing insights and contributing to a better understanding of the impact. Two case studies are evaluated as follows: one focusing on a single transmission line approach and the other on a combined transmission line data with reduced weather parameters.

1. Single transmission line estimated corona loss (two-minute time step) with reduced features

   This section conducts a thorough analysis, examining the forecasting of corona losses using various sets of reduced weather parameters. By employing a detailed approach, we aim to uncover key insights and patterns that may not be immediately evident. The ERF regression model is trained for different sets of input features, including the full datasets, which are as follows:

   - RP_1 – Air Temperature, Air Pressure, Humidity, Precipitation, Visibility, Wind Direction and Wind Speed
   - RP_2 – Air Temperature, Air Pressure, Humidity, Precipitation, Wind Direction and Wind Speed
   - RP_3 – Air Temperature, Air Pressure, Humidity, and Precipitation
   - RP_4 – Air Temperature, Air Pressure and Precipitation
   - RP_5 – Air Temperature, Air Pressure and Humidity.

   In relation to the five regression models with distinct reduced parameter inputs, the forecasting behaviour of corona loss is examined to better understand how each model operates and the unique advantages they offer in various contexts. To illustrate, residual prediction and its histogram as depicted in Fig. 58, considering the top 10% of higher corona losses. In addition, to evaluate the performance of the

regression model, the error metrics RMSE and MAE for the entire test dataset are presented in Fig. 59.

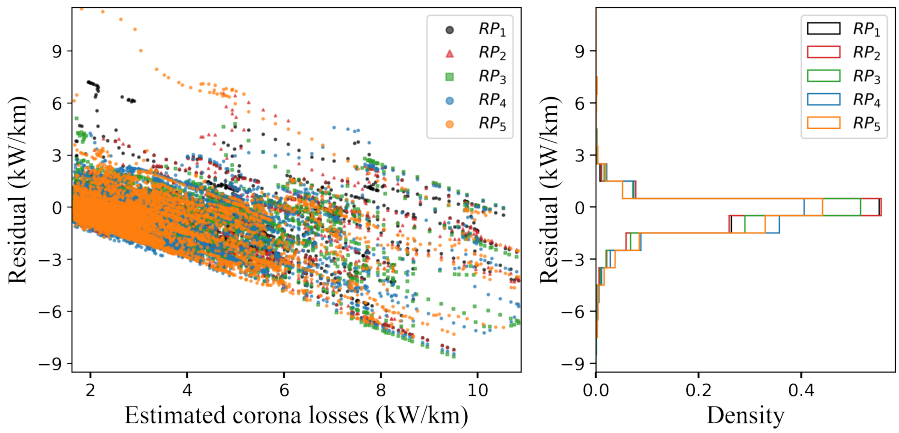


Figure 58: Residual plot of Corona loss forecasting for the Estonian transmission network using ERF model for single-line data with different reduced interpolated weather parameters [6].

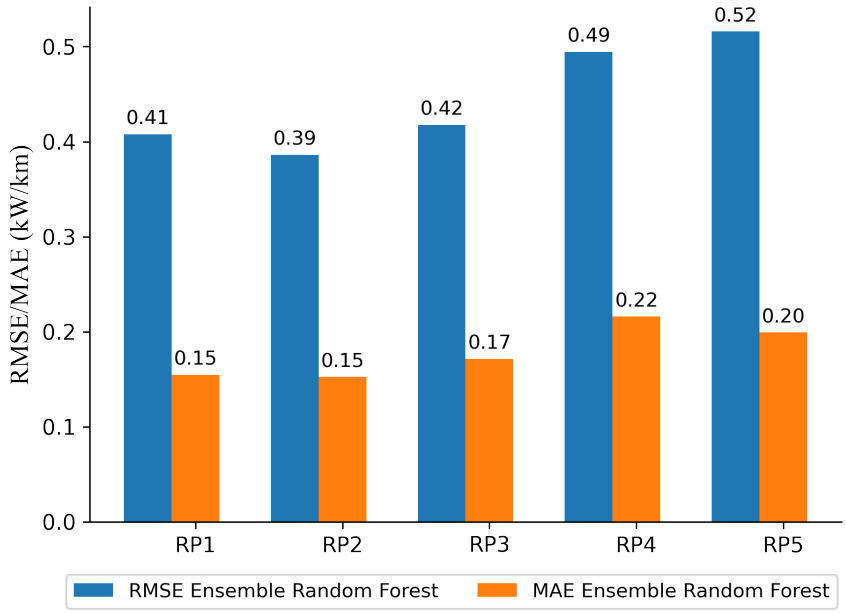


Figure 59: RMSE and MAE of ERF model for forecasting corona loss in the Estonian transmission network using single-line data with different reduced interpolated weather parameters [6].

The results suggest that identical performance is observed for the first two sets of features, i.e., when only visibility is removed, with little difference between them. However, as more features are removed, the error metric values increase significantly, degrading performance and model accuracy. The accuracy is still relatively high and has not significantly degraded, which is encouraging for our continuous analysis efforts. Overall, it demonstrates a complex relationship, but it is clear that

the air temperature, air pressure, humidity, and precipitation have the greatest impact, resulting in changes with a higher percentage. Though wind speed and direction also have some impact with lower accuracy, and it's very useful. Using a minimum set of weather parameters, such as air temperature, air pressure, and humidity, the machine learning regression model can make more accurate forecasts of higher corona losses, which provides valuable insights.

2. Combined two transmission lines corona losses (two-minute time steps) with reduced features

   To enhance our understanding and provide additional clarity, the ERF regression model is further evaluated for the combined transmission line data with reduced important weather parameters – air temperature, air pressure, and humidity. The residual prediction with a histogram is plotted and shown in Fig. 60 for the combined transmission line dataset, taking into account all weather parameters and important reduced parameters, the top 10% of corona losses. From the histogram, it is clearly visible that the prediction accuracy is decreased, which is also reflected with increasing error metrics RMSE and MAE presented in Table 8 with a smaller set of input weather data, i.e., reduced important weather parameters. However, the prediction continues to hold significant relevance and valuable insight, demonstrating its effectiveness even with reduced weather features for monitoring purposes, benefiting forecasting scenarios, i.e., training a machine learning model is still feasible.

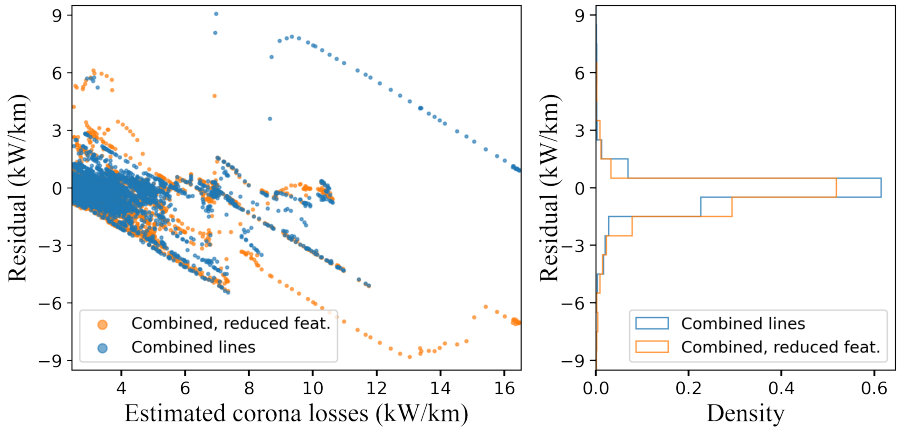


Figure 60: Residual plot of combined line data using ERF model considering all weather parameters and reduced weather parameters [6].

Table 9: RMSE and MAE comparison of corona loss forecasts for a larger network with combined data considering all weather and reduced parameters [6].

Feature Set	ERF Model Error (kW/km)	
	RMSE	MAE
RP_1	0.384	0.147
RP_5	0.489	0.193

4.5 Incorporating Corona Loss Forecasting Scenarios into Operational Planning Activities of Transmission System Operators

Monitoring corona losses is a crucial task for Transmission System Operators (TSOs) as it enhances the reliability, efficiency, and performance of power network asset management. By incorporating forecasting scenarios for day-ahead or intra-day operational planning, TSOs can effectively monitor corona loss situations and evaluate conditions, especially in Extra High Voltage (EHV) and Ultra High Voltage (UHV) networks, where corona losses are more significant. Various scenario-based forecasting analyses will assist TSOs in assessing trade-offs in the energy market, balancing the need to minimize losses, maintain voltage stability, and adhere to operational constraints. This encourages industrial firms to utilize their energy loss data, including PMU measurements, weather, and various other data.

This research has focused on the feasibility of forecasting corona losses based on weather monitoring. There could be potential for developing a practical implementation for short term forecasts in transmission networks, however, this would require considerable additional work. Any such work should start from a review of data availability, e.g. what kind of PMU measurements are being recorded and how much data has been gathered, which sources of weather monitoring data are available and how much historic data has been stored, and what kind of weather forecasts could be used as inputs to corona losses forecasting. If suitable monitoring devices are not available then they should be installed. If sufficient amount of historic data is not available, a data gathering effort has to be started.

Once suitable data has been gathered, there should be a preliminary study to validate if the WAMS data can be used to estimate corona losses on the given transmission lines. Relevant parts of the presented work can be used as reference for that task. This could be followed by establishing computational tools to process the estimated corona losses data for all applicable transmission lines for the available time periods.

After forming a suitably large dataset of historic corona loss data and corresponding weather monitoring data, a forecasting model could be trained. The presented research has explored a few different types of models and their configurations, but different methods may produce better results on larger or otherwise different datasets. Initially, a forecasting model could be trained on historic monitoring data. This could be followed up by testing the forecasting models not on weather monitoring data but recorded weather forecast data. Based on that a truer test can be carried out to see what accuracy could be achieved in practice.

Finally, when sufficiently good predictive models have been trained, a forecasting framework should be set up. This would handle the data gathering, inference, potential model updating, and follow-up testing. It would allow the network operator to continuously gather the necessary inputs and receive the forecasts.

4.6 Intermediate Summary

In this Chapter, the main objective was to discuss the forecasting model developed for corona losses in high-voltage lines using the PMU data and various input weather parameters collected along the overhead transmission lines. To achieve the goal, different forecasting scenarios were considered using different machine learning models. Here, three prominent and effective machine learning techniques, i.e., ensemble random forest, multi-linear regression, and XGBoost regression model, which can greatly enhance data analysis and predictive modelling to forecast corona losses, were used. The machine learning models were trained for different scenarios, where the testing results suggest

that the comparative evaluation of regression models exhibited varying levels of performance. In conclusion, the evaluation results showed that the ERF model outperformed all other regression models assessed, demonstrating its superior predictive capabilities to forecast corona losses. This indicates that the ERF model is particularly effective in capturing the real-time complexities of weather data, making it a valuable tool for future analyses of corona loss. To test the effectiveness, error metrics RMSE and MAE, along with the residual graph, are plotted.

While there are some limitations due to the complex nature of varying weather conditions, the forecast results can be viewed as valuable opportunities for monitoring and improving future research. By addressing the challenges posed by weather conditions, it is possible to develop innovative solutions using different machine learning algorithms to forecast corona losses in the electrical power networks, where corona losses are currently ignored. The research effectively illustrates the significant findings for short term corona loss forecasting – a few hours to a few days ahead. Such a short-term projection of corona loss and insights gathered through rigorous analysis of the dataset, and employing statistical graphs using different machine learning methodologies, are the key highlights and trends improving the operation of the electrical network.

The following four different scenarios were considered in order to investigate the performance of the regression model in forecasting corona loss, i.e., analysing data with different time resolution, applying seasonal models, combining the corona losses of two transmission lines, and using reduced sets of weather parameters. The first initial case presented a compelling illustration of the complexities involved in different time-step models using interpolating weather data to forecast corona losses, which shows improving accuracy of the machine learning models. The second case demonstrates forecasting scenarios for different seasonal conditions. The third case pertains to a complex situation involving more than one transmission line analysis, i.e., combined transmission, which will help us to gain valuable insight to monitor corona losses in larger power systems networks. Lastly, in the final case study, which presents reduced weather parameter analysis, it helps us draw meaningful insights, illustrating the possibility of using less input data with fair resolution and accuracy.

The key takeaway from this Chapter is that, by examining the weather data and considering various perspectives, we can implement a forecasting model in practice using machine learning to forecast corona losses in overhead transmission lines. The best case performance was an error in the range of ± 0.5 kW/km for 98 % of the test data points and a root mean squared error of 0.16 kW/km. This method has the potential to significantly improve outcomes by leveraging new machine learning techniques and comprehensive strategies, promising forecasting of corona losses for larger areas using measurement data from more transmission lines, which can be further used in an in-process dynamic monitoring system using short-term weather forecasting as an input. Future work could lead to significant improvements in the model accuracy using advanced methods and possibly look into the interdependence of sequential time points, e.g., by applying time series methods.

Conclusions and Further Work

This thesis discusses the research work related to forecasting of electrical network corona losses in modern power systems using synchronized phasor measurements and weather station data as inputs. All of the hypotheses of this thesis have been confirmed and proved to be possible and applicable. Within the framework of this thesis, fundamental knowledge on the dependency of corona losses on weather parameters has been presented, along with the development of an approach to estimate corona losses based on machine learning. In general, the results presented in the thesis are directly applicable to transmission system operators, enabling them to reduce costs related to power loss purchases, facilitate more detailed network loss monitoring, and increase system reliability through more effective network operational planning.

Corona discharge is a phenomenon that occurs when air is ionized around a high-voltage-carrying conductor exposed to a high electrical field, causing a violet glow, a humming noise, and the release of ozone gas. The energy loss in the electrical network due to the corona discharge phenomenon is termed corona loss. This loss depends on the weather and has a complicated relationship. In view of this, it is of interest to assess the impact of weather and develop a more sophisticated approach to estimating losses for greater accuracy and reliability. Within the framework of understanding the corona and its related losses, a series of experiments were conducted to analyse and visualise corona discharges. During the research, various methods were implemented to estimate the corona discharge, with the optical method proving to be the most effective. This method involves capturing corona discharge images with a camera to visually detect corona using image processing techniques. The test objects for the experiments included spherical electrodes, needles-to-planes, and bare conductors. After assessing the corona as a phenomenon, the research continued with a more detailed assessment of the dependence of corona losses on various weather parameters.

To assess corona losses and their dependency on weather parameters, two years of synchronized phasor measurement data from the Estonian transmission system operator were used. The corresponding weather parameter measurements were obtained from the Estonian Environmental Agency. A comprehensive analysis that included robust statistical analysis, various feature importance methods, e.g. MI, Tree classifier, XGBR, Boruta techniques, and detailed linear regression models to evaluate the impact of various weather parameters on corona losses, was performed. Based on the analysis it can be concluded that the identified relationships are complex, and this consequently suggests that multiple interdependent factors contribute to the occurrence of corona losses. The following highlights the key weather parameters that are the most and least influential on corona losses, as outlined below.

- Air Temperature (influence level: high) – Lower air temperature forms icing sharp edge, ice crystals, and heavy snowfall on the conductor surface, which strengthens corona losses in an overhead transmission line.
- Air Pressure (influence level: high) – It shows a complex relation where high air pressure accompanied by low winter temperatures result in higher corona losses. However, in summer, it shows lower corona losses values. In low-pressure regions, high altitude increases audible noise, affecting corona loss.
- Humidity (influence level: high) – Higher moisture levels increase ion availability and their movement, which decreases the corona inception voltage. This effect is further amplified in the presence of atmospheric phenomena, such as fog, dew, or accumulated surface pollutants, which increase air conductivity.

- Precipitation (influence level: medium) – It depends upon the season, where precipitation, considering weak/strong rain or snow, should be studied. In general, precipitation influences the surface condition of conductors and reduces the air's dielectric strength, leading to corona.
- Wind Speed (influence level: low–medium) – Higher wind speed disperses space charge and ionized air particles, which may either slightly suppress or intermittently intensify corona discharge; whereas lower wind speed reflects a higher corona loss value. The overall influence is secondary and variable depending on several environmental factors.
- Wind Direction and Visibility (influence level: low) – The wind direction correlation is weak, as a stable and uniform distribution of corona loss with varied angles was seen. Similarly, in the case of visibility, higher corona was observed with decreasing temperature. However, it indirectly and briefly influences corona losses unless accompanied by moisture-related effects.

Another aspect of interest was to understand the level of corona losses in case of different aged transmission lines. For the analysis, two transmission lines with an age difference of 50 years were assessed. Seasonal variability was also assessed in the analysis. The results indicate that corona losses tend to be higher on the older transmission line, which is likely a combination of weather effects on ageing and different design configurations.

Optimal operation of the electrical network includes the forecasting of network losses in an adequate manner. The reason is that these losses are purchased from the electricity market, and therefore, it is of interest that the loss estimation is made as accurately as possible. Generally, the forecasting of network losses can be divided into two parts, i.e., forecasting losses related to energy transmission through network elements and then forecasting losses related to corona. Forecasting of the first is relatively easy as it depends on network elements and their parameters, which are quite well known. More challenges are related to the forecasting of corona losses, as there is no straightforward approach that can be implemented. Therefore, machine learning based approaches were implemented and studied within the framework of this thesis. The general objective was to implement machine learning-based approaches to improve the accuracy of forecasting network losses. During the analysis, various approaches were assessed. In addition, different scenarios were considered, e.g., analysing data with different time resolution, applying seasonal models, combining the corona losses of two transmission lines, and reducing sets of weather parameters. The primary conclusion drawn from the findings is that the forecasting of corona losses can be effectively implemented with high accuracy through the application of machine learning techniques. In the most favourable scenarios, the forecasting error was comprised within $\pm$0.5 kW/km for 98 % of the analysed data points, yielding a root mean squared error of 0.16 kW/km. Therefore, it can be concluded that the approach presented in this thesis has the potential to significantly improve the outcome of estimating losses by leveraging new machine learning techniques and comprehensive strategies.

Further Work

Future work could focus on improving the accuracy of loss forecasting using more advanced machine learning techniques. It could also be interesting to assess the interdependence of sequential time points by, for example, applying time series methods. An-

other interesting topic is the applicability of the proposed method to larger areas using measurement data from multiple transmission lines.

In addition, the research presented in this thesis has the potential to be expanded to explore the interdependency between corona losses and higher harmonics. This could provide valuable information for transmission system operators to understand alternative aspects when analysing network operation and determining the level of network losses.

List of Figures

1 Connection between main domain of the corona phenomenon and corona losses monitoring, forecasting and decision making, research questions and publication used in the thesis. 14
2 Experimental setup to test the sensitivity of the instruments [2]. 19
3 Experimental set-up to visualize corona discharge. 19
4 Flowchart of image processing method to detect corona discharge using high voltage setup. 20
5 Intensity of the corona discharges calculated from the images taken with the DSLR camera for the sphere-plane geometry with the sphere placed 8 cm above the ground plane for different 3 mm spheres for various supplies [1]. 21
6 Intensity of the corona discharges calculated from the images taken with the DSLR camera for the sphere-plane geometry with the sphere placed 8 cm above the ground plane for different spheres under AC supply [1]. 21
7 Aged conductor corona discharge at 50kV AC supply 22
8 New conductor corona discharge at 50kV AC supply 22
9 The framework of estimating corona loss and weather parameters selection using feature selection. 28
10 Map of North-Eastern Estonia depicting transmission line (solid blue line) and respective weather stations (black dots) [3]. 29
11 Probability density distribution of weather parameters from single stations and the average of multiple stations [3]. 30
12 Tower configuration of the transmission line [3]. 31
13 Dependence between varying air temperature and corona losses [3]. 32
14 Dependence between varying air pressure and corona losses [3]. 33
15 Dependence between varying relative humidity and corona losses [3]. 33
16 Dependence between varying air density and corona losses [3]. 34
17 Dependence between precipitation and corona losses under varying air temperature [3]. 35
18 Dependence between varying wind speed and corona losses [3]. 36
19 Dependence between varying wind direction and corona losses [3]. 37
20 Dependence between visibility and corona losses [3]. 37
21 Dependence between conductor temperature and corona losses under varying air temperature [3]. 38
22 Dependence between voltage and corona losses [3]. 39
23 Dependence between different seasons and corona losses [3]. 40
24 Dependence between weather parameters and corona losses using Pearson correlation. 40
25 Dependence between weather parameters and corona losses using Kendall correlation. 41
26 Dependence between weather parameters and corona losses using Spearman correlation. 41
27 Weather parameters importance score and ranking using feature importance methods for 10 mins and hourly data 2018-19 [3]. 42
28 Dependence between air temperature and corona losses, hourly data [3]. . 43
29 Tower configuration of the new line [4]. 47
30 Tower configuration of the old line [4]. 47
31 Monthly comparison of hourly corona loss for New line [4]. 49

32 Monthly comparison of hourly corona loss for Old line [4]. 49
33 Comparison of corona loss over the span of one week [4]. 50
34 Comparison of air temperature over the span of one week [4]. 50
35 Comparison of corona loss over the span of one week [4]. 51
36 Comparison of air temperature over the span of one week [4]. 51
37 Dependence of corona loss and conductor temperature for New line [4]. ... 52
38 Dependence of corona loss and conductor temperature for Old line [4]. 52
39 New line conductor temperature dependence on corona losses under varying air temperature [5]. 53
40 Old line conductor temperature dependence on corona losses under varying air temperature [5]. 54
41 New line air temperature dependence on corona losses [5]. 55
42 Old line air temperature dependence on corona losses [5]. 55
43 New line air pressure dependence on corona losses under varying air temperature [5]. 56
44 Old line air pressure dependence on corona losses under varying air temperature [5]. 56
45 New line wind speed dependence on corona losses under varying air temperature [5]. 57
46 Old line wind speed dependence on corona losses under varying air temperature [5]. 58
47 Flowchart of corona loss forecasting model using machine learning [6]. 61
48 Map of monitored 330 kV transmission line corridors. Lines are coloured blue, substations are denoted by square markers and weather stations by round dots [6]. 62
49 Network architecture of Ensemble Random Forest Regression [6]. 65
50 Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the hourly dataset [6]. 67
51 Evaluation of error metric MAE of forecasted corona loss values using machine learning models for different time steps [6]. 69
52 Residuals and their histograms comparing the performance of the ERF model's forecasting capability with different time steps [6]. 70
53 Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the 2 minutes data set [6]. 70
54 Comparison of different forecasting models and corona loss estimated using PMU data for 2 minutes data. Estimated measured values have been ordered descendingly from the highest recorded value. 71
55 Measured corona loss (a) and forecasted corona loss (b) using the ERF model, for different seasons of the Estonian transmission network with a two-minute time step of interpolated weather parameters inputs [6]. 72
56 Residual plot using the ERF model, for different seasons of the Estonian transmission network with a two-minute time step of interpolated weather parameters inputs [6]. 73
57 Residual plot of combined lines (i.e. two transmission lines) and a single line for the Estonian transmission network using the ERF model considering interpolated two-minute weather parameters considering 10 percent of higher losses [6]. 74

58 Residual plot of Corona loss forecasting for the Estonian transmission network using ERF model for single-line data with different reduced interpolated weather parameters [6]. 76
59 RMSE and MAE of ERF model for forecasting corona loss in the Estonian transmission network using single-line data with different reduced interpolated weather parameters [6]. 76
60 Residual plot of combined line data using ERF model considering all weather parameters and reduced weather parameters [6]. 77

List of Tables

1 Weather factors influencing corona discharge in an overhead transmission line 16
2 Factors influencing corona discharge in an overhead transmission line 16
3 CIV Observed with Electrical PD Detector Method and the Camera [1]..... 20
4 Overhead transmission line design parameters and measured value of the line impedance [3]..... 30
5 Design Parameters of Transmission line [4]. 48
6 Hyperparameters tuned for ERF Regressor 65
7 RMSE and MAE comparison of corona loss forecasts for different seasons using a two-minute time step model [6]. 73
8 RMSE and MAE comparison of corona loss forecasts for a single line and larger network with combined data considering all weather parameters [6]. 75
9 RMSE and MAE comparison of corona loss forecasts for a larger network with combined data considering all weather and reduced parameters [6]... 77

References

[1] P. K. Gupta, J.-R. R. Ruiz, K. Tuttelberg, and J. Kilter, "Quantification of corona discharge intensity applied to sphere-plane configurations," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2025.

[2] P. K. Gupta, J.-R. R. Ruiz, P. Casals-Torrens, J. T. i Garvín, K. Tuttelberg, and J. Kilter, "Sensitivity analysis of corona discharges measuring instruments using different electrodes and high voltage supplies," *IEEE Sensors Letters*, vol. 8, no. 12, 2024.

[3] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power and Energy Systems*, vol. 155, p. 109537, 2024.

[4] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023.

[5] P. K. Gupta, K. Tuttelberg, and J. Kilter, "The weather impact on corona losses of 330 kV aging transmission lines," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE), Dubrovnik, Croatia*, pp. 1–5, 2024.

[6] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Forecasting corona losses on high voltage transmission lines using machine learning," *IEEE Transactions on Power Delivery*, vol. 40, no. 5, pp. 2696–2705, 2025.

[7] F. W. Peek, "The law of corona and the dielectric strength of air," *Proceedings of the American Institute of Electrical Engineers*, vol. 30, no. 7, pp. 1485–1561, 1911.

[8] G. Faccioli, "Tests of losses on high tension lines," *Proceedings of the American Institute of Electrical Engineers*, vol. 30, no. 1, pp. 99–117, 1911.

[9] F. Peek, "The law of corona and dielectric strength of air—ii," *Proceedings of the American Institute of Electrical Engineers*, vol. 31, no. 6, pp. 1085–1126, 1912.

[10] F. Peek, "Comparison of calculated and measured corona loss curves," *Transactions of the American Institute of Electrical Engineers*, vol. 34, no. 1, pp. 269–278, 1915.

[11] F. Peek, "Voltage and current harmonics caused by corona," *Journal of the American Institute of Electrical Engineers*, vol. 40, no. 6, pp. 455–461, 1921.

[12] F. Peek, "The law of corona and the dielectric strength of air-iv the mechanism of corona formation and loss," *Transactions of the American Institute of Electrical Engineers*, vol. 46, pp. 1009–1024, 1927.

[13] J. S. Carroll and B. Cozzens, "Corona loss measurements for the design of transmission lines to operate at voltages between 220 kV and 330 kV," *Transactions of the American Institute of Electrical Engineers*, vol. 52, no. 1, pp. 55–62, 1933.

[14] J. S. Carroll and M. M. Rockwell, "Empirical method of calculating corona loss from high-voltage transmission lines," *Transactions of the American Institute of Electrical Engineers*, vol. 56, no. 5, pp. 558–565, 1937.

[15] W. Peterson, B. Cozzens, and J. S. Carroll, "Desert measurements of corona loss," *Electrical Engineering*, vol. 69, no. 10, pp. 907–912, 1950.

[16] A. Timascheff, "Effective dielectric constant of the atmosphere during snowstorms and its influence upon corona losses," *IEEE Transactions on Power Apparatus and Systems*, vol. 83, no. 5, pp. 492–495, 1964.

[17] J. J. Cladt and C. H. Gary, "Predetermination of corona losses under rain: experimental interpreting and checking of a method to calculate corona losses," *IEEE Transactions on Power Apparatus and Systems*, no. 5, pp. 853–860, 1970.

[18] Y. Liu, S. Chen, and S. Huang, "Evaluation of corona loss in 750 kV four-circuit transmission lines on the same tower considering complex meteorological conditions," *IEEE Access*, vol. 6, pp. 67427–67433, 2018.

[19] Y. Solovyev, I. Gutman, Ø. Byrkjedal, and B. H. Thorsteinsson, "Practical methodology for calculation of corona losses induced by hoarfrost using operational measurements," *CIGRE Sci. Eng.*, vol. 10, pp. 77–87, 2018.

[20] A. Diachenko, K. Sidorov, A. Hussein, and Y. Kononov, "Investigation of online corona losses as a function of operating voltage based on PMU data," in *2019 16th Conference on Electrical Machines, Drives and Power Systems (ELMA)*, pp. 1–5, IEEE, 2019.

[21] K. Lahti, M. Lahtinen, and K. Nousiainen, "Transmission line corona losses under hoar frost conditions," *IEEE Transactions on Power Delivery*, vol. 12, no. 2, pp. 928–933, 1997.

[22] M. Pfeiffer and C. M. Franck, "Impact of conductor surface type and rain intensity on HVDC corona losses," *IEEE Transactions on Power Delivery*, vol. 30, no. 5, pp. 2284–2292, 2015.

[23] G. He, Q. Hu, L. Shu, X. Jiang, H. Yang, D. Yang, and R. Sundararajan, "Impact of icing severity on positive corona generated audible noise characteristic of rime ice-covered conductor," *IEEE Transactions on Industry Applications*, vol. 55, no. 5, pp. 5269–5276, 2019.

[24] P. Bleuler, S. Hedtke, and C. Franck, "Corona performance of DC overhead lines in outdoor experiments during wet weather transitions and under varying humidity," *CIGRE Digital-E-Session 2020: Papers and Proceedings*, 2020.

[25] X. Bian, L. Wang, Y. Liu, Y. Yang, and Z. Guan, "High altitude effect on corona inception voltages of DC power transmission conductors based on the mobile corona cage," *IEEE Transactions on Power Delivery*, vol. 28, no. 3, pp. 1971–1973, 2013.

[26] F.-C. Lu, S.-H. You, Y.-P. Liu, Q.-F. Wan, and Z.-B. Zhao, "Ac conductors' corona-loss calculation and analysis in corona cage," *IEEE Transactions on Power Delivery*, vol. 27, no. 2, pp. 877–885, 2012.

[27] T. Sugimoto, "DC corona loss of experimental transmission line at shiobara test station," *IEEE Transactions on Power Apparatus and Systems*, no. 4, pp. 1440–1446, 1978.

[28] S. Hedtke, M. Pfeiffer, M. Gobeli, P. Bleuler, R. Braeunlich, and C. M. Franck, "Setup of an outdoor hybrid AC/DC test line for corona measurements," in *VDE High Voltage Technology 2018; ETG-Symposium*, pp. 1–6, VDE, 2018.

[29] S. Huang, L. Ha, and Y. Liu, "The altitude correction term of the corona loss on alternating-current transmission lines," *Energy Reports*, vol. 9, pp. 2149–2152, 2023.

[30] M. M. Saied and E. A. Oufi, "An assessment of the harmonic pollution due to line corona," *Electric Machines and Power Systems*, vol. 21, no. 1, pp. 127–139, 1993.

[31] W. Gao, D. Ding, and W. Liu, "Research on the typical partial discharge using the uhf detection method for gis," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2621–2629, 2011.

[32] J.-R. Riba and P. Bas-Calopa, "Use of dslr and sonic cameras to detect and locate high-voltage corona discharges," *Sensors*, vol. 22, no. 19, p. 7250, 2022.

[33] J. Carpenter, J. Kresge, and C. Musick, "Ultrasonic corona detection in transformers," *IEEE Transactions on Power Apparatus and Systems*, vol. 84, no. 8, pp. 647–651, 2009.

[34] H. I. Uckol and S. Ilhan, "Corona discharge modes and their detections under DC and AC voltages," *IEEE Sensors Journal*, 2024.

[35] D. S. Prasad and B. S. Reddy, "Digital image processing techniques for estimating power released from the corona discharges," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 1, pp. 75–82, 2017.

[36] H. I. Uckol and S. Ilhan, "Dc corona discharge mode identification based on the visible light images via the yolov8," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[37] J.-R. Riba, "Linking digital image intensity to carrier density in low-pressure corona discharges," *Sensors and Actuators A: Physical*, vol. 359, p. 114474, 2023.

[38] J.-R. Riba, A. Morosini, and F. Capelli, "Comparative study of AC and positive and negative DC visual corona for sphere-plane gaps in atmospheric air," *Energies*, vol. 11, no. 10, p. 2671, 2018.

[39] P. Bas-Calopa, J.-R. Riba, and M. Moreno-Eguilaz, "Measurement of corona discharges under variable geometry, frequency and pressure environment," *Sensors*, vol. 22, no. 5, p. 1856, 2022.

[40] "IEC Standard 60270: High-voltage test techniques: partial discharge measurements," *International Electrotechnical Commission*, pp. 13–31, 2000.

[41] J.-R. Riba, Á. Gómez-Pau, and M. Moreno-Eguilaz, "Insulation failure quantification based on the energy of digital images using low-cost imaging sensors," *Sensors*, vol. 20, no. 24, p. 7219, 2020.

[42] M. Asprou and E. Kyriakides, "Identification and estimation of erroneous transmission line parameters using PMU measurements," *IEEE Transactions on Power Delivery*, vol. 32, no. 6, pp. 2510–2519, 2017.

[43] K. Tuttelberg and J. Kilter, "Estimation of transmission loss components from phasor measurements," *International Journal of Electrical Power & Energy Systems*, vol. 98, pp. 62–71, 2018.

[44] K. Tuttelberg, M. Löper, and J. Kilter, "Correcting systematic errors in corona losses measured with phasor measurement units," *IEEE Transactions on Power Delivery*, vol. 34, no. 6, pp. 2275–2277, 2019.

[45] K. Tuttelberg and J. Kilter, "Uncertainty propagation in PMU-based transmission line monitoring," *IET Generation, Transmission & Distribution*, vol. 12, no. 3, pp. 745–755, 2018.

[46] P. Heroux, P. S. Maruvada, and N. G. Trinh, "High voltage AC transmission lines: Reduction of corona under foul weather," *IEEE Transactions on Power Apparatus and Systems*, no. 9, pp. 3009–3017, 1982.

[47] H. Kirkham, "The influence of rain rate on transmission line corona performance," *IEEE Transactions on Power Apparatus and Systems*, no. 1, pp. 420–430, 1981.

[48] F. Yin, M. Farzaneh, and X. Jiang, "Laboratory investigation of AC corona loss and corona onset voltage on a conductor under icing conditions," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 3, pp. 1862–1871, 2016.

[49] F. Yin, M. Farzaneh, and X. Jiang, "Corona investigation of an energized conductor under various weather conditions," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 1, pp. 462–470, 2017.

[50] F. J. Sollerkvist, A. Maxwell, K. Rouden, and T. M. Ohnstad, "Evaluation, verification and operational supervision of corona losses in sweden," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 1210–1217, 2007.

[51] N. Vasilenko, M. Gadzhiev, I. Galiaskarov, K. Zhgun, V. Korobka, V. Ryabchenko, and Y. Sharov, "Corona losses reduction of OHL 500 kV of omsk electric power system based on signal processing of PMU," *CIGRE Paris Session*, 2020.

[52] I. Pavičić, N. Holjevac, I. Ivanković, and D. Brnobić, "Model for 400 kV transmission line power loss assessment using the PMU measurements," *Energies*, vol. 14, no. 17, p. 5562, 2021.

[53] I. Witten, E. Frank, M. Hall, and C. Pal, "*Data Mining: Practical Machine Learning Tools and Techniques*," 4^{th} Edition, Elsevier, 2016.

[54] T. Hastie, R. Tibshirani, and J. Friedman, *"The Elements of Statistical Learning: Data Mining, Inference, and Prediction"*. 2^{nd} Edition, Springer, 2009.

[55] C. Freeman, D. Kulić, and O. Basir, "Feature-selected tree-based classification," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1990–2004, 2013.

[56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *The Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[57] M. B. Kursa and W. R. Rudnicki, "Feature selection with the boruta package," *Journal of Statistical Software*, vol. 36, pp. 1–13, 2010.

[58] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," *Noise Reduction in Speech Processing*, pp. 1–4, 2009.

[59] scikit-learn SelectKBest, "https://scikit-learn.org/dev/modules/generated/sklearn.feature_selection.selectkbest.html."

[60] P. E. Meyer, C. Schretter, and G. Bontempi, "Information-theoretic feature selection in microarray data using variable complementarity," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 3, pp. 261–274, 2008.

[61] Y. Wang, W. Cui, N. K. Vuong, Z. Chen, Y. Zhou, and M. Wu, "Feature selection and domain adaptation for cross-machine product quality prediction," *Journal of Intelligent Manufacturing*, pp. 1–12, 2023.

[62] J. R. Vergara and P. A. Estévez, "A review of feature selection methods based on mutual information," *Neural Computing and Applications*, vol. 24, no. 1, pp. 175–186, 2014.

[63] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[64] S. M. Kasongo and Y. Sun, "Performance analysis of intrusion detection systems using a feature selection method on the unsw-nb15 dataset," *Journal of Big Data*, vol. 7, no. 1, p. 105, 2020.

[65] N. K. Poona and R. Ismail, "Reducing hyperspectral data dimensionality using random forest based wrappers," in *IEEE International Geoscience and Remote Sensing Symposium-IGARSS*, pp. 1470–1473, 2013.

[66] M. B. Kursa, A. Jankowski, and W. R. Rudnicki, "Boruta–a system for feature selection," *Fundamenta Informaticae*, vol. 101, no. 4, pp. 271–285, 2010.

[67] C. Zhang, Y. Li, Z. Yu, and F. Tian, "Feature selection of power system transient stability assessment based on random forest and recursive feature elimination," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 1264–1268, IEEE, 2016.

[68] M. L. Waskom, "Seaborn: statistical data visualization," *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, 2021.

[69] M. Sobhani, A. Campbell, S. Sangamwar, C. Li, and T. Hong, "Combining weather stations for electric load forecasting," *Energies*, vol. 12, no. 8, p. 1510, 2019.

[70] F. J. Sollerkvist, A. Maxwell, K. Rouden, and T. M. Ohnstad, "Evaluation, verification and operational supervision of corona losses in sweden," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 1210–1217, 2007.

[71] D. Rodriguez, R. S. Gorur, and P. M. Hansen, "Effect of humidity on the breakdown characteristics of air in non-uniform fields at 30 kHz," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 17, no. 1, pp. 45–52, 2010.

[72] C. Gallo, J. Germanos, and J. Courtney, "The effect of humidity and temperature variations on the behavior of wire-to-plane coronas," *Applied Optics*, vol. 8, no. 101, pp. 111–119, 1969.

[73] P. S. Maruvada, R. Dallaire, P. Heroux, and N. Rivest, "Long-term statistical study of the corona electric field and ion-current performance of a±900-kV bipolar HVDC transmission line configuration," *IEEE Transactions on Power Apparatus and Systems*, no. 1, pp. 76–83, 1984.

[74] M. Khalifa and R. Morris, "A laboratory study of the effects of wind on DC corona," *IEEE Transactions on Power Apparatus and Systems*, no. 3, pp. 290–298, 1967.

[75] G. Reid and H. Vermeulen, "Effects of conductor temperature on corona inception," in *2014 49th International Universities Power Engineering Conference (UPEC)*, pp. 1–5, 2014.

[76] G. W. Corder and D. I. Foreman, *Nonparametric statistics for non-statisticians*. John Wiley & Sons, Inc., 2011.

[77] M. G. Kendall, *Rank correlation methods*. C. Griffin, 1948.

[78] "IEEE Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors," *IEEE Standard 738-2012*, pp. 1–67, 2012.

[79] B. H. Chudnovsky, *Electrical Power Transmission and Distribution: Aging and Life Extension Techniques*. CRC Press, 2013.

[80] L. Yunpeng, H. Feng, L. Chen, L. Yongshuang, T. Jian, C. Shaoshuai, L. Daran, C. Sijia, and H. Shilong, "Corona loss of the bundle conductors on EHV/UHV AC power lines under sandy and dusty conditions in high-altitude areas," *Journal of Electrostatics*, vol. 107, p. 103476, 2020.

[81] S. Ghosh, N. Ahmed, and S. Banerjee, "Impact of weather (fog) on corona loss and its geographical variation within eastern region," in *2018 20th National Power Systems Conference (NPSC)*, pp. 1–6, IEEE, 2018.

[82] H. Kirkham, "The influence of rain rate on transmission line corona performance," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 1, pp. 420–430, 1981.

[83] F. W. Peek, *Dielectric phenomena in high-voltage engineering*. McGraw-Hill Book Company, Incorporated, 1929.

[84] P. S. Maruvada, *Corona performance of high-voltage transmission lines*. Research Studies Press Baldock, UK, 2000.

[85] X. Bian, X. Zhao, J. Cao, L. Gu, L. Wang, and Z. Guan, "Aging effects of conductor surface conditions on AC corona discharge," in *2010 Annual Report Conference on Electrical Insulation and Dielectic Phenomena*, pp. 1–4, IEEE, 2010.

[86] J. C. Matthews, "The effect of weather on corona ion emission from AC high voltage power lines," *Atmospheric Research*, vol. 113, pp. 68–79, 2012.

[87] K. D. Rao, G. V. Rao, G. Ramesh, and K. Govardhanachari, "Comparison of used conductor for corona, radio interference voltage (RIV) characteristics and audible noise values with unused conductor," in *Power Electronics and High Voltage in Smart Grid*, pp. 79–91, Springer, 2022.

[88] J.-R. Riba, S. Bogarra, Á. Gómez-Pau, and M. Moreno-Eguilaz, "Experimental study of the corona performance of aged sand-cast substation connectors," *Energies*, vol. 13, no. 11, p. 2785, 2020.

[89] A. Parizad, S. Dehghan, H. Saboori, and A. Kazemi, "Transmission network augmentation planning considering the impact of corona power loss," in *2011 IEEE Trondheim PowerTech*, pp. 1–6, IEEE, 2011.

[90] J. Tulensalo, J. Seppänen, and A. Ilin, "An lstm model for power grid loss prediction," *Electric Power Systems Research*, vol. 189, p. 106823, 2020.

[91] F. Pandžić, I. Sudić, T. Capuder, and I. Pavičić, "On the practical aspects of machine learning based active power loss forecasting in transmission networks," *IET Generation, Transmission & Distribution*, vol. 18, no. 14, pp. 2452–2463, 2024.

[92] S. I. Sulakov, "Forecasting hourly corona losses applying statistical approach," in *2016 19th International Symposium on Electrical Apparatus and Technologies (SIELA)*, pp. 1–4, IEEE, 2016.

[93] C. Ding, Y. Zhou, Q. Ding, and Z. Wang, "Loss prediction of ultrahigh voltage transmission lines based on eemd–lstm–svr algorithm," *Frontiers in Energy Research*, vol. 10, p. 811745, 2022.

[94] N. Dalal, M. Mølna, M. Herrem, M. Røen, and O. E. Gundersen, "Day-ahead forecasting of losses in the distribution network," *AI Magazine*, vol. 42, no. 2, pp. 38–49, 2021.

[95] H. Huang, "Line loss prediction of distribution network based on bp neural network," *Scientific Programming*, vol. 2022, no. 1, p. 6105316, 2022.

[96] A. Söderlind, "Day-ahead grid loss forecasting: A study of linear and non-linear models when modelling electrical grid losses," 2022.

[97] I. Baran, T. Leonida, and D. Sidea, "Using numerical weather forecast to predict power losses on transmission lines," in *2013 4th International Symposium on Electrical and Electronics Engineering (ISEEE)*, pp. 1–8, IEEE, 2013.

[98] M. Rejc and M. Pantos, "Short-term transmission-loss forecast for the slovenian transmission power system based on a fuzzy-logic decision approach," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1511–1521, 2011.

[99] Estonia Environmental Agency, "https://www.ilmateenistus.ee/."

[100] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[101] E. T. Iorkyase, C. Tachtatzis, I. A. Glover, P. Lazaridis, D. Upton, B. Saeed, and R. C. Atkinson, "Improving RF-based partial discharge localization via machine learning ensemble method," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1478–1489, 2019.

[102] M. Venkateswara Reddy and R. Sodhi, "A modified s-transform and random forests-based power quality assessment framework," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 78–89, 2018.

[103] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[104] P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "An adaptive class topper optimization for scheduling thermal generation with cubic fuel cost function," in *2025 15th International Conference on Power, Energy, and Electrical Engineering (CPEEE)*, pp. 195–199, IEEE, 2025.

[105] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Adaptive chaotic class topper optimization to solve non-convex chped problems in power systems," in *2024 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 1–6, IEEE, 2024.

[106] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Class topper optimization approach to solve non-convex combined heat and power economic dispatch problems in power systems," in *2024 IEEE International Conference on Power and Energy (PECon)*, pp. 179–184, IEEE, 2024.

[107] P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "Ameliorated quantum class topper optimization to solve economic load dispatch problems," in *2024 3rd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)*, pp. 1–5, IEEE, 2024.

Acknowledgements

I would like to express my gratitude to everyone who has contributed to the creation of this thesis.

My deepest gratitude goes to my family. Without your continuous support, patience, and encouragement, writing this thesis would not have been possible. I feel a deep sense of gratitude towards my parents, who have been a part of my vision and taught me the good things that truly matter in life.

I would like to express my sincere gratitude to my supervisors, Prof. Jako Kilter and Dr. Kaur Tuttelberg, for their valuable guidance, suggestions, and moral support throughout my studies, which culminated in the completion of my PhD work. This thesis would not have been possible without their continuous encouragement and involvement.

I would like to extend my sincere thanks to my colleagues at TalTech for their valuable suggestions and support throughout my studies. I would also like to thank all of my friends for helping me to develop a sense of duty, perfectionism, and sincerity in our efforts.

I would also like to express my acknowledgement to Dr. Pau Bas and Prof. Jordi Rozer Riba Ruiz, with whom I had the opportunity to work during my stay at UPC Universitat Politècnica de Catalunya, Barcelona. Thank you for all the mentoring, guidance, and help.

Finally, I am grateful to the Ministry of Education for their financial support throughout my PhD studies as well as the Estonian Research Council, grant no: TEM-TA134, which made this research possible as part of the project “Optimal and reliable control of power systems in the framework of large-scale renewable energy”.

Abstract

Forecasting of Electrical Network Corona Losses Based on Machine Learning Methods

The availability of synchronized phasor measurements has provided more sophisticated options for the reliable and optimal control of modern electrical networks. This thesis uses these types of measurements, together with weather parameter measurements, to forecast network corona losses. The main objectives of the research were to analyse electrical network corona as a phenomenon together with consequent losses, and to examine their dependence on various weather parameters. The research also aimed to develop a framework for forecasting the corona losses, considering the applicability of synchronized phasor measurements and respective weather parameter data.

Electrical network corona as a phenomena and the respective corona losses have been studied for a long time; however, the development of sophisticated measurement and advanced analysis tools provides additional options for more accurate corona loss forecasting and performance analysis. To understand the corona and its related losses, a series of experiments was conducted to analyze and visualize the corona discharge. As a result, the effectiveness of optical methods was identified. Based on the correlation analysis of network corona losses and weather parameters, it can be concluded that correlation exists and the most influential weather parameters are air temperature, air pressure, and relative humidity. While the correlations are visible, it was shown that the dependence is complex. Different methods that were applied indicate similar rankings when defining the most influential weather parameters. From the perspective of data availability (10 minutes and hourly data), it was seen that finer resolution data may be more illustrative, but is not necessary for most analyses. In addition to the dependency analysis of the weather parameter and corona losses, a more in-depth assessment was made of the relationship between corona losses and the ageing of overhead transmission lines. The results indicate that corona losses tend to be higher on older transmission lines due to ageing of the conductor and less optimized design in the case of older lines.

Transmission network operators are responsible for coordinated control of their network, and one important aspect within this task is to accurately estimate the network losses, as these have to be purchased from the market. Network losses can be divided into multiple categories, one of which is corona-related losses, which depend on weather parameters. The objective of this research was to develop a corona losses forecasting model for the electrical network using machine learning. During the research, various machine learning algorithms were implemented, including XGBR, linear regression and ensemble random forest. In conclusion, the ensemble random forest model outperformed all other models. During the analysis, various aspects were considered, i.e., using data with different time resolution, applying seasonal models, combining the corona losses of two transmission lines, and reduced sets of weather parameters. The primary conclusion drawn from the findings is that it is possible to implement corona loss forecasting with sufficient accuracy in practice using machine learning. The results presented in this thesis can be implemented by the owners of transmission networks to improve their network losses forecasts and thereby assist in cost savings.

Keywords: ageing transmission lines, conductor surface, corona discharge, corona loss, high voltage, feature analysis, partial discharge, phasor measurement unit, real-time monitoring, synchronized phasor measurements, transmission system operator, weather dependency, weather parameters.

Kokkuvõte
Masinõppe meetoditel põhinev elektrivõrgu koroonakadude prognoosimine

Sünkroniseeritud faasimõõtmiste kasutuselevõtt võimaldab täiendavaid võimalusi tänapäeva elektrivõrkude töökindlamaks ja optimaalsemaks juhtimiseks. Selles doktoritöös esitatakse lähenemisviis, kus neid mõõtmisi kasutatakse koos ilmastikuparameetrite mõõtmistega elektrivõrgu koroonakadude prognoosimiseks. Doktoritöö peamiseks eesmärgiks on analüüsida elektrivõrgu koroonanähtust koos sellega kaasnevate kadudega ja näidata koroonakadude sõltuvust erinevatest ilmastikuparameetritest. Samuti oli uurimistöö eesmärgiks välja töötada raamistik koroonakadude prognoosimiseks, arvestades sünkroniseeritud faasimõõtmiste ja vastavate ilmastikuparameetrite andmete rakendatavust.

Elektrivõrgu koroonanähtust ja sellega seotud koroonakadusid on uuritud pikka aega, kuid tänapäeva arengud elektrivõrgu mõõtmis- ja analüüsivahendite raamistikus võimaldavad välja töötada täiendavaid lähenemisviise koroonakadude täpsemaks prognoosimiseks ja elektrivõrgu talitluse analüüsiks. Koroonanähtuste ja sellega seotud kadude mõistmiseks viidi läbi mitmeid katseid ja nende tulemusena leiti, et koroonalahenduse tuvastamiseks on kõige efektiivsemad optilised meetodid. Elektrivõrgu koroonakadude ja ilmastikuparameetrite korrelatsioonianalüüsi põhjal saab järeldada, et vastavasisuline korrelatsioon on olemas ja kõige mõjukamad ilmastikuparameetrid on õhutemperatuur, õhurõhk ja suhteline õhuniiskus. Siiski tuleb välja tuua, et kuigi vastavasisulised korrelatsioonid on nähtavad, siis vastastikused sõltuvused on keerulised. Erinevate rakendatud meetodite alusel saadud tulemused näitavad sarnast järjestust kõige mõjukamate ilmastikuparameetrite määratlemisel. Andmete kättesaadavuse seisukohast (10-minuti väärtused ja tunnikeskmised andmed) nähtus, et väiksema eraldusvõimega andmed võivad olla illustratiivsemad, kuid enamiku analüüside jaoks pole need tarvilikud. Lisaks ilmastikuparameetrite ja koroonakadude sõltuvusanalüüsile hinnati põhjalikumalt ka koroonakadude ja õhuliinide vananemise vahelist seost. Tulemused näitavad, et vanematel ülekandeliinidel on koroonakaod tavaliselt suuremad juhi vananemise ja vanemate liinide vähem optimeeritud konstruktsiooni tõttu.

Ülekandevõrguoperaatorid on otseselt vastutavad oma elektrivõrgu koordineeritud juhtimise eest ja selle ülesande üheks oluliseks aspektiks on elektrivõrgu kadude täpne prognoosimine, kuna need tuleb osta elektriturult. Elektrivõrgu kaod saab jagada mitmesse kategooriasse, milledest üks on elektrivõrgu koroonanähtusega seotud kadu, mis sõltub ilmastikuparameetritest. Selle doktoritöö üheks eesmärgiks oli välja töötada masinõppel põhinev elektrivõrgu koroonakadude prognoosimise mudel. Uurimistöö käigus rakendati erinevaid masinõppe algoritme, sealhulgas XGBR, lineaarne regressioon ja juhuslik mets. Kokkuvõtteks võib öelda, et juhusliku metsa mudel edestas kõiki teisi mudeleid. Analüüsi käigus arvestati erinevate aspektidega, st erineva ajaresolutsiooniga andmete kasutamist, hooajaliste mudelite rakendamist, kahe ülekandeliini koroonakadude kombineerimist ja ilmastikuparameetrite vähendatud komplekte. Tulemustest tulenev peamine järeldus on see, et masinõppe abil on praktikas võimalik tagada piisavalt täpne koroonakadude prognoosimine. Käesolevas doktoritöös esitatud tulemusi saavad ülekandevõrguoperaatorid kasutada oma elektrivõrgu kadude täpsemaks prognoosimiseks ja seeläbi kulude kokkuhoiuks.

Märksõnad: faasimõõteseade, ilmastikuparameetrid, ilmastikusõltuvus, juhi pind, koroonakadu, koroonalahendus, kõrgepinge, omaduste analüüs, osalahendused, reaalajas jälgimine, sünkroniseeritud faasimõõtmised, vananevad ülekandeliinid, ülekandevõrguoperaator.

Appendix 1

I

P. K. Gupta, J.-R. R. Ruiz, K. Tuttelberg, and J. Kilter, “Quantification of corona discharge intensity applied to sphere-plane configurations,” *IEEE Transactions on Dielectrics and Electrical Insulation*, 2025

Quantification of Corona Discharge Intensity Applied to Sphere-Plane Configurations

Pradeep Kumar Gupta, Jordi-Roger Riba Ruiz, Kaur Tuttelberg, Jako Kilter

***Abstract*— This paper proposes several experimental methods for the quantification of the intensity of corona discharges under alternating current (AC) and positive and negative direct current (+DC and –DC) power supplies. These methods are based on the measurement of voltage and corona current, the energy associated with the corona pulses using high-frequency probes and a high-frequency oscilloscope, the apparent charge using a partial discharge (PD) detector coupled to a coupling capacitor, and the intensity of images taken with a digital camera. The experimental results presented are based on a sphere-plane electrode. However, many other electrode geometries can be studied based on the developments made in this work. In addition, an analysis of the advantages and disadvantages of the different methods is presented.**

***Index Terms*— corona inception voltage, corona losses, high voltage, optic methods, partial discharges, visual corona.**

I. Introduction

CORONA discharges are caused by the ionization of the gas insulation surrounding an energized electrode as a result of a high intensity electric field. They produce light, electromagnetic and acoustic emissions, chemical reactions and power loss [1], [2]. Among other factors, corona loss intensity depends on applied voltage amplitude, type (AC, DC, sinusoidal or pulsed, frequency, etc.), electrode geometry, surface condition, and environmental conditions.

The sphere-plane configuration has been analyzed by many researchers for various purposes, such as analyzing the electrical stress distribution on the breakdown voltage [3], the corona discharge characteristics [4], the spark breakdown conditions [5], ion and corona current flow [6]. However, little research has been done on the corona loss in the sphere-plane geometry. The current-voltage characteristics and the effects on the corona onset voltage and current for different electrode shapes under DC supply using brass and copper electrodes were measured in [7]. The effect of gap spacing, polarity and waveform of the impulse voltage, and environmental conditions on the sparkover behavior of sphere-rod gaps was investigated in [8]. An experimental study analyzing different inter-electrode spacings using the sphere-plane geometry based on both positive and negative corona discharge at atmospheric pressure is presented in [9]. The behavior of the breakdown voltage in sphere-plane epoxy coated electrodes under AC supply has been analyzed in [10]. Sphere-plane electrode configurations were also analyzed in [11] based on experimental measurements using an ultraviolet camera and finite element method (FEM) simulations.

It is widely recognized that among the various corona and partial discharge (PD) diagnostic techniques, PD measurement based on a PD detector using either the direct or indirect circuit is an effective tool. PD testing is often used as a quality assurance test for electrical insulation in high-voltage equipment [12]. For this purpose phase-resolved partial discharge (PRPD) plots are very useful to visualize the number of discharge pulses, their location along the 360º of an AC cycle, and to classify their origin [13] from cluster plots [14]. PRPD plots have been widely used for the early identification of electrical trees in polymeric insulation systems [15], especially in insulated medium and high voltage cables [16], to assess the condition of the insulation of electrical machines [17], or in gas insulated switchgear [18], among others.

In recent years, there has been a growing interest in the detection of corona discharge by optical methods based on both ultraviolet (UV) and visible methods [19]. An experimental work to obtain the visual corona extinction voltage of spherical plane geometries under AC, +DC and –DC supplies was presented in [4] using a digital camera. Other works have proposed the use of digital single lens reflex (DSLR) [20] and sonic cameras [21] or even gas-filled tube solar blind ultraviolet (UV) sensors [22] to detect and locate corona discharges. Imaging sensors are attractive because they can be used for both AC and DC supplies, overcoming the difficulties of quantifying PD, especially under DC supply [23]. In [24] it was shown that the electrical energy associated with the discharges occurring in ethylene-tetrafluoroethylene (ETFE) insulated wires correlates with the intensity of the images taken by a digital camera, thus opening the possibility of using optical methods to determine the intensity of the corona discharges and locate the discharge

This project received funding from grant PID2023-147016OB-I00, by MICIU/AEI/10.13039/501100011033/ and by ERDF "A way of making Europe," by the European Union and from the Agència de Gestió d'Ajuts Universitaris i de Recerca-AGAUR (2021 SGR 00392). This work was also funded by the European Union and co-funded by the Estonian Ministry of Education and Science (Project No. TemTa-134).

Pradeep Kumar Gupta, Kaur Tuttelberg, and Jako Kilter are with the Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Estonia (e-mail: pradeep.gupta@taltech.ee; kaur.tuttelberg@taltech.ee; jako.kilter@taltech.ee).

Jordi-Roger Riba Ruiz is with the Department of Electrical Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain (e-mail: jordi.riba-ruiz@upc.edu).

This article has been accepted for publication in IEEE Transactions on Dielectrics and Electrical Insulation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDEI.2025.3562196

sites. However, it is of utmost importance to test the method under different environmental conditions with different electrodes. To fill this gap, this article explores the potential applications of the optical method in a high-voltage laboratory using spherical electrodes to demonstrate its reliability, accuracy, and effectiveness in detecting and localizing electrical discharges.

It is known that under certain conditions pulseless modes exist for both positive +DC and negative –DC coronas. This work shows the advantage of using digital cameras for corona detection and localization, especially for pulseless corona modes, such as the pulseless Hermstein corona glow under positive DC supply, where no discharge activity can be detected by conventional methods based on pulse detection.

There is little information in the technical literature on comparative studies to determine the intensity and power loss of corona discharges, so this paper contributes to the field. Four methods are proposed. The first method is based on the voltage-current characteristic of the corona, which allows the quantification of the corona losses and the value of the corona inception voltage (CIV). The second method is based on the measurement of the electrical energy of the discharge pulses, which also allows the determination of the CIV value. The third method quantifies the intensity of the discharges from the phase-resolved partial discharge (PRPD) pattern and the apparent charge using a PD detector, which are widely applied especially under AC supply. This method allows the CIV value to be determined as well as to count the number of PD pulses per unit time and their average apparent charge. Finally, the fourth method is based on the intensity determined from digital corona images taken with a digital camera. It is shown that the fourth method allows not only the direct localization of the exact discharge locations, but also the quantification of the intensity of the discharges even in the presence of pulseless discharges, i.e. when the methods based on the analysis of the discharge pulses may fail.

II. EXPERIMENTAL SETUP

This section describes the experimental setup used in this work to measure corona discharges, which is shown in Fig. 1.

In this work, three high voltage generators were used, one AC generator and two HVDC generators. The AC generator was a calibrated high-voltage transformer capable of delivering up to 130 kV (BK-130, 0-130 kV RMS value, Phenix Technologies, Accident, Maryland, USA), while the calibrated HVDC generators are capable of delivering up to +120 kV and –120 kV, respectively (Phenix 4120-10, 0-120 kV, Accident, Maryland, USA). The high voltage source (either AC, +DC or –DC) was connected to the test object (sphere to ground plane configuration) via a corona-free conductor.

A 150 kV partial discharge free coupling capacitor (150 kV, 1000 pF, Hua Gao, Wuhan, China) was used for PD measurement together with a PD instrument (PD Base II, Techimp-Altanova Group, Zola Predosa, Italy) and a measuring impedance or quadripole (CPL 542, Omicron, Vacaville, CA, USA). Prior to testing, a PD calibrator (PDCAL, Techimp-Altanova Group, Zola Pedrosa, Italy) was used to convert the measured voltage to equivalent charge in pC by averaging 500 calibration pulses. PD pulses were measured using the conventional indirect circuit based on the IEC 60270 [25] standard.

A high-frequency current transformer (MCT 120, 80 kHz to 40 MHz, Omicron, Vacaville, CA, USA) was used to measure the discharge current flowing from the high-voltage electrode to ground under both AC and DC supply, with the ground wire positioned inside the ferromagnetic loop of the HFCT.

A high performance oscilloscope (MDO3024, 200 MHz, 2.5 Gsamples/s, Tektronix) was used to record the voltage and discharge current waveforms simultaneously. A Matlab® code was used to determine the power and energy of the corona pulses from the oscilloscope waveforms.

The AC voltage waveform was measured using a high-voltage transformer (VKPE-36, single-phase, 0-36 kV, 600 VA, Laboratorio Electrotécnico, Cornellà de Llobregat, Spain) connected to the MDO3024 oscilloscope through a high-frequency voltage probe (RT-ZI10, 500 MHz, 1000 V, attenuation 10:1, Rohde & Schwarz, Munich, Germany).

A DSLR digital camera (EOS-70D, a 20.2 megapixel CMOS APS-C sensor with dimensions of 2.5 mm × 15 mm, Canon, Tokyo, Japan) was used to detect and locate visual corona discharges. After several tests, it was found that the maximum sensitivity was achieved in bulb mode, ISO-400 sensitivity, f/5.6 aperture, and tungsten color temperature. To increase its sensitivity to the faint light emitted by the corona discharges, 30 s long exposure photographs were taken to capture more photons emitted by the corona.

A calibrated digital thermo-hygro-barometer (D4130, Commeter, Commet System, Bezrucova, Czech Republic) was used to measure the ambient temperature, relative humidity, and atmospheric pressure during the tests.

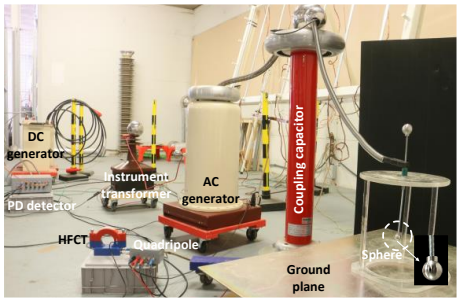


Fig. 1. Experimental setup and sphere-plane electrode.

A set of polished metal spheres (3 mm, 7 mm, 9 mm, and 14.5 mm in diameter) were used as high-voltage electrodes, suspended vertically using a metal rod support with an appropriate radius to ensure that no corona was generated below the CIV value of the sphere. To minimize the risk of dust or debris accumulation that could alter the results, the polished spheres were cleaned with isopropanol prior to testing. The lowest point of the sphere was placed 8 cm above the ground plane in order to obtain CIV values compatible with the

This article has been accepted for publication in IEEE Transactions on Dielectrics and Electrical Insulation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDEI.2025.3562196

requirements of our experimental setup.

The experimental tests in the high voltage laboratory were carried out under approximate atmospheric conditions of 20 ºC, 40 % relative humidity and 989 hPa. Great care was taken to ensure that the discharges originated from the sphere-plane electrode and not from other elements of the experimental setup. The conductor between the output of the high voltage supply and the vertical tube to which the spheres were attached was of sufficient diameter, and the connection points were protected with toroidal or spherical corona protections. Other difficulties were related to the fluctuating nature of DC discharges in air, as DC discharges are more unstable than AC discharges, and also to the lack of pulses under corona glow conditions.

The electric field E [kV/mm] at the bottom point of the different spheres studied when placed 8 cm above the ground plane was characterized as a function of the applied voltage V [kV] by means of finite element method simulations using the Comsol Multiphysics software. This results in a linear relationship $E = a + b \cdot V$, where the parameters a and b are given in Table I for the different spheres analyzed.

TABLE I
PARAMETERS OF THE LINEAR FIT OF E VERSUS V

Sphere diameter	Parameter a [kV/mm]	Parameter b [1/mm]	Coefficient of determination R^2
3 mm	$-9 \cdot 10^{-7}$	0.5349	≈ 1
7 mm	$+5 \cdot 10^{-6}$	0.2545	≈ 1
9 mm	$-6 \cdot 10^{-6}$	0.2050	≈ 1
14.5 mm	$-4 \cdot 10^{-7}$	0.1353	≈ 1

III. DETECTION AND QUANTIFICATION OF CORONA DISCHARGE INTENSITY

A. Corona Power Loss Determined from the Voltage-Current Characteristic

If the voltage-current characteristic of the corona discharges is known, it is straightforward to determine the corona loss. This section applies this approach to AC, +DC and –DC supplies.

Fig. 2 shows the equivalent circuit of a corona discharge. According to [26], the fundamental component of the corona current is in phase with the applied voltage waveform, while the other current harmonics do not contribute to the corona loss. Fig. 2 shows an equivalent circuit representing the corona loss and the capacitive current due to the capacitance of the analyzed discharge electrode. Under this assumption, it can be seen that $I_{discharge}$ is a resistive current that determines the corona losses.

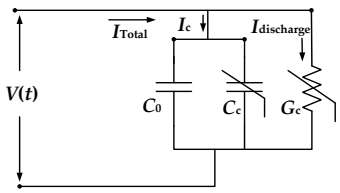


Fig. 2. Corona discharge equivalent circuit.

Note that C_0 is the geometric capacitance of the electrode, C_c is the nonlinear capacitance due to the corona discharge, and G_c is the conductance that causes the corona losses. Below the corona inception point, $C_c \approx 0$ and $G_c \approx 0$. However, above the corona inception point G_c increases very rapidly but C_c increases very slowly since any noticeable increase in C_c is well above the corona inception point. According to Fig. 2 and assuming $C_c \approx 0$, the relationship between $I_{discharge}$ (corona discharge current), I_{total} (measured current) and I_c (capacitive current due to the electrode geometry) is given by,

$$I_{discharge} = \sqrt{I_{total}^2 - I_c^2} \quad (1)$$

Under DC supply, since $I_c \approx 0$, it follows from equation (1) that $I_{discharge} \approx I_{total}$. The test object is a 3 mm diameter sphere placed 8 cm above a grounded plane (sphere-plane geometry).

Fig. 3b shows the total current and the capacitive and discharge currents under AC supply for the sphere-plane gap using a sphere of 3 mm diameter placed 8 cm above a conductive plane. In this case, a linear increase of the capacitive current with the applied voltage is observed up to 12.0 kV in RMS value (17.0 kV peak), corresponding to the corona inception voltage (CIV) value. Below this point $I_{total} \approx I_c$, and above the CIV value equation (1) must be applied to determine $I_{discharge}$, where I_c can be determined by a linear extrapolation, as shown in Fig. 3b. Similarly, the discharge current as a function of the applied voltage was measured again under +DC and –DC supply. These results are summarized in Fig. 3c. It can be seen that under DC supply $I_c \approx 0$, so that $I_{discharge} \approx I_{total}$. Under DC supply the CIV value is approximately 16.5 kV for both +DC and –DC supply.

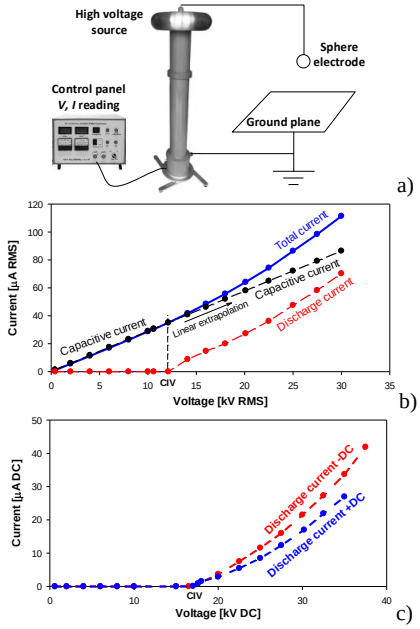


Fig. 3. 3 mm diameter sphere at 8 cm above the ground plane. Data obtained from the display of the high voltage AC generator. a) Experimental setup. Voltage and current characteristics curve under b) AC supply. c) DC supply.

This article has been accepted for publication in IEEE Transactions on Dielectrics and Electrical Insulation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDEI.2025.3562196

Next, the discharge power is obtained from the data shown in Fig. 3, which is calculated as,

$$P_{\text{discharge}} = V_{\text{applied}} \cdot I_{\text{discharge}} \quad (2)$$

Note that V_{applied} and $I_{\text{discharge}}$ are RMS values for AC supply, whereas they are DC values for +DC and –DC supplies. While the voltage is taken directly from the source display, the discharge current is calculated as indicated in Figure 3 from the current acquired by the source.

Fig. 4 shows the corona power losses for AC, +DC and –DC supplies, with similar values for the different supplies. The results presented in Fig. 4 show a similar trend of power loss due to corona for AC, +DC and –DC supplies. This method is useful for finding the CIV, which is the starting point of the corona discharge phenomenon. Although this method allows the corona activity to be detected and the corona power losses to be quantified, it does not allow the discharge points to be located.

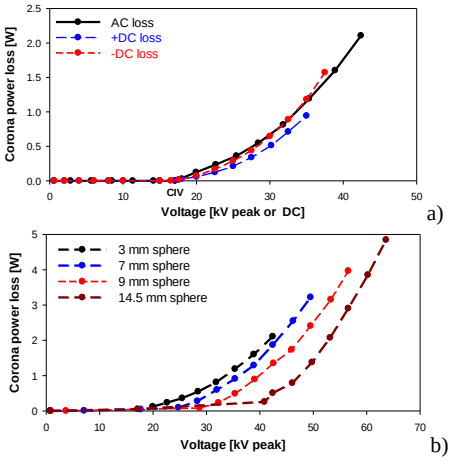


Fig. 4. Corona losses of spheres placed 8 cm above the ground plane. a) 3 mm diameter under AC, +DC and –DC supplies. b) Spheres of different diameters under AC supply.

B. Measurement of the electrical energy of the discharge pulses

This section proposes the simultaneous acquisition of the voltage and current waveforms using an HFCT for current measurement, an RT-ZI10 high-frequency voltage probe, and a high-frequency oscilloscope, as shown in Fig. 5.

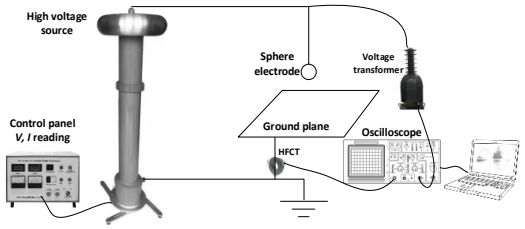


Fig. 5. Experimental setup for measuring the electrical energy of the discharge pulses.

Mathematically, the electrical energy of each discharge pulse can be calculated from the instantaneous values of the voltage and current pulses as [24],

$$E_{\text{i}} = \sum_{\text{j=1}}^{\text{duration of pulse } i} V_{\text{j,i}} I_{\text{j,i}} \Delta t \quad (3)$$

where E_{i} is the energy of discharge i-th, $V_{\text{j,i}}$ and $I_{\text{j,i}}$ are the instantaneous voltage and current values of discharge i-th, respectively, and Δt is the selected time step. The electrical energy of the N_{d} discharges occurring in one complete electrical cycle (20 ms) can be calculated as,

$$E_{\text{total}} = \sum_{\text{i=1}}^{N_{\text{d}}} E_{\text{i}} \quad (4)$$

The discharges appear as very fast pulses, with distinctive and unique patterns. Their typical duration is a fraction of a microsecond, so the acquisition system must be very fast. The complete acquisition takes one electrical period, i.e. 20 ms, consisting of 10 million samples with a sampling time of 2 ns, as shown in Fig. 6.

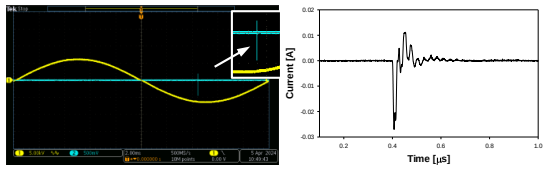


Fig. 6. 7 mm diameter sphere at 8 cm above the ground plane at 23.8 kV peak value under CIV conditions. Simultaneous acquisition of voltage and current with a sampling rate of 2 ns during 20 ms. The same procedure was used for AC, +DC and –DC supplies. a) Voltage (yellow) and current (blue) waveforms. b) Zoom on the PD current waveform.

Table II shows the number of pulses at the CIV value for the sphere-plane geometry (3 mm diameter sphere placed 8 cm above the ground plane), and the signal to noise ratio, calculated as the amplitude of the PDs divided by the standard deviation of the noise measured by the HFCT.

TABLE II
3 MM SPHERE PLACED 8 CM ABOVE THE GROUND PLANE. ELECTRICAL ENERGY OF THE SINGLE DISCHARGE PULSE AT THE CIV VALUE FOR AC, +DC AND –DC SUPPLY

Supply	Electrical energy for 20 ms from (4) [µJ]	Signal to noise ratio	Number of pulses
15.0 kVpeak AC	5.35	194.9	1
17.0 kV +DC	9.87	222.0	1
-16.5 kV –DC	5.00	122.7	1

Fig. 7a shows the electrical energy associated with the activity of the corona discharges in the case of AC supply for different spheres placed 8 cm (bottom of sphere to ground) above the ground plane. As expected, the CIV value increases as the diameter of the sphere increases. Fig. 7b shows the evolution of the number of discharge pulses in one electrical period with increasing values of the applied AC voltage for different sphere-plane geometries under AC supply. It can be

seen that the number of discharges is significantly reduced as the sphere diameter increases.

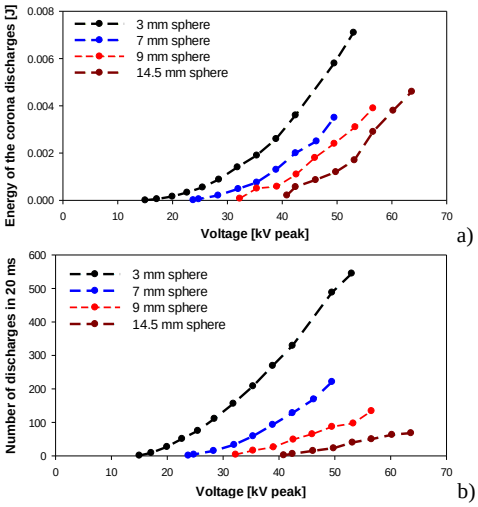


Fig. 7. Sphere-plane geometries (spheres placed 8 cm above the ground plane) in the case of AC supply. a) Calculated electrical energy of the corona discharges. b) Measured number of discharges.

Fig. 8 shows the electrical energy associated with the activity of the corona discharges as well as the number of discharges in the case of +DC and –DC supply for the 3 mm diameter sphere placed 8 cm above the ground plane.

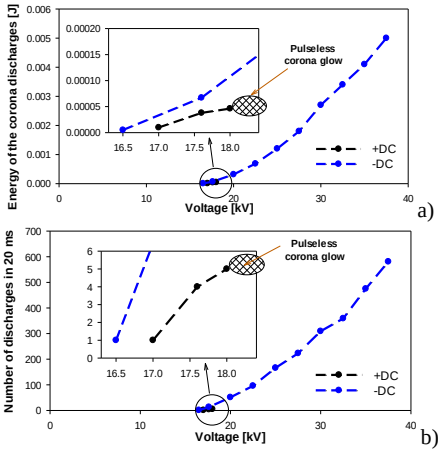


Fig. 8. Sphere-plane geometries (3 mm diameter sphere placed 8 cm above the ground plane) under +DC and –DC supply. a) Electrical energy of the discharges. b) Number of pulses.

As shown in Fig. 8, the pulseless corona glow mode occurs for +DC supply above 18 kV, and therefore the oscilloscope cannot acquire new pulses. Similar to the previous method, this method allows the detection of corona activity and the quantification of the energy associated with the corona activity during a predefined time interval, but it does not allow the localization of the discharge locations sites. In addition, it has been experimentally verified that this system does not provide any signal under pulseless corona glow conditions.

C. Discharge intensity from the PRPD pattern and the apparent charge using the PD detector

According to the technical bibliography different PD sources have distinct phase-resolved PD (PRPD) patterns that make them easily identifiable. Fig. 9 shows the patterns generated by the test platform with the PD detector using the indirect circuit based on the coupling capacitor, where the corona activity is clearly visible in the negative cycle.

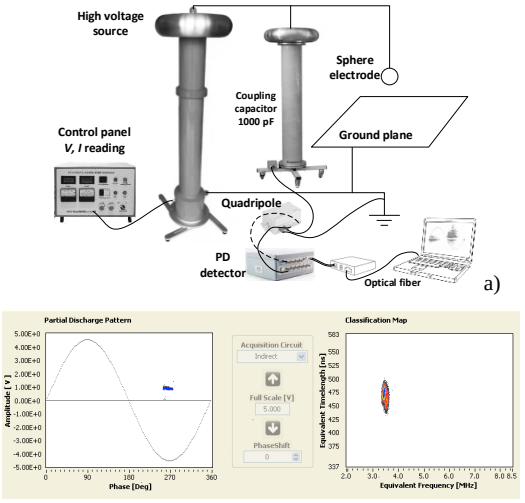


Fig. 9. a) Experimental setup. b) Results obtained with the PD detector, including the PRPD pattern, classification map, and statistical discharge information for a 7 mm diameter sphere placed 8 cm above the ground plane at the CIV value.

The cluster of points in Fig. 9b shows the pattern of the corona discharges in the negative half-cycle of the AC voltage waveform. Because the measurements in Fig. 9b are taken near the CIV value, the number of discharges in the positive half-cycle is much smaller than the number of discharges in the negative half-cycle.

The charge moved by the avalanche is often used to quantify the intensity of the discharge, because it allows an indirect estimate of the number of hot electrons, those with an energy level to ionize [27]. Since it is very difficult to calculate the true charge Q, most PD detectors provide the value of the apparent charge Q'. Therefore, the apparent charge is an essential attribute for quantifying the PD intensity, and this quantity is commonly used for testing many high-voltage devices [28] and for quantifying the discharge power [25]. The apparent charge Q' [pC] induced on the electrodes of the insulation system can be calculated based on the Shockley-Ramo theorem as [29],

This article has been accepted for publication in IEEE Transactions on Dielectrics and Electrical Insulation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDEI.2025.3562196

$$Q^{'} = Q(E/U)g \quad (5)$$

where Q [pC] is the true charge, E/U [-] is the voltage reduced background field at the position of the discharge and g is a geometry factor that depends on the shape of the discharge point and the relative permittivity of the surrounding dielectric. In general, the higher the PD charge, the higher the discharge energy.

From Fig. 10 it can be seen that beyond the CIV value, the mean apparent charge of the corona discharges occurring during the negative half-cycle increases with the applied voltage, since the charge depends on the product of the system capacitance and the applied voltage. In addition, the CIV increases with the radius of the sphere, as expected.

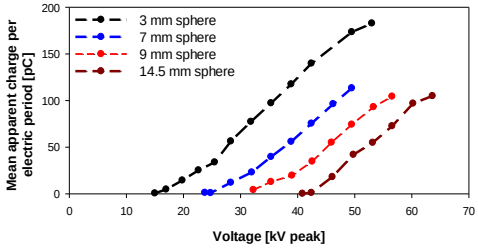


Fig. 10. Mean apparent charge of the PD pulses occurring in one electric period observed by the PD detector.

The advantage of using the PD detector is its ability to collect data for statistical evaluation and identification of different types of discharges and their characteristics. This method is also useful for determining the CIV value, since it allows the PD pulses to be visualized and counted using the PRPD pattern and the classification map. However, its main disadvantage is that in the case of pulseless corona glow, the PD detector cannot detect any trace of discharge activity [30]. This method allows the detection of corona activity and the quantification of the apparent charge due to the discharges, but it does not allow the precise localization of the discharge sites.

D. Discharge intensity determination from digital corona images

As mentioned above, corona discharges emit UV and visible light that can be used to locate the discharge points and quantify the intensity of the discharges using optoelectronic sensors. This approach has many advantages, such as direct location of the discharge points and high immunity to electromagnetic and acoustic noise. The optical method can only detect and locate the discharge if it occurs at visible locations, so PDs occurring within solid insulation materials cannot be detected by this method.

In this paper, a DSLR camera is used to detect the corona discharges, localize the discharge points, and quantify the intensity of the corona discharges. Once the corona photographs are taken, some image processing is required for optimal estimation of the intensity of the discharges, which is summarized in Fig. 11. A 30 s long exposure photograph of the sphere-plane air gap is taken with the DSLR camera in total darkness. The image is then converted into a triplet array with coordinates in the range 0-255. Next, the discharge region is identified by identifying the pixels in the image whose colors match the characteristic wavelengths of the corona light. Next, to filter out the noise in the image, pixel clusters smaller than 100 pixels are removed. Next, the RGB pixels of the areas identified as corona are transformed to grayscale as [31],

$$I_{\text{grayscale}}(i,j) = 0.299I_{\text{R}}(i,j) + 0.587I_{\text{G}}(i,j) + 0.114I_{\text{B}}(i,j) \quad (6)$$

where (i,j) are the horizontal and vertical coordinates of the image, and $I(i,j)$ is the intensity of each pixel in the image, ranging from 0-255. Finally, the intensity of the corona discharges is calculated as [32],

$$Intensity = \sum_{i=1}^{m}\sum_{j=1}^{n} I_{\text{grayscale}}(i,j) \quad (7)$$

where m and n are the number of rows and columns in the image.

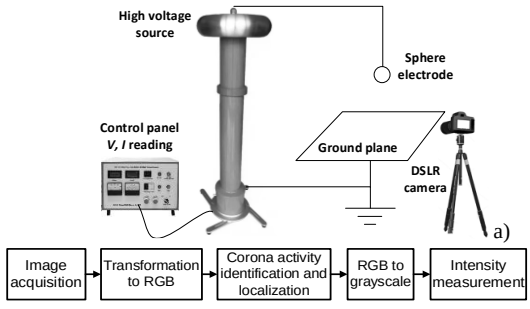


Fig. 11. a) Experimental setup. b) Flowchart of the image processing method to detect corona discharges and quantify their intensity.

Fig. 12 shows the corona images taken by the camera with the different supplies for the sphere-plane electrode geometry.

Fig. 12. Corona images of a 14.5 mm sphere 8 cm above the ground plane. a) AC supply at 31.8 kV peak. b) +DC supply at 30 kV. b) –DC supply at -30 kV.

Fig. 13.a shows the intensity of the corona images for different sphere-plane geometries under AC supply. It is observed that for all the geometries analyzed, the intensity of the images tends to increase with the applied voltage level. Similarly, Fig. 13.b shows the intensity level using the 3 mm diameter sphere under AC, +DC and –DC supply. Note that at the CIV voltage the intensity of the discharges is almost the same for the different voltage sources (AC, +DC and –DC). Fig. 13b shows that for the 3 mm sphere, the intensity of the digital corona images has the highest values under –DC supply, while it has the lowest value under +DC supply. According to the results shown in Fig. 13b, the intensity of the corona discharges under +DC supply suddenly decreases at the starting point of the corona glow phenomenon (pulseless Hermstein's glow), i.e. above 19 kV for the analyzed geometry. It should be noted

This article has been accepted for publication in IEEE Transactions on Dielectrics and Electrical Insulation. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDEI.2025.3562196

that in this region the PD detector and the current pulses acquired with the HFCT connected to the oscilloscope showed no signs of corona activity. In the other regions, under AC, and positive and negative DC supply, the intensity of the digital corona images increases almost linearly with the applied voltage.

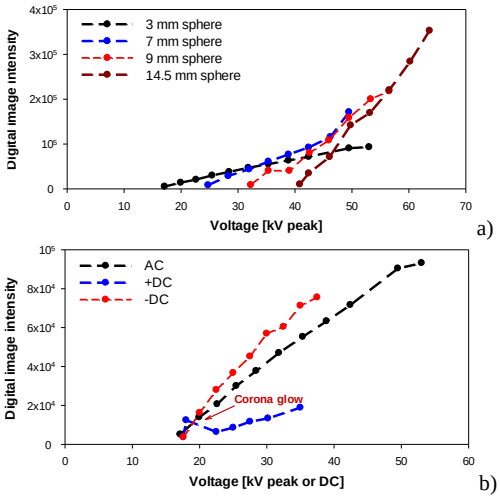


Fig. 13. Intensity of the corona discharges calculated from the images taken with the DSLR camera for the sphere-plane geometry with the sphere placed 8 cm above the ground plane. a) Different spheres under AC supply. b) 3 mm sphere under AC, +DC and –DC supplies.

The method described in this section has several advantages. First, it allows not only a straightforward localization of the discharge sites, but also a quantification of the intensity of the discharges. Second, it is a very simple method that does not require sophisticated equipment and uses affordable commercial digital cameras. Finally, the DSLR camera is able to clearly detect corona activity even in the presence of pulseless discharges, i.e. when the methods based on the analysis of the discharge pulses are unable to detect any trace of discharge activity.

IV. Discussion

This section discusses the main results of this work and compares the four methods analyzed.

The first method, based on the voltage-current characteristic, allows the determination of the CIV value, the detection of the corona activity and the quantification of the corona losses, but it does not allow the localization of the discharge sites.

The second method, based on the measurement of the electrical energy of the discharge pulses, allows the detection of corona activity, the determination of the CIV value and the quantification of the energy associated with the corona discharge pulses during a predefined time interval, but it does not allow the localization of the discharge sites. In the case of pulseless corona activity, this method cannot detect any trace of corona discharges.

The third method is based on measuring the intensity of the discharges from the PRPD pattern and the apparent charge using a PD detector. It allows the CIV value to be determined and the PD pulses to be counted using the PRPD pattern and the classification map. This method allows detection of corona activity and quantification of the apparent charge due to the discharges, but does not allow accurate localization of the discharge sites. In the case of a pulseless corona glow, the PD detector cannot detect any trace of discharge activity.

Finally, the method based on the intensity determined from digital corona images offers several advantages. First, it allows not only the direct localization of the exact discharge sites, but also the quantification of the intensity of the discharges even in the presence of pulseless discharges, i.e., when conventional equipment based on PD pulse detection is unable to detect any trace of discharge activity. This method is also easy to use and does not require sophisticated equipment, only an inexpensive digital camera. The main disadvantage of this method is that it requires total darkness [33].

Table III summarizes the technical advantages and disadvantages of the analyzed methods.

TABLE III
COMPARISON OF THE DIFFERENT METHODS

Method	Technical advantages and disadvantages
Corona power loss determined from the voltage-current characteristics	- Allows the CIV value to be found - Allows quantification of corona losses ($V \cdot I$) - Does not allow localization of corona sites - Ease of use - Low sensitivity
Measurement of the electrical energy of the discharge pulses	- Allows the CIV value to be found - Allows quantification of the electrical energy of the discharge pulses - Does not allow localization of corona sites - Insensitive under pulseless corona - Requires expensive equipment
Discharge intensity from the PRPD pattern and the apparent charge using a PD detector	- Allows the CIV value to be found - Allows quantification of apparent charge of discharges - Does not allow direct localization of corona sites - Insensitive under pulseless corona - Requires expensive equipment
Discharge intensity determined from digital corona images	- Allows the CIV value to be found - Allows quantification of discharge intensity - Allow direct localization of corona sites - Sensitive under pulseless corona - Requires simple equipment but in total darkness - Not a standard method

V. Conclusion

This paper has analyzed various experimental methods for detecting, locating and quantifying the intensity of corona discharges under AC, +DC and –DC supplies. The performance of the different methods has been evaluated, assessed and compared to provide researchers and practitioners in the field with the most convenient method to implement in practice, depending on the environment and equipment availability. Although the sphere-plane electrode configuration has been analyzed, the results presented here can be extrapolated to many

other practical geometries. Based on the comparison between visual and experimental results, the results suggest that if it is possible to work in total darkness, the use of a DSLR camera is a very sensitive, effective, affordable and simple method for corona detection and quantification of its intensity. It has the advantage over other methods of allowing direct localization of the discharge points, which is an interesting option for inspection, periodic maintenance, health monitoring, or factory quality control and design optimization. In addition, the use of a DSLR camera allows the detection, accurate localization and quantification of corona activity even under pulseless corona conditions, when the methods based on corona pulse detection are unable to detect any corona activity. In the future, it would be interesting to obtain additional experimental results for different types of electrodes and to compare the results of the different methods. It may also be interesting to integrate data from several sensors to optimize the sensitivity of corona detection, location and intensity quantification.

REFERENCES

[1] H. I. Uckol and S. Ilhan, "DC corona discharge mode identification based on the visible light images via the YOLOv8," *IEEE Trans. Instrum. Meas.*, vol. 73, 2024.

[2] G. A. David, P. O. C. Junior, F. R. L. Dotto, and B. R. Dos Santos, "New Signal Processing-Based Methodology for Optimal Feature Selection of Corona Discharges Measurement in HVDC Systems," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023.

[3] N. K. Kishore, G. S. Punekar, and H. S. Y. Shastry, "Sparkover in sphere gaps with alternating voltages and perturbed electric fields," *Annu. Rep. - Conf. Electr. Insul. Dielectr. Phenomena, CEIDP*, pp. 634–637, 2009.

[4] J.-R. Riba, A. Morosini, and F. Capelli, "Comparative Study of AC and Positive and Negative DC Visual Corona for Sphere-Plane Gaps in Atmospheric Air," *Energies*, vol. 11, no. 10, p. 2671, Oct. 2018.

[5] A. Pedersen, J. Lebeda, and S. Vibholm, "Analysis of Spark Breakdown Characteristics for Sphere Gaps," *IEEE Trans. Power Appar. Syst.*, vol. PAS-86, no. 8, pp. 975–978, Aug. 1967.

[6] V. Amoruso and F. Lattarulo, "A graphical approach to the unipolar point-to-plane corona," *J. Electrostat.*, vol. 39, no. 1, pp. 41–51, Jan. 1997.

[7] A. Lagăr, C. M. Diniş, and G. N. Popa, "Experimental analysis of direct current corona discharge," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 163, no. 1, p. 012035, Jan. 2017.

[8] P. N. Mikropoulos and C. A. Stassinopoulos, "Impulse sparkover characteristics of sphere-rod gaps," *IEE Proc. Sci. Meas. Technol.*, vol. 152, no. 4, pp. 169–174, Jul. 2005.

[9] S. I. Wais and D. D. Giliyana, "Sphere-to-Plane Electrodes Configuration of Positive and Negative Plasma Corona Discharge," *Am. J. Mod. Phys.*, vol. 2, no. 2, pp. 46–52, 2013.

[10] M. Higashiyama *et al.*, "Breakdown voltage of interelectrode gap between sphere-plane electrodes under ac voltage," *Conf. Rec. IEEE Int. Symp. Electr. Insul.*, 2010.

[11] H. Wu, J. Wang, P. Liu, Z. Peng, Z. Deng, and X. Wei, "Research on corona inception electric-field strength and critical electric-field values for valve hall fittings," in *2014 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, 2014, pp. 308–311.

[12] G. C. Stone, "Partial discharge diagnostics and electrical equipment insulation condition assessment," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 12, no. 5, pp. 891–904, Oct. 2005.

[13] A. Simard, M. Levesque, E. David, and H. Provencher, "Study on the Characterization of the Electromagnetic Probe with PRPD Pattern Analysis on Spherical Cavities in Epoxy," *2023 IEEE Electr. Insul. Conf. EIC 2023*, 2023.

[14] A. Rodrigo Mor, L. C. Castro Heredia, D. A. Harmsen, and F. A. Muñoz, "A new design of a test platform for testing multiple partial discharge sources," *Int. J. Electr. Power Energy Syst.*, vol. 94, pp. 374–384, Jan. 2018.

[15] A. Cavallini, M. Conti, G. C. Montanari, C. Arlotti, and A. Contin, "PD inference for the early detection of electrical treeing in insulation systems," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 11, no. 4, pp. 724–735, Aug. 2004.

[16] H. Kumar, M. Shafiq, K. Kauhaniemi, and M. Elmusrati, "A Review on the Classification of Partial Discharges in Medium-Voltage Cables: Detection, Feature Extraction, Artificial Intelligence-Based Classification, and Optimization Techniques," *Energies*, vol. 17, no. 5, p. 1142, Feb. 2024.

[17] D. Verginadis, T. Iakovidis, A. Karlis, M. Danikas, and J.-A. Antonino-Daviu, "A Study on the Effectiveness of Partial Discharge Models for Various Electrical Machines' Insulation Materials," *Machines*, vol. 11, no. 2, p. 230, Feb. 2023.

[18] V. N. Tuyet-Doan, Y. W. Youn, H. S. Choi, and Y. H. Kim, "Shared Knowledge-Based Contrastive Federated Learning for Partial Discharge Diagnosis in Gas-Insulated Switchgear," *IEEE Access*, vol. 12, pp. 34993–35007, 2024.

[19] D. S. Prasad and B. S. Reddy, "Digital image processing techniques for estimating power released from the corona discharges," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 1, pp. 75–82, Feb. 2017.

[20] H. I. Uckol and S. Ilhan, "Corona Discharge Modes and Their Detections under DC and AC Voltages," *IEEE Sens. J.*, May 2024.

[21] J. R. Riba and P. Bas-Calopa, "Use of DSLR and Sonic Cameras to Detect and Locate High-Voltage Corona Discharges," *Sensors*, vol. 22, no. 19, p. 7250, Sep. 2022.

[22] P. Bas-Calopa, J.-R. Riba, and M. Moreno-Eguilaz, "Measurement of Corona Discharges under Variable Geometry, Frequency and Pressure Environment," *Sensors*, vol. 22, no. 5, p. 1856, Feb. 2022.

[23] P. N. Mikropoulos and P. K. Samaras, "Surface partial discharge characteristics under repeated positive triangular high voltages (1kV/s)," *IEEE Trans. Dielectr. Electr. Insul.*, 2024.

[24] P. Bas-Calopa, J.-R. Riba, and J. A. Ortega, "Low-Pressure Optical Detection, Location, and Quantification of Electrical Discharges in Aircraft Wiring Systems," *Aerospace*, vol. 10, no. 1, p. 3, Dec. 2022.

[25] International Electrotechnical Commission, *IEC 60270:2000 High-voltage test techniques - Partial discharge measurements*, 3.0. International Electrotechnical Commission, 2000.

[26] P. S. Maruvada and S. Bisnath, *Corona in Transmission Systems: Theory, Design and Performance*. Johannesburg, South Africa: Crown Publications, 2011.

[27] A. Cavallini, G. C. Montanari, and M. Tozzi, "PD apparent charge estimation and calibration: A critical review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 17, no. 1, pp. 198–205, Feb. 2010.

[28] A. G. Ovsyannikov, S. M. Korobeynikov, and D. V. Vagin, "Simulation of apparent and true charges of partial discharges," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 6, pp. 3687–3693, Dec. 2017.

[29] B. Fruth and L. Niemeyer, "The Importance of Statistical Characteristics of Partial Discharge Data," *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 60–69, 1992.

[30] M. Kaufhold, S. S. Bamji, and A. T. Bulinski, "Optical detection of partial discharges in gas-insulated systems," *Conf. Electr. Insul. Dielectr. Phenom. (CEIDP), Annu. Rep.*, vol. 2, pp. 618–623, 1996.

[31] J.-R. Riba, P. Bas-Calopa, Y. A. Qolla, M. Pourraz, and B. Ozsahin, "Using CMOS Image Sensors to Determine the Intensity of Electrical Discharges for Aircraft Applications," *Appl. Sci.*, vol. 12, no. 17, p. 8595, Aug. 2022.

[32] J.-R. Riba, Á. Gómez-Pau, and M. Moreno-Eguilaz, "Insulation Failure Quantification Based on the Energy of Digital Images Using Low-Cost Imaging Sensors," *Sensors*, vol. 20, no. 24, 2020.

[33] P. Das, S. Chatterjee, and C. Koley, "Recurrence Plot Aided Partial Discharge Detection Framework Employing HFCT Sensor and Customized Convolutional Neural Network," *IEEE Trans. Dielectr. Electr. Insul.*, 2024.

Appendix 2

II

P. K. Gupta, J.-R. R. Ruiz, P. Casals-Torrens, J. T. i Garvín, K. Tuttelberg, and J. Kilter, “Sensitivity analysis of corona discharges measuring instruments using different electrodes and high voltage supplies,” *IEEE Sensors Letters*, vol. 8, no. 12, 2024

Sensor applications

Sensitivity Analysis of Corona Discharges Measuring Instruments Using Different Electrodes and High Voltage Supplies

Pradeep Kumar Gupta[1]*, Jordi-Roger Riba Ruiz[2], Pau Casals-Torrens[2], Jordi Talens i Garvín[2], Kaur Tuttelberg[1], and Jako Kilter[1]**

[1]*Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, 19086 Tallinn, Estonia*
[2]*Department of Electrical Engineering, Universitat Politècnica de Catalunya, 08222 Terrassa, Spain*
**Graduate Student Member, IEEE*
***Senior Member, IEEE*

Manuscript received 23 October 2024; revised 5 November 2024; accepted 19 November 2024. Date of publication 25 November 2024; date of current version 6 December 2024.

Abstract—Corona discharge testing is critical to ensuring the safety, performance and reliability of high-voltage (HV) equipment and systems. Due to the high cost of shielded laboratories, many manufacturers test their components in unshielded laboratories for applications, such as product research and development or quality control. This letter compares the sensitivity of two corona detection instruments in an unshielded laboratory. The sensitivity of an electrical detector according to IEC 60270 in combination with a measuring impedance and a coupling capacitor is compared with that of a digital single lens reflex (DSLR) camera. In addition, needle-to-plane and sphere-to-plane electrode geometries are studied under different types of high-voltage power supplies, i.e., 50 Hz ac, positive dc, and negative dc. Experimental results performed in an unshielded high-voltage laboratory show that although the DSLR camera is not a standard method for partial discharge (PD) detection, it has similar sensitivity to the conventional electrical method according to IEC 60270 in all cases analyzed, and that the digital camera behaves much better in pulseless glow corona mode.

Index Terms—Sensor applications, camera, corona discharge, electrical partial discharge (PD) detector, sensitivity.

I. INTRODUCTION

Partial discharges (PDs) occur in high-voltage (HV) insulation systems and are broadly classified as internal, surface, and corona. Corona discharges occur in the gas insulation surrounding a HV electrode. They produce some observable effects, such as chemicals, ultraviolet and visible radiation, electromagnetic waves up to the ultra high frequency band (3 MHz to 3 GHz) or higher in some insulation materials, a hissing noise and power loss. The intensity of corona discharges depends on the surface and aging conditions [1], the amplitude and frequency of the applied voltage, and the weather conditions [2]. PD detection has become increasingly important for assessing the insulation condition of high-voltage equipment [3] for indoor laboratory tests and field tests [4]. PD measurements, regulated by the IEC 62478 standard, are an effective tool for insulation assessment and are widely used to improve maintenance practices. Various researchers have proposed different methods for PD detection. These include applications such as a piezoelectric sensors [5], acoustic sensors [6], optoacoustic sensors [7], high-sensitivity fiber-optic sensors [8], wireless Internet of Things (IoT) systems to detect PDs in high-voltage lines [9], or the use of a digital single-lens reflex [digital single lens reflex (DSLR)] camera to detect corona discharge patterns for ac and dc applications [10], [11], among others.

One of the most commonly used methods for PD detection is the electrical method described in the IEC 60270 standard [12], which requires a PD instrument [13]. It is widely used in offline applications due to its sensitivity and accuracy. This method requires the electrical PD detector according to IEC 60270 (electrical PD detector or PD instrument from now on) to be connected to a coupling capacitor and a measuring impedance. In the case of field tests, the electrical PD detector can be connected to a sensitive current sensor, such as a high-frequency current transformer or a Rogowski coil, to measure the discharge pulses present in the leakage current to ground. When using the PD instrument, it is very useful to display the discharge data over the 360° of the ac cycle using phase-resolved PD (PRPD) plots. PRPD plots are also used for acoustic detection of PD patterns [14].

The corona current is known to be directly correlated to the number of photons generated in the discharge process, so the ultraviolet and visible photons generated by the corona are very important in the discharge process, both under ambient and low pressure conditions (e.g., high altitude power lines or aircraft electrical systems) [15]. Optical PD detection using imaging sensors offers some advantages, such as high immunity to electromagnetic and acoustic interference, direct localization of the discharge areas, reduced sensor size and, in many cases, low cost. Visible images combined with appropriate image processing methods have been used to detect corona activity in several papers [10], [16], [17]. It is important to detect corona discharges early, because as the intensity of corona discharges increases, so do the potential harmful effects.

Despite the growing interest in using imaging sensors for corona detection, there is a lack of works analyzing their sensitivity, so this letter contributes to this field. To this end, the sensitivity of a DSLR camera with a back-illuminated complementary metal–oxide semiconductor (CMOS) sensor is compared with that of a conventional PD instrument based on the detection of current or voltage pulses, which is probably the most widely used method for PD detection since it is described in the IEC 60270 standard [12]. Although CMOS sensors have been used previously, there is a lack of works analyzing their sensitivity.

Corresponding author: Jordi-Roger Riba Ruiz (e-mail: jordi.riba-ruiz@upc.edu).
Associate Editor: Hamza Shakeel.
Digital Object Identifier 10.1109/LSENS.2024.3504551

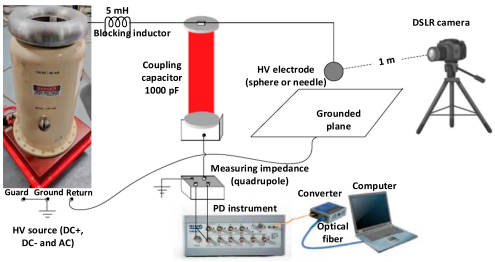


Fig. 1. Schematic of the experimental setup.

The measurements presented in this paper show that the sensitivity of the DSLR camera is better than 100 pC (apparent charge). Two commonly used high-voltage electrodes, such as the needle-to-plane and the sphere-to-plane configurations are analyzed under three types of power supplies, i.e., power frequency ac, negative direct current(dc-) and positive direct current (dc+) supplies. This comparison shows that the imaging sensor, has a high sensitivity, comparable or better than that of the PD instrument, when the corona discharges are measured in an unshielded laboratory. Note that unshielded labs are more common due to the high investment required for shielded laboratories, but add additional difficulties to measuring low PD levels due to higher noise levels. In addition, the images taken by the DSLR camera allow an accurate localization of the corona discharge regions.

II. EXPERIMENTAL SETUP

In this section, we describe the experimental setup used to detect the corona discharges generated by two high-voltage electrodes, i.e., the needle-to-plane and sphere-to-plane configurations. Fig. 1 shows a schematic of the experimental setup including the HV source, the instrumentation used and the electrodes.

Since this letter deals with ac, dc−, and dc+ corona, three HV sources (ac, dc−, and dc+) are required to perform the experimental tests. For this purpose, a calibrated high-voltage ac generator (BK-130, variable output voltage from 0 to 130 kV-rms, voltage accuracy 1%, 50 Hz, of reading, Phenix Technologies, Accident, MD, USA) and two dc generators of positive and negative polarity, respectively (4120–10, variable output voltage from 0 to 120 kV, voltage accuracy 0.5% of full scale, Phenix Technologies, Accident, MD, USA), are used.

The PD activity was measured with a commercial electrical PD instrument (PDBase II, 200 Msamples/s, 16 kHz–48 MHz bandwidth, Techimp, Zola Pedrosa, Italy). The PDBase II instrument expresses the measured voltage as equivalent charge in pC. Prior to testing, the test electrodes were calibrated using a PD calibrator (PDCAL, Techimp, Zola Pedrosa, Italy) by averaging 500 calibration pulses to minimize the effect of background noise. PD pulses were measured using the standard indirect circuit based on IEC 60270, which requires a measuring impedance (CPL 542 quadrupole, Omicron, Vancouver, Canada) in series with a coupling capacitor (50 kV-rms, 1000 pF, PD free, Hua Gao, Wuhan, China). The coupling capacitor in series with the quadrupole are connected in parallel with the test object (needle-to-plane or sphere-to-plane electrode). This is the most common circuit used in practice [18]. The quadrupole is required to reject the electromagnetic noise from the HV source and to prevent the PD signal from being shorted by the source. Care was taken to keep the ground conductors as short as possible to minimize inductance and the effect of electromagnetic noise on the test results.

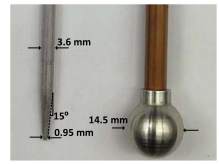


Fig. 2. Detail of the needle and sphere electrodes analyzed in this work.

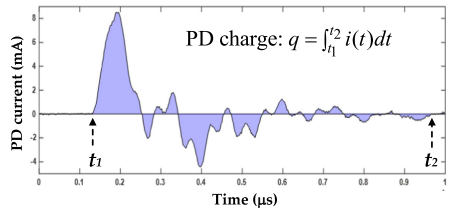


Fig. 3. Current discharge waveform of a 3 mm diameter sphere.

A DSLR camera (Canon EOS-70D digital, 20.2-megapixel CMOS APS-C sensor, bulb mode, ISO-800 sensitivity, f/5.6 aperture, color tungsten temperature, 60-s long exposure, Canon Inc., Tokyo, Japan) was also used to detect corona activity.

Two test objects are analyzed (results are presented in Section III), a needle-to-plane geometry and a sphere-to-plane geometry. The tip of the needle and the lowest point of the sphere were placed vertically at different heights above the ground plane. The dimensions of the electrodes are shown in Fig. 2. Their size and distance from the ground were chosen to ensure that corona discharges take place within the specified parameters and operating range of the equipment used.

A calibrated digital thermo-hygro-barometer (Commeter D4130, Comet, Wünnewil-Flamatt, Switzerland) was used to monitor air pressure, relative humidity, and temperature.

III. EXPERIMENTAL RESULTS

This section first describes the method that is used to measure the sensitivity of the DSLR camera, which is almost independent of the electrode used, as it is based on the number of photons hitting the CMOS sensor. Fig. 3 shows a discharge acquired with a high frequency current transformer (MCT 120, 0.08 to 40 MHz, Omicron, Vacaville, USA) connected to a high performance oscilloscope (MDO3024, 200 MHz, 2.5 Gs/s, Tektronix). Different tests comparing the acquisition of this setup with the images of the DLR camera showed that the sensitivity of the camera is better than 100 pC (apparent charge) when placed at 1 m from the discharge electrode. As it reflects the physical source of the discharge, the apparent charge and the corona inception voltage (CIV) are the most used indicators for PD diagnosis.

Electrical PD detectors are widely used in research, industrial laboratories, and field tests. They generate PRPD patterns, which are very useful for visualizing the amplitude and frequency of discharges over the 360° of the ac cycle and to classify the type of discharge. During testing, care was taken to disconnect all nonessential electrical equipment to reduce noise levels. In the case of using a DSLR camera, long exposure photographs allow to increase its sensitivity because they are based on a larger number of photons. The photographs allow not only the detection but also the localization of the corona discharges. The DSLR camera needs to be operated in complete darkness to

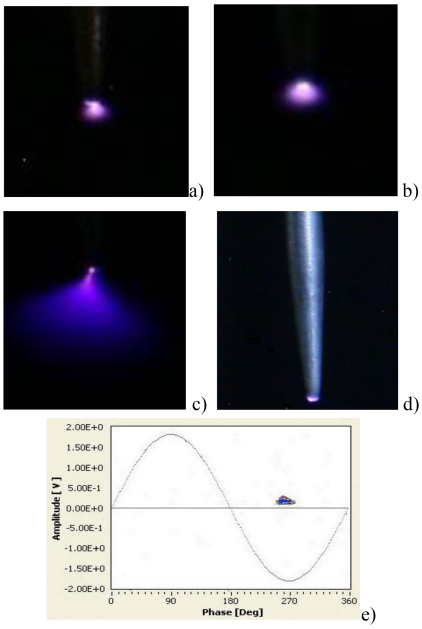


Fig. 4. Needle-to-plane electrode with the tip of the needle 17.5 cm above the ground plane. (a) AC supply at 14.14 kV peak. (b) DC- supply at 13 kV. (c) DC+ supply at 13 kV. (d) DC+ supply at 24 kV (corona glow). (e) PRPD pattern under ac supply at 11.3 kV peak (CIV of PD detector).

optimize its sensitivity. The distance between the camera and the test object is always 1 m.

Corona discharges were generated in the setup shown in Fig. 1 using needle-to-plane and sphere-to-plane electrodes. Both the PD instrument and the digital camera were used to detect the discharges under different voltage supplies (dc+, dc−, and ac). The needle electrode was placed vertically with the tip at 11.5, 17.5, and 21.5 cm above the ground plane, while the bottom of the sphere was placed at 8 and 19.5 cm above the ground plane.

Fig. 4(a)–(d) shows some 60 s exposure corona images of the needle-to-plane electrode taken with the DSLR camera, while Fig. 4(e) shows a PRPD plot corresponding to the CIV of the PD detector.

Comparing Fig. 4(c) (dc+, 13 kV) and Fig. 4(d) (dc+, 24 kV), it is clear that the intensity of the corona discharges decreases under corona glow conditions (pulseless mode), despite the higher voltage applied. It is worth noting that for the dc+ supply under pulseless corona glow conditions, the camera clearly shows the corona activity, but the PD instrument does not detect any pulse. The needle-to-plane corona glow mode occurs at different voltage levels for the three air gaps, i.e., 12.5 kV for 11.5 cm, 14.6 kV for 17.5 cm air gap, and about 16.2 kV for 21.5 cm. The PRPD plot of Fig. 4(e) shows a corona pattern because the pulses are in the negative semi-cycle of the voltage, while the average phase angle is close to 270° (maximum peak value).

Fig. 5(a) shows a corona image at 42.42 kV under ac supply, Fig. 5(b) under 40 kV dc- supply, and Fig. 5(c) under 45.5 kV dc+ supply (streamer discharge) for a sphere-to-plane electrode with an air gap of 8 cm. The experimental results shown in Fig. 5(d) show that the pure corona glow mode appears only at 19.5 cm, since for smaller air

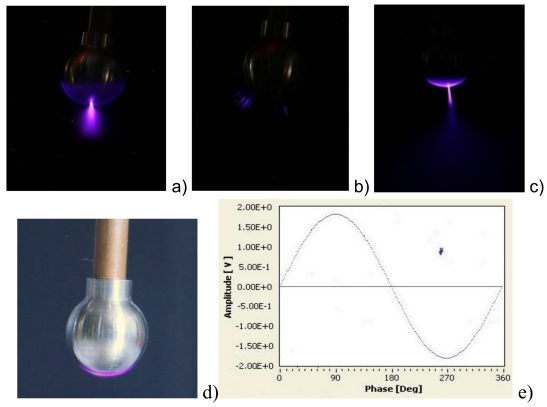


Fig. 5. Sphere-to-plane electrode. (a) AC supply at 42.42 kV peak (8 cm air gap). (b) DC- supply at 40 kV (8 cm air gap). (c) DC+ supply at 45.5 kV (8 cm air gap). (d) DC+ glow phenomenon at 65 kV (19.5 cm air gap). (e) PRPD pattern under ac supply at 46.2 kV and air gap of 19.5 cm (CIV of PD detector).

Table 1. CIV Observed With the Electrical Method and the Camera

Electrodes	Source	CIV (kV or kV_{peak})		
		Electrical PD detector	Camera	Difference
Needle-to-plane (air gap = 11.5 cm)	DC+	11.0 kV	11.0 kV	0%
	DC−	10.5 kV	9.5 kV	−10.5%
	AC	10.2 kV	9.6 kV	−6.3%
Needle-to-plane (air gap = 17.5 cm)	DC+	11.5 kV	11.3 kV	−1.8%
	DC−	10.1 kV	9.5 kV	−6.3%
	AC	11.3 kV	10.2 kV	−10.8%
Needle-to-plane (air gap = 21.5 cm)	DC+	14.4 kV	11.7 kV	−23.1%
	DC−	11.6 kV	7.2 kV	−61.1%
	AC	12.3 kV	10.0 kV	−23.0%
Sphere-to-plane (air gap = 8 cm)	DC+	45.0 kV	42.0 kV	−7.1%
	DC−	38.0 kV	32.0 kV	−18.8%
	AC	38.2 kV	37.6 kV	−1.6%
Sphere-to-plane (air gap = 19.5 cm)	DC+	Pulseless glow	65.0 kV	–
	DC−	63.0 kV	42.3 kV	−48.9%
	AC	46.2 kV	45.7 kV	−1.1%

gaps the corona glow is combined with streamer discharges. Fig. 5(e) shows a PRPD pattern under ac supply at 19.5 cm.

Table 1 shows the CIV obtained with the PD instrument (1 to 3 discharge pulses per electrical cycle) and the digital camera for the needle-to-plane and sphere-to-plane electrodes under dc+, dc−, and ac supplies. It should be noted that the CIV value is the minimum voltage level at which corona activity can be detected when the voltage is gradually increased from a small value.

The tests were performed under room conditions (25°C, 40% relative humidity, 980 hPa). All the results shown in Table 1 were obtained with the camera and the electrical PD detector operating simultaneously under the same test conditions and that the duration of the measurement is similar for both instruments. The results shown in Table 1 correspond to the CIV value, but the images shown in Figs. 4 and 5 do not correspond exactly to the this value. This is because the corona light is very weak at the CIV value. Due to the sharpness of the needle-to-plane electrode, its CIV values are much lower than those corresponding to the sphere-to-plane electrode. The results presented in Table 1 show that in an unshielded laboratory, the digital camera detects the CIV value at a lower voltage than the electrical PD detector

in all the cases analyzed (different electrodes, different distances from ground and different voltage supplies). These results also show that as the distance from the electrode to the ground plane increases, i.e., as the electric field becomes more nonuniform, the differences between the camera and the PD instrument tend to increase. In a shielded laboratory, the CIV results are expected to remain unchanged with the digital camera, but the sensitivity of the PD instrument is expected to increase. This means that in an unshielded laboratory, the digital camera can be slightly more sensitive than the electrical PD detector, while still allowing a clear localization of the discharge area. These results also show that in most cases the maximum difference between the CIV detected by the digital camera and the electrical PD detector corresponds to the dc- supply, and this difference increases with the air gap distance. This may be due to the fact that negative corona activity tends to be less stable than positive corona, with more frequent transitions between different discharge modes. Negative corona is also more susceptible to the effects of cathode surface contamination, which affects emission characteristics and discharge behavior. This makes it more difficult to maintain and reproduce experimental conditions. These effects make it more difficult to obtain consistent and repeatable measurements under negative dc corona.

Under corona glow mode, the electrical PD detector cannot measure any trace of PD activity due to the pulseless nature of this discharge, while the digital camera clearly detects the corona glow activity. In all cases analyzed, the CIV obtained with the DSLR camera is lower compared to the results of the electrical PD detector, allowing the user to detect corona activity at an earlier stage with the DSLR camera. These results clearly show the usefulness of using a DSLR camera to detect and locate corona discharges in industrial test facilities for product optimization and regular product inspection.

IV. CONCLUSION

Corona discharge testing is critical for HV components, such as substation connectors, fittings, assemblies, insulators, and other HV components, to ensure efficient and reliable operation. This letter has compared the performance of two corona detection instruments, an electrical PD detector combined with a measuring impedance and a coupling capacitor and a DSLR camera. Two commonly used electrode geometries, a needle-to-plane and a sphere-to-plane air gap, were analyzed. Both electrodes were tested under different types of HV power supplies such as AC, DC+ and DC−. Tests conducted in an unshielded laboratory have shown that the DSLR camera has comparable or even better sensitivity than the electrical PD detector, with sensitivity better than 100 pC (apparent PD charge). It has also been shown that in glow pulseless corona mode, the electrical PD detector is unable to detect corona activity, while the DSLR camera clearly detects and identifies the corona discharge regions. Note that the DSLR camera is not intended to replace the PD detector as it is not a standard method regulated by international standards, but it can be used as a complementary method to locate discharge areas and assist in decision making and interpretation of results. While the DSLR camera offers simplicity and cost-effectiveness while clearly identifying the exact corona locations, it requires operation in total or partial darkness, while the electrical PD detector requires a much more expensive (by a factor of about 50) and complex experimental setup (the electrical PD detector itself, a measuring impedance, and a coupling capacitor).

ACKNOWLEDGMENT

This work was supported in part by the Ministerio de Ciencia e Innovación de España under Grant PID2020-114240RB-I00 and in part by the Generalitat de Catalunya under Grant 2021 SGR 00392.

REFERENCES

[1] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Rep.*, vol. 9, pp. 215–219, Nov. 2023.

[2] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *Int. J. Electr. Power Energy Syst.*, vol. 155, Jan. 2024, Art. no. 109537.

[3] G. C. Stone, "Partial discharge diagnostics and electrical equipment insulation condition assessment," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 12, no. 5, pp. 891–904, Oct. 2005.

[4] A. Rodrigo Mor, L. C. Castro Heredia, D. A. Harmsen, and F. A. Muñoz, "A new design of a test platform for testing multiple partial discharge sources," *Int. J. Electr. Power Energy Syst.*, vol. 94, pp. 374–384, Jan. 2018.

[5] S. N. Meitei, "Partial discharge detection using piezoelectric sensors on power transformer: A review," *IEEE Sens. J.*, vol. 24, no. 9, pp. 13730–13742, May 2024.

[6] P. Casals-Torrens, A. González-Parada, and R. Bosch-Tous, "Online PD detection on high voltage underground power cables by acoustic emission," *Procedia Eng*, vol. 35, pp. 22–30, Jan. 2012.

[7] Z. Wu, Y. Guo, J. Zhou, X. Fan, Z. Pei, and Q. Zhang, "Optical sensing technology for acoustic detection of partial discharges: A review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 31, no. 2, pp. 591–606, Apr. 2024.

[8] G. M. Ma, H. Y. Zhou, M. Zhang, C. R. Li, Y. Yin, and Y. Y. Wu, "A high sensitivity optical Fiber sensor for GIS partial discharge detection," *IEEE Sens. J.*, vol. 19, no. 20, pp. 9235–9243, Oct. 2019.

[9] P. Bleuler et al., "A unified sensor platform for investigating corona effects on overhead lines," *IEEE Trans. Instrum. Meas.*, vol. 72, Sep. 2023, Art. no. 9005011.

[10] J. R. Riba and P. Bas-Calopa, "Use of DSLR and sonic cameras to detect and locate high-voltage corona discharges," *Sensors*, vol. 22, no. 19, Sep. 2022, Art. no. 7250.

[11] H. I. Uckol and S. Ilhan, "Corona discharge modes and their detections under DC and AC voltages," *IEEE Sens. J.*, vol. 24, no. 10, pp. 17019–17026, May 2024.

[12] *High-Voltage Test Techniques — Partial Discharge Measurements, 3.0*, Standard IEC 60270:2000, International Electrotechnical Commission, 2000.

[13] S. Kaziz et al., "Radiometric partial discharge detection: A review," *Energies*, vol. 16, no. 4, Feb. 2023, Art. no. 1978.

[14] L. Kirkcaldy, G. Lees, R. Rogers, and P. Lewin, "Time synchronized distributed acoustic sensing of partial discharge at the oil-pressboard interface," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 29, no. 6, pp. 2348–2353, Dec. 2022.

[15] J.-R. R. Riba, "Linking digital image intensity to carrier density in low-pressure corona discharges," *Sensors Actuators A Phys*, vol. 359, 2023, Art. no. 114474.

[16] D. S. Prasad and B. S. Reddy, "Study on corona activity using an image processing approach," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 4008–4014, Jul. 2017.

[17] Z. Yuan, Q. Ye, Y. Wang, and Z. Guo, "State recognition of surface discharges by visible images and machine learning," *IEEE Trans. Instrum. Meas.*, vol. 70, Oct. 2020, Art. no. 5004511.

[18] E. Lemke et al., "Guide for electrical partial discharge measurements in compliance to IEC 60270," *Electra*, vol. 241, pp. 61–67, 2008.

Appendix 3

III

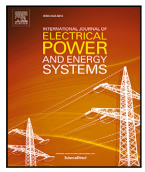
P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power and Energy Systems*, vol. 155, p. 109537, 2024

Electrical Power and Energy Systems 155 (2024) 109537

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes



Weather dependency of corona losses on 330 kV overhead transmission lines

Pradeep Kumar Gupta *, Kaur Tuttelberg, Jako Kilter

Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Tallinn, Estonia

ARTICLE INFO

Keywords:
Corona losses
Feature analysis
Phasor Measurement Unit
Real-time monitoring
Transmission System Operator
Weather dependency
Weather parameters

ABSTRACT

This paper presents an analysis of how corona losses on transmission lines depend on weather conditions in their vicinity. Two years of synchronized phasor measurement data has been collected to monitor corona losses over a longer period of time. From the same time period, both hourly and 10 min weather monitoring data has been obtained for comparison. The dependence between various weather parameters and corona losses is presented with graphical data analysis. In addition, feature importance methods have been applied to rank the weather parameters and determine which of them have the largest influence on the occurrence of significant corona losses. The results presented in this paper are directly applicable for transmission systems operators for adequate planning of system operation.

1. Introduction

Transmission System Operators (TSO) are responsible for adequate and cost effective control of their power networks. One of the important elements of that is the management of system running costs. Among other, these include costs related to buying the network losses from the market. These costs can be significant, e.g. according to [1] the network losses in 2021 for the whole Baltic synchronous area were 38 MEUR, and therefore respective understanding on origin and dependencies of these losses is of interest. In general, network losses include a technical part (dependent on overhead line construction and voltage level) and a part that is referred to as the corona losses which have strong dependence on weather. This paper is dealing with understanding the challenges related to corona losses estimation and its weather dependency.

Corona is a discharge phenomenon due to a rise in air ionization across the conductors. In high voltage overhead transmission lines, the effects of corona are radio interference, audible noise (hissing sound) and ozone emission. From network operational point of view corona discharge can be considered as an additional cause of losses and the level of it is influenced by different weather parameters (e.g. air density, air humidity, solar radiation, air conductivity) and overhead design related factors (e.g. voltage level, conductor type, electric field intensity). In order to better understand how the intensity of corona losses varies and to forecast such losses with hourly precision over short time periods (e.g. day ahead), it is important to understand these weather dependencies better.

Weather dependence has been a part of many studies on corona losses, including the early works developing empirical equations for corona loss estimation [2–4]. Experimental work that lead to Peek's and Peterson's equations included observing corona losses in various weather conditions. Weather dependence can be seen in some parameters, e.g. the air density correction factor, which is a function of air pressure and ambient temperature [2]. In other cases, correction factors for various categorized weather conditions have been developed for the empirical equations [3,4]. However, factors for categorized weather conditions or a single combined weather parameter are too generalized for analyses over shorter time periods where different weather parameters change significantly and independently of each other.

Various other studies have been carried out both as laboratory and field experiments to study the effects of weather conditions (either real or emulated) on corona losses [5–10]. Large corona losses have commonly been an issue in colder climates and many studies have focused on various winter conditions. The work in [5,6] proposed the analysis of corona losses considering different weather conditions such as rime ice, glaze ice, rain etc., where the highest corona loss was seen in the case of glaze ice. To see the impact of icing severity, an experimental approach was proposed in [7] using a corona cage in an artificial climate chamber. In case of hoar frost condition, an experimental setup with corona cage was used to monitor corona losses, which resulted in the highest losses as compared to sunny or rainy weather [8].

Other studies have focused on rainy weather conditions. The authors in [9] executed an experimental approach where direct current (DC) was energized to the conductor and investigated the effect of conductor surface type and rain intensity. The results indicate a slight increase in corona current for increasing rain intensity. To observe the influence of outdoor environment on corona losses, an outdoor experimental

* Corresponding author.
E-mail address: pradeep.gupta@taltech.ee (P.K. Gupta).

https://doi.org/10.1016/j.ijepes.2023.109537
Received 16 February 2023; Received in revised form 12 August 2023; Accepted 27 September 2023
Available online 7 October 2023

setup was installed in [10]. The author's in [11] explored the relation of corona losses in case of heavy rain and sandy-dusty weather with varying altitude, also an annual sum of corona loss was determined for rainy, snow, foggy and fair weather conditions. Even though these studies give insights into how various weather conditions influence the intensity of corona losses, it is difficult to quantify the dependencies and differences between them.

Another study gathered a larger and more varied set of measurement data of operational transmission lines over a longer period of time [12]. In this case, it was possible to analyze the impact of various weather parameters on corona losses such as conductor temperature, relative atmospheric humidity, dewpoint, wind speed, clouds, visibility and precipitation. While this study provides more quantifiable results, it raises new questions. It is not clear which weather parameters have the strongest influence and it is not very clear how strong and consistent the presented dependencies are.

In recent years, various studies have utilized Phasor Measurement Units (PMU) for corona loss estimation [13–15]. The present research leverages the more widespread adoption of PMUs to monitor corona losses over a longer period of time without needing any additional experimental equipment. Corona loss monitoring data has been gathered over a period of two years. This has been supplemented by weather monitoring data from weather stations close to the transmission line over the same period. Based on the combined data, the dependence of observed corona losses on various weather parameters is analyzed.

This paper presents a graphical data analysis of the dependence of corona losses on each observed weather parameter. From this analysis, it can be seen how each parameter independently correlates with changes in losses. Following that, feature importance analysis is applied to determine importance scores and rankings of the different weather parameters. This improves the understanding of which weather parameters have the strongest influence on corona losses. The study has been carried out based on both hourly and 10 min weather monitoring data. The different time-steps in weather monitoring have been compared to determine whether more commonly available hourly data is sufficient for such analyses.

The paper is organized as follows. Section 1 presents a short introduction and literature on corona loss estimation approaches and influence based on various weather parameters. Section 2 discusses the methodology and theoretical background. Section 3 deals with the results analysis which shows the effects of weather parameters on corona losses and includes features importance analysis. The paper ends with conclusions in Section 4.

2. Methodology

2.1. Data collection

For corona loss estimation and monitoring, data measured with PMUs was used. These PMUs are normally a part of a Wide Area Monitoring System (WAMS) which handles data gathering and storage. For monitoring one transmission line, measurements from both ends of the line are required. For this study, data from a period of two years (2018–2019) was collected. The measured phasor data was first resampled at 1 sample per second rate to reduce the amount of data. From this data, loss estimates were calculated, based on principles summarized in the next subsection. In order to compare loss data to weather data, the values were further smoothed by averaging to match the corresponding time steps of weather data.

Weather data was obtained from weather stations of the Estonian Environment Agency. These are conventional weather stations used and operated by the national weather service [16]. The values were sampled with two different time steps, i.e., 10 min and 1 hr. The following quantities and values were monitored referred to as weather parameters in the paper.

- Air temperature
- Air pressure
- Wind speed
- Wind direction
- Visibility
- Precipitation
- Humidity
- Automatic sensor-based classification of weather conditions

2.2. Corona loss estimation

In this paper, the estimation of corona losses is based on the approach according to [15] and the systematic error is reduced using the correction method from [17]. In order to estimate corona losses for different weather conditions, PMU data from both ends of the overhead transmission line is needed. The difference between sending and receiving end power (i.e. losses) can be expressed as

$$\underline{V}(L)\,\underline{I}^{*}(L) - \underline{V}(0)\underline{I}^{*}(0) = \Delta P + j\Delta Q \tag{1}$$

where $\underline{V}(L)$ and $\underline{I}(L)$ are the sending end voltage and current phasors, $\underline{V}(0)$ and $\underline{I}(0)$ are the phasors quantities at receiving end, ΔP are total active losses, ΔQ are total reactive losses (or generation), and $*$ denotes the complex conjugate.

The active transmission loss and reactive power balance can be represented by integrating over the line length as following:

$$\Delta P = r\int_0^L \left|\underline{I}(l)\right|^2 dl + g\int_0^L \left|\underline{V}(l)\right|^2 dl \tag{2}$$

$$\Delta Q = x\int_0^L \left|\underline{I}(l)\right|^2 dl - b\int_0^L \left|\underline{V}(l)\right|^2 dl \tag{3}$$

where $\underline{V}(l)$ and $\underline{I}(l)$ are the voltage and current phasor values at distance l along the line. Here, the losses dissipated in the line are resistive loss through resistance (r), reactive power consumed in reactance (x), losses in conductance (g), which is mainly due to corona or leakage current, and reactive power generation in susceptance (b) — all per unit distance values.

For the estimation itself, the line parameters above can be expressed from the distributed parameter transmission line equations and computed as

$$\underline{\gamma} = \frac{1}{L}\,arcosh\left(\frac{\underline{V}(L)\,\underline{I}(L) + \underline{V}(0)\,\underline{I}(0)}{\underline{V}(0)\,\underline{I}(L) - \underline{V}(L)\,\underline{I}(0)}\right) \tag{4}$$

$$\underline{z}_c = \frac{\underline{V}(L) - \underline{V}(0)\cosh(\underline{\gamma}L)}{\underline{I}(0)\sinh(\underline{\gamma}L)} \tag{5}$$

where L is the length of the line (km) and per unit impedance and admittance are given by $\underline{z} = r + jx = \underline{\gamma}\,\underline{z}_c$ [Ω/km] and $\underline{y} = g + jb = \underline{\gamma}/\underline{z}_c$ [S/km]. For the purpose of this study, it is assumed that leakage losses in the g term are negligible and it represents corona losses.

2.3. Error correction method

Previous work on estimating components of transmission losses from PMU measurements has provided a simplified correction method to reduce some systematic errors in the estimates of corona losses [17]. In the method, corona losses from fair weather conditions are analyzed in order to determine the dependence of the systematic errors on the power flow on the line. The correction is further supplemented by the average fair weather corona loss value calculated from the design parameters of the line.

The corrected corona losses value is determined as

$$P_C(t) = T_C(t, y(t)) - B(P(t)) \tag{6}$$

where the systematic error was approximated by $B(P(t))$ which is defined based on a linear approximation model.

$$B(P(t)) \cong D\,P(t) + C \tag{7}$$

where $P(t)$ is the power flow at time t and the constant values D and C are found for the given transmission line.

2.4. Conductor temperature estimation

In addition to weather and loss data, the conductor temperature has also been approximated. This was necessary to determine the influence of conductor surface status and the load current which regulates the conductor temperature. A high load current will inhibit the formation of snow or hoar frost condition on the conductor surface which reduces corona loss. The current–temperature relationship of a bare overhead conductor is estimated using the nonlinear heat-balance equation from IEEE Std 738-2012 [18]. For the estimation, line loading data has been taken from the PMU measurements and it has been combined with received weather monitoring data.

2.5. Feature importance analysis

In the study, each weather parameter was analyzed separately in order to understand how it affects the occurrence and level of corona losses. However, this does not necessarily show whether a certain weather parameter has a greater influence on corona losses than another one. In order to get a better understanding of this, a feature importance analysis has also been carried out. Four feature importance methods were applied – Mutual Information (MI) [19], Extreme Gradient Boosting Regression (XGBR) [20], Tree based classifier [21], and Boruta method [22].

In general, MI method is defined as the dependence between two random variables, where one random variable obtains the information given another one using the equation

$$I(X;Y) = H(X) - H(X/Y) \tag{8}$$

where, $I(X;Y)$ = MI of X and Y, $H(X)$ = entropy of X, $H(X/Y)$ = conditional entropy of X given Y. MI using Venn diagram is defined in [23] providing more information how the two variables are related to each other. Also, the levels of individual relevance are defined in terms of MI.

XGBR method is an efficient ensemble learning method which is an implementation of gradient boosted decision trees. To support the weighted classification among different variables and ranking function along with user defined objective functions XGBoost package was implemented [24]. The ranking of weather parameters is evaluated based on gain and score values which ranges between zero and one.

In Tree Based Classifiers approach, feature importance is built into the method where it gives a score for each feature between 0 – 1. The most important input weather features are selected based on the score achieved — the higher the score, the more relevant the feature to influence corona losses is. In order to apply a classifier model, the loss values have been categorized. Based on minimum and maximum percentile values, the corona losses have been classified as low (P_c < 0.16 MW), medium (0.16 MW $\leq$ P_c $\leq$ 0.19 MW) and high (P_c > 0.19 MW). The Tree Based Classifiers is used from Scikit Learn libraries for feature selection techniques [21,25].

Boruta method evaluates the variables or parameters importance by creating an ensemble of corresponding artificially added 'shadow' variables randomly sampled obtained from the dataset and determine the importance ranking of features or parameters [26]. In Boruta approach the computation is made iteratively where the Z-score is compared with each variable to the shadow variable and then feature importance is assessed [22,27].

3. Influence of weather parameters on corona losses in one estonian power system 330 kV transmission line

The main motivation of this investigation is to assess the influence of different weather parameters on corona losses in high voltage overhead transmission lines. For this study, one 170 km long overhead line (north-south direction) in Estonian 330 kV transmission network was considered. For a more detailed understanding of weather parameters behavior, two years of weather data sets were taken from different stations along the transmission line corridor. For the same time period, PMU data were used for corona loss measurement.

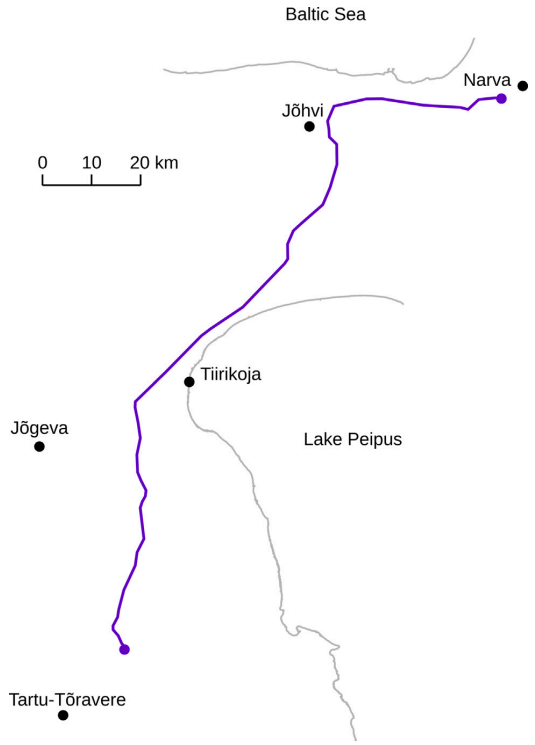


Fig. 1. Map of North-Eastern Estonia depicting transmission line (solid blue line) and respective weather stations (black dots).

3.1. Overhead transmission line overview and data

The observed transmission line was constructed in 1959, it is located in the eastern part of Estonia and follows a mostly north to south corridor—from Narva to Tartu. The approximately 170 km long line corridor together with locations of nearby weather stations is illustrated in Fig. 1. The most characteristic tower configuration of the transmission line is shown in Fig. 2 [28]. Respective design parameters and the measured impedance of the transmission line are presented in Table 1. While the estimation of loss components involves the estimation of line parameters, the present paper does not concentrate on this topic. Previous work has evaluated parameter estimation and its uncertainties [29,30].

Environmental conditions cause deterioration of the transmission line over time. Components and properties of the line are subject to gradual changes, for example the surface of the conductor is affected and its properties change over time. Over long periods of time (several years or decades) this can likely also influence how corona losses occur on the line. In this work, it is assumed that over the two year monitoring period this effect is negligible, but in other scenarios the effect may have to be considered.

3.2. Weather conditions along transmission line

To analyze the rate of change in weather conditions along the line, weather parameters data from four weather stations, including stations close to both ends of the transmission line, was collected. The distance

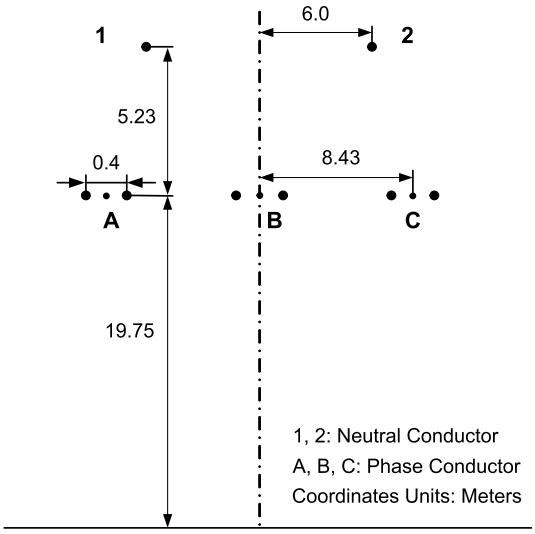


Fig. 2. Tower configuration of the transmission line.

Table 1
Overhead transmission line design parameters and measured value of the line impedance.

Description	Value
Line length	170 km
Conductor type	GOST 839-59
Bundle spacing	400 mm
Conductor radius	14 mm
Conductors per bundle	2
Bundle conductor GMR	66 mm
Positive sequence impedance	55.0 ∠ 83.3°Ω

between the weather station from the transmission line tower for the Narva end is around 5 km, for the Tartu end 20 km, and for all other weather stations is below 15 km. The difference between minimum and maximum elevation along the transmission line is about 55 m, with generally flat topography. The site mapping of 330 kV overhead transmission line is shown in Fig. 1, indicated with a blue color and the location of different weather stations are projected using the dot.

There are various methodologies developed for combining weather data from multiple locations in different manners – linear combination, exponential combination, complex weighting methods, geometric mean combination, etc. [31]. However, it can be difficult to assess which method gives the most characteristic value to describe the whole area. In this research a simple average of multiple weather stations to the values from a single station is compared. A comparison of mean values and standard deviations for the weather parameters is given in Table 2. Based on this analysis it can be concluded that the differences between the averages of multiple stations and a single station are small.

In order to analyze the differences in more detail, Fig. 3 shows the probability density distributions of selected weather parameters (air pressure, air temperature, wind speed and air density) from Narva station, Tartu station and along the transmission line (averaged value of multiple weather stations). The kernel density estimate (KDE) is used from the Python Seaborn library to represent the probability density distribution [32] of weather parameters. The most significant difference can be seen in wind speed values—this is assumed to be caused by the more coastal wind conditions in the Narva end compared to the slightly more inland Tartu end. Overall, it was observed that the differences between the different stations and an average of the stations are small and in general weather conditions along the line corridor can be assumed to be close to uniform. Therefore, in the following analysis, the values from a single weather station are used.

3.3. Weather parameters influence analysis on corona losses

The impact of weather parameters on corona loss under varying temperature conditions is demonstrated from Figs. 4 to 16. As it is seen from the analysis, air temperature is one of the stronger factors to influence corona losses and thus the same grouping is used in the figures for all weather parameters to illustrate their combined effect. For the analysis four different air temperature levels, i.e., low ($T \leq -10.0$ °C), medium-low (-10.0 °C $< T \leq 2.8$ °C), medium-high (2.8 °C $< T \leq 15.6$ °C) and high ($T > 15.6$ °C), are considered with almost equal division considering the minimum and maximum values. In order to illustrate the density and distribution of overlapping data points in the figures, the 75th and 95th percentile lines have also been plotted.

3.3.1. Air temperature

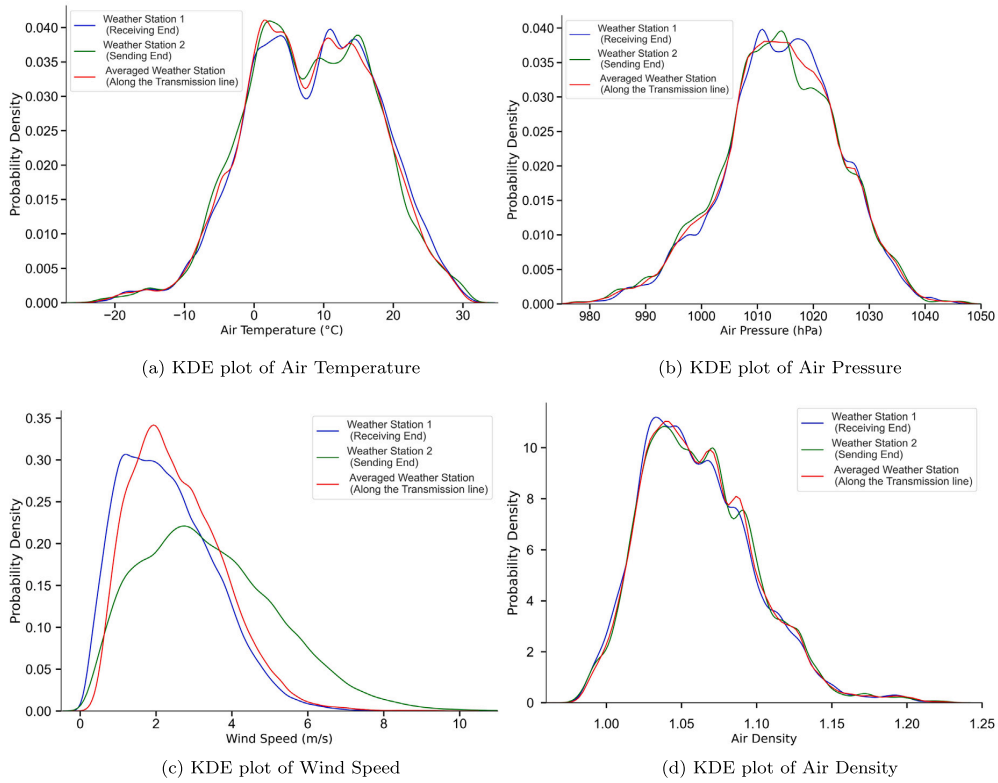
The influence of varying temperatures on corona losses is shown in Fig. 4. It can be seen quite clearly that higher corona losses occur when air temperatures are lower. This is in agreement with common knowledge that larger corona losses occur in winter conditions. Previous studies have shown that higher levels of corona losses tend to appear in conditions of hoarfrost, wet snow or icing [8,12]. The load current could have an effect if it keeps the conductor temperature higher, which could prevent the formation of icing or snow build-up on the conductor surface.

It can also be seen that there is no simple dependence between air temperature and corona losses. There is also a large number of data points where temperature is low, but significant corona losses are not present. Also, similar corona loss intensity may appear in very cold weather (−15 °C and lower) and mildly cold conditions (around 0 °C). However, based on the percentile lines it can be seen that in milder temperatures higher corona losses occur less frequently at the same temperature. In the coldest temperature ranges higher corona losses are more likely to appear.

3.3.2. Conductor temperature

In order to understand whether it could be a better indicator of corona losses, conductor temperature has also been estimated. Fig. 5 shows the dependence between conductor temperature and corona losses. In general, the results are quite similar, i.e., due to loading, conductor temperatures are higher, but larger losses were still occurring in lower temperatures. Again, lower temperatures have not always led to higher corona losses and there is no simple dependence between the two quantities. The percentile data lines for conductor temperature exhibit similar behavior as air temperature where the variation of high corona loss values occurs mostly during winter and the share of higher corona losses is the highest in the coldest temperature range.

In experiments using a corona cage, the inception of corona was studied under different conductor temperatures [33]. In the study, it was observed that increasing conductor temperature at higher excita-



(a) KDE plot of Air Temperature

(b) KDE plot of Air Pressure

(c) KDE plot of Wind Speed

(d) KDE plot of Air Density

Fig. 3. Probability density distribution of weather parameters from single stations and the average of multiple stations.

Table 2
Comparison of mean and standard deviation of weather parameters from a single station and the average of multiple stations.

Weather parameters	Single weather station		Averaged weather station	
	Mean	Standard deviation	Mean	Standard deviation
Air temperature (°C)	8.7	9.2	7.3	9.4
Conductor temperature (°C)	14.4	11.4	12.4	11.4
Air pressure (hPa)	1014.4	10.3	1013.8	11.4
Wind speed (m/s)	2.30	1.24	2.62	1.24
Air density	1.06	0.04	1.06	0.04
Visibility (km)	23.6	10.4	21.9	7.6
Precipitation (mm)	0.012	0.109	0.011	0.052

tion voltage increases corona intensity. In the present study it was not possible to observe the same kind of effect; however, it should be kept in mind that in this study it was not possible to analyze the effect of any given parameter in isolation.

3.3.3. Air pressure

The dependence between air pressure and corona loss intensity has been plotted in Fig. 6. It can be seen that the value of air pressure is not a good indicator of larger corona losses. However, there are several traces visible in the figure, i.e., cases where corona losses have been higher and air pressure has been changing between data points. This indicates that corona losses can coincide with periods when air pressure is changing, which usually means that weather conditions are changing.

In winter, low temperatures usually coincide with clear skies and higher air pressure, thus, it can be expected that higher corona losses are present at higher air pressure values. However, in summer, high air pressure means higher air temperatures and low corona losses. According to Peek's formula lower pressure should lead to higher corona loss [2], whereas, in the presented data such a dependence is not clearly visible. The percentile lines mostly follow the peaks in the data. The highest peak in the low pressure values is mainly due to a small number of data points and one event of larger losses. The highest peak in high air pressure values corresponds to a higher concentration of low temperature value data points.

3.3.4. Wind speed

The dependence between wind speed and corona losses is presented in Fig. 7. It can be seen that the largest corona losses were present when wind speeds were low and at higher wind speeds no high corona losses occurred. At the same time there is a large number of data points where

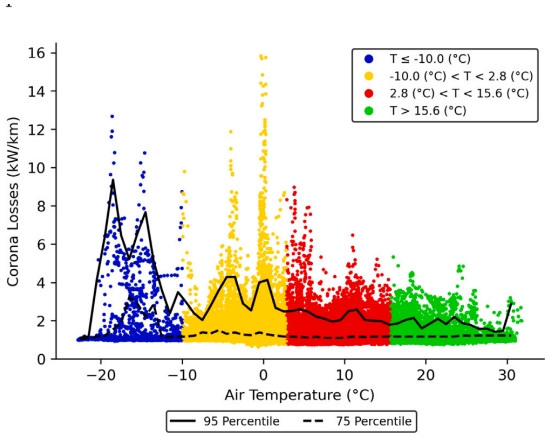


Fig. 4. Dependence between air temperature and corona losses, 10 min data.

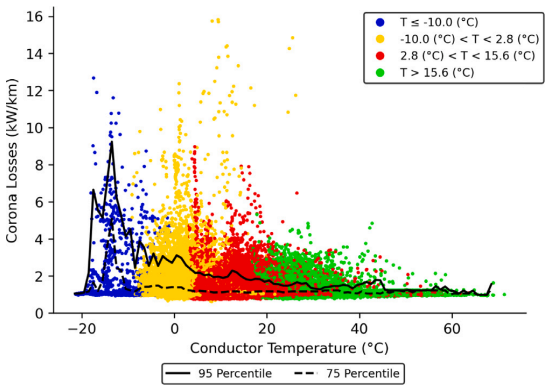


Fig. 5. Dependence between conductor temperature and corona losses under varying air temperature.

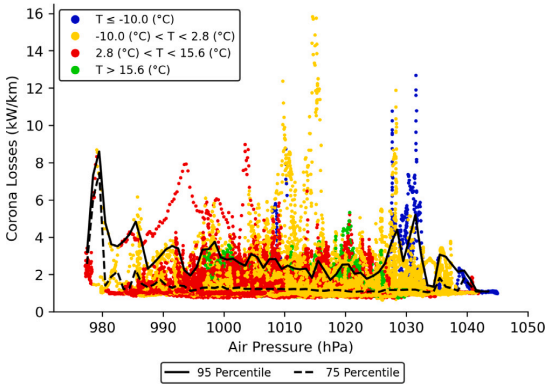


Fig. 6. Dependence between air pressure and corona losses under varying air temperature.

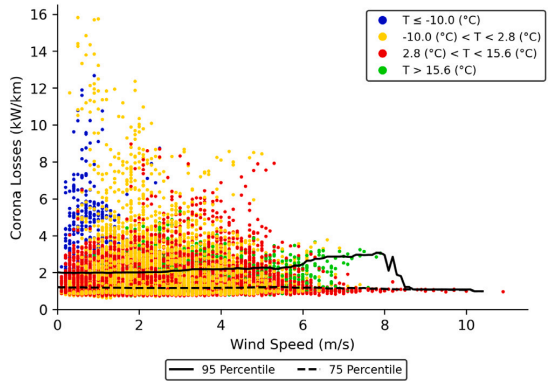


Fig. 7. Dependence between wind speed and corona losses under varying air temperature.

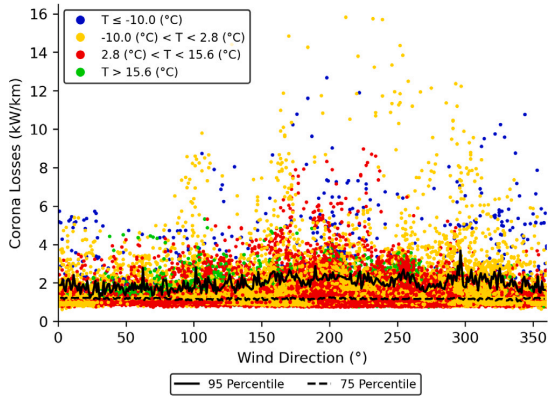


Fig. 8. Dependence between wind direction (0–360 degree) and corona losses under varying air temperature.

wind speed is low. This is illustrated by the nearly flat percentile lines. Also, the lowest air temperatures coincide with low wind speeds. While low wind speeds do not necessarily imply high corona losses, higher wind speeds mean that higher corona losses are very unlikely.

3.3.5. Wind direction

The relationship between wind direction and corona losses is illustrated in Fig. 8. As described earlier, the transmission line lies in a north–south direction. It can be seen that the more significant corona losses are not related to specific wind directions nor does it seem to matter whether the wind is blowing perpendicular to the line or along the line. The percentile lines also show that there is no wind direction where any higher influence would appear more frequently. In a laboratory study, it has been found that for HVDC, corona losses increase with the magnitude of the perpendicular component of wind [34,35], however the same does not appear significantly in the presented data.

3.3.6. Air density

The effect of the air density correction factor (defined as $\delta = \frac{3.92p}{273+T}$ where p is air pressure in cmHg and T temperature in °C [2]) is presented in Fig. 9. It can be seen that the scatter plot and the percentile lines are very similar to what was in Fig. 4 for air temperature, only the x–axis is reversed. According to Peek's formulas, a lower value of δ causes higher corona loss as it is inversely proportional and the

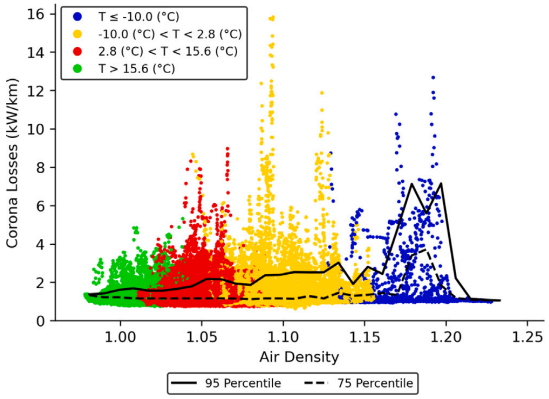


Fig. 9. Dependence between air density correction factor and corona losses under varying air temperature.

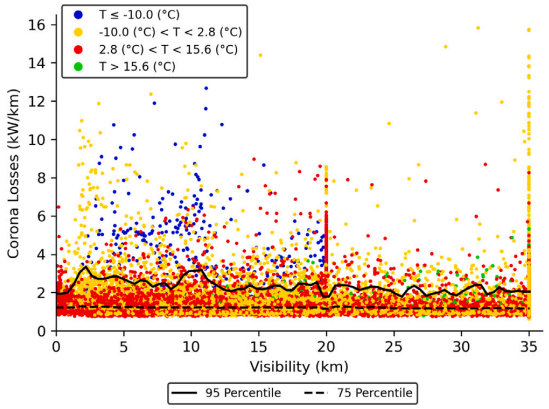


Fig. 10. Dependence between visibility and corona losses under varying air temperature.

disruptive critical voltage decreases when the δ value is decreased [2]. However, from the data it does not appear that the air density correction factor is a clear indicator of higher corona losses and the strongest losses occurred at medium or high values instead.

3.3.7. Visibility

The dependence of corona losses on visibility is presented in Fig. 10. It should be noted here that during the observation period, the maximum value for recording visibility data was changed from 20 km to 35 km. It can be seen that there are data points of higher corona losses across the range of visibility values, but there is a higher concentration of points at high and low visibility. The slight increases in the 95th percentile line around lower values show that higher corona loss values have a slightly larger share in the data points there. At the maximum values, there are high corona loss values but there is also a very large number of low loss points at the same value of visibility. This shows further that there are different sets of weather conditions where higher corona losses tend to be present.

3.3.8. Precipitation

The effects of precipitation on corona losses have been presented in Fig. 11. A few data points of very large precipitation values have been omitted to show the low end where the majority of data points lie. It

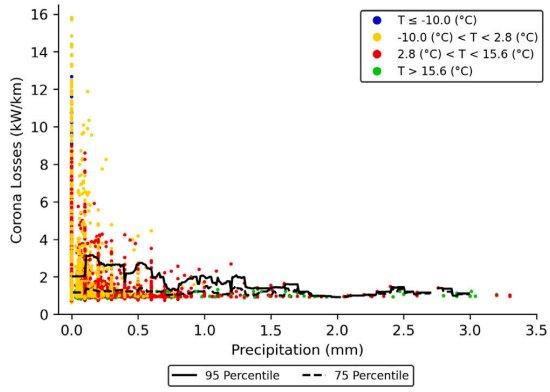


Fig. 11. Dependence between precipitation and corona losses under varying air temperature.

appears that the largest losses occur when there is no precipitation but there is also a region of very low precipitation when significant corona losses can appear. This is also in agreement with what was seen in the visibility data as high visibility would require no precipitation and low precipitation would mean low visibility. It should be noted that in low temperature ranges precipitation values are also low. The peak at zero is similar to the peaks in visibility, i.e., there are data points of higher corona losses but there is also a large number of data points of low losses. The 95th percentile line indicates a slightly larger share of high corona loss data points at low values of precipitation.

3.3.9. Humidity

The relationship between relative humidity and corona losses is shown in Fig. 12. The data for humidity was available only in hourly steps and all values were rounded to full percentage points, resulting in a sparser plot. It can be seen that more significant corona losses occur when humidity is at 75% or higher and air temperature is low. Based on theoretical analysis, higher corona intensity could be expected with higher humidity [36]. Similarly to other observed influences, based on the analyzed data, this effect does not always hold. Nevertheless, it can be seen that at lower values of humidity, higher corona losses are unlikely. The 95th percentile line shows that at higher humidity levels (above 75%) the effect is visible but does not increase at higher relative humidity.

3.3.10. Seasonal weather conditions

In addition to the presented measured weather parameters, the categorical weather conditions provided by automatic sensors have also been analyzed. The categorical weather conditions include atmospheric phenomena such as fog, rain, snow, and also fair weather conditions. Corona losses in various weather conditions have been plotted in Fig. 13. It can be observed that in the coldest temperature range, the largest losses occur in snowy conditions; however, medium losses can be present in both fair and snowy conditions. In medium-low temperature, on the other hand, the highest losses appear in fair weather, while medium losses can be present in foggy, rainy or snowy conditions. In the medium-high temperature range, rainy conditions appear to be the main contributor to higher corona loss levels. In general, the results are in agreement with what was observed in the above analysis of the different weather parameters.

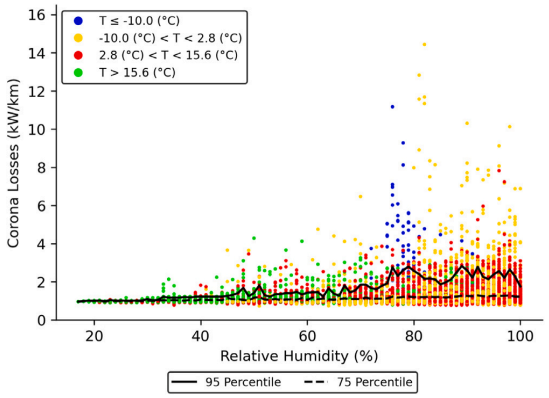


Fig. 12. Dependence between relative humidity and corona losses under varying air temperature.

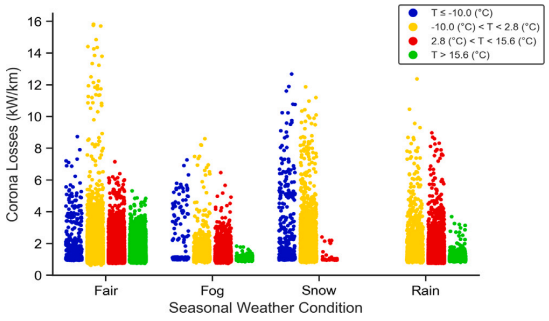


Fig. 13. Dependence between seasonal weather conditions and corona losses under varying air temperature.

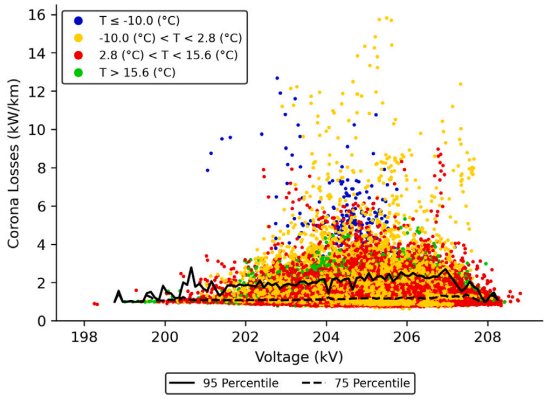


Fig. 14. Dependence between phase voltage and corona losses under varying air temperature.

3.3.11. Phase voltage

Fig. 14 shows the dependence between the measured corona loss and phase voltage which is obtained from the PMU data. In general, voltage variations on the line are quite small and there does not appear to be a significant dependence of corona losses on voltage. The percentile lines do not show any considerable variation with voltage either. While voltage is an important factor in most of the empirical equations for corona losses, its effect in practice is small in the presented data. From the presented data it appears that weather conditions have a stronger influence on corona losses.

3.4. Result analysis of weather parameters ranking based on features importance and feature selection methods

In Section 3.3 the effects of various weather parameters on corona losses were analyzed. While it was seen from the analyzed data that some weather parameters have a stronger effect, it was difficult to determine which of them are the most important contributors. In addition to graphical analysis, feature selection methods were applied in order to analyze the importance of the observed weather parameters.

For the feature importance analysis, four methods were applied: MI, XGBR, Tree-based classifier and the Boruta method. The results obtained with these methods are presented in Fig. 15. For presenting the analysis results in comprehensive manner the vector of all mutual information scores has been normalized so that all importance scores for each method add up to one. The Boruta method can report equal ranking for multiple features and in the results of this analysis this is the case for most of the features. The analysis was carried out on both 10 min and hourly data.

For comparison, linear regression analysis is also used to analyze the relationship between losses and various weather parameters. In this case, the dependence is described by the linear equation $Y = MX + C$, where Y is corona loss and X is the selected weather parameter. The R-squared value is used for result assessment and here a higher R-squared value reflects a better correspondence.

It can be seen from the results that the measured weather parameters that have the strongest influence are air temperature, air pressure, and humidity. Air density and conductor temperature also show similarly strong influence, but these are parameters derived from measured quantities, not directly measured themselves i.e., air density combines temperature and pressure, while conductor temperature is estimated from a larger number of measured quantities. Precipitation appears to be the least influential while wind speed and direction and visibility show varying levels of importance with different methods and different time-steps in the data. According to the R-squared value results it can be concluded that simple linear dependencies are very weak.

With this analysis the objective was to assess the relative importance of weather parameters by evaluating the importance scores obtained by different feature importance methods. Regarding the 10 min and hourly data analysis using linear regression, it can be concluded that despite of different R-squared values a similar importance ranking of the weather parameters is visible. The analysis shows that there is a correlation between corona losses and different weather parameters in a complex form where the most influential parameters are air temperature, air pressure and relative humidity.

3.5. Comparison of 10 min and hourly data

For the presented study the observed weather data was obtained with both a 10 min and a one hour time-step. This was done in order to analyze whether a finer resolution in weather data shows more details in the dependence between weather conditions and corona losses. Normally, weather data recorded with hourly time-steps is more readily available and it is simpler to carry out similar analyses on such data. To illustrate the differences, hourly data for air temperature has been presented in Fig. 16.

When comparing results on data with different time resolution, it can be seen that the graphs are mostly similar. In 10 min data there is of course a larger number of data points, but the temperature ranges

Parameter	MI Score / Tree Classifier Method / XGBR Method — Importance Score		Boruta — Ranking		Linear Regression — R-squared	
	Hourly	10 Mins	Hourly	10 Mins	Hourly	10 Mins
Air Temperature	0.1311, 0.1298, 0.1947	0.152, 0.1509, 0.2167	1	1	0.0366	0.0287
Air Pressure	0.1084, 0.1607, 0.1434	0.12, 0.191, 0.1872	1	1	0.0172	0.0044
Air Density	0.1023, 0.1268, 0.1454	0.1346, 0.1553, 0.2194	1	1	0.0215	0.0206
Conductor Temperature	0.1326, 0.1218, 0.1424	0.1177, 0.1458, 0.1783	1	1	0.0308	0.0252
Wind Direction	0.0966, 0.1283, 0.0266	0.0882, 0.1523, 0.0507	1	1	4.8e-05	0.0005
Humidity	0.0954, 0.1257, 0.1211		1	-	0.0221	-
Wind Speed	0.0598, 0.109, 0.0358	0.1379, 0.1217, 0.0127	2	1	0.0001	2.3e-05
Visibility	0.2221, 0.0742, 0.1671	0.1174, 0.0709, 0.1252	1	1	0.0044	0.0028
Precipitation	0.0513, 0.0233, 0.0231	0.1318, 0.0117, 0.0095	3	2	0.0008	5.4e-05

Fig. 15. Weather parameters importance score and ranking using feature importance methods for 10 min and hourly data 2018–19.

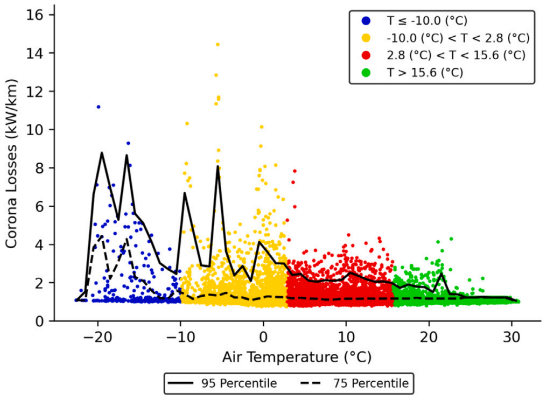


Fig. 16. Dependence between air temperature and corona losses, hourly data.

where higher corona losses occur are the same. In case of events of larger corona losses the finer resolution data shows some extra detail, but the general dependence remains the same. For the more influential weather parameters, feature importance analysis with 10 min and hourly data also gave very similar results. This result is expected as weather conditions do not normally change multiple times within one hour.

In the case of 10 min data, the quantity of data points is higher, allowing more precise identification of dependencies and correlations between the weather parameters and corona loss. Availability of higher frequency data points enable to demonstrate the irregularities of behavior in greater detail, such as unexpected peaks in a very short time period that are overlooked in hourly data. Nevertheless, the findings from the analysis indicate that the use of hourly data is sufficient to understand the relation between weather parameters and corona losses. More detailed data can be used if available but it cannot be said that it provides more comprehensive results with respect to corona loss estimation. The main benefit of using hourly data is that it is more readily available. Most weather reporting is done with an hourly time-step and large amounts of public data are available. Generally, there are also not many cases where it is necessary to analyze the behavior of losses with a resolution below one hour.

4. Conclusion

In modern power systems operation, the understanding of network loss composition is of interest due to related monetary aspects. This paper clarifies in detailed manner the aspects related to corona losses and their weather dependence. For the analysis, actual network PMU data from the Estonian 330 kV transmission network and weather monitoring data over the period of two years were used. In order to comprehensively understand corona loss dependency on different weather parameters, statistical analysis together with feature importance methods and linear regression were applied.

The analysis and results presented in this paper contribute to determine the dependence of different weather parameters on corona loss in high voltage (330 kV) overhead transmission lines in a more comprehensive manner. Based on the analysis it can be concluded that there is a correlation between corona losses and weather data, and the most influential weather parameters are air temperature, air pressure and relative humidity. While the correlations are visible, it was shown that the dependence is complex. Different methods that were applied indicate similar rankings when defining the most influential weather parameters. From the perspective of data availability (10 min and hourly data), it was seen that finer resolution data may be more illustrative, but is not necessary for most analyses.

For future research, the results presented in this paper will be used to improve the prediction quality of corona losses.

CRediT authorship contribution statement

Pradeep Kumar Gupta: Conceptualization, Methodology, Software, Validation, Investigation, Formal analysis, Writing – original draft. **Kaur Tuttelberg:** Visualization, Investigation, Supervision, Project administration, Writing – review & editing. **Jako Kilter:** Investigation, Supervision, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgment

The Estonian TSO, Elering provided measurement data for this study. Weather data from the Estonian Environmental Agency has been used in this work.

References

[1] Elering AS. Elering annual report. 2021.
[2] Peek FW. The law of corona and the dielectric strength of air. Proc Am Inst Electr Eng 1911;30(7):1485–561.
[3] Carroll JS, Cozzens B. Corona loss measurements for the design of transmission lines to operate at voltages between 220 kV and 330 kV. Trans Am Inst Electr Eng 1933;52(1):55–62.
[4] Carroll JS, Rockwell MM. Empirical method of calculating corona loss from high-voltage transmission lines. Trans Am Inst Electr Eng 1937;56(5):558–65.
[5] Yin F, Farzaneh M, Jiang X. Corona investigation of an energized conductor under various weather conditions. IEEE Trans Dielectr Electr Insul 2017;24(1):462–70.
[6] Yin F, Farzaneh M, Jiang X. Laboratory investigation of AC corona loss and corona onset voltage on a conductor under icing conditions. IEEE Trans Dielectr Electr Insul 2016;23(3):1862–71.
[7] He G, Hu Q, Shu L, Jiang X, Liu Y, Wu W, Li Y, Peng H. Impact of icing severity on corona performance of glaze ice-covered conductor. IEEE Trans Dielectr Electr Insul 2017;24(5):2952–9.
[8] Lahti K, Lahtinen M, Nousiainen K. Transmission line corona losses under hoar frost conditions. IEEE Trans Power Deliv 1997;12(2):928–33.
[9] Pfeiffer M, Franck CM. Impact of conductor surface type and rain intensity on HVDC corona losses. IEEE Trans Power Deliv 2015;30(5):2284–92.
[10] Bleuler P, Hedtke S, Franck C. Corona performance of DC overhead lines in outdoor experiments during wet weather transitions and under varying humidity. In: CIGRE digital-E-session 2020: papers and proceedings. CIGRE; 2020.
[11] Liu Y, Chen S, Huang S. Evaluation of corona loss in 750 kV four-circuit transmission lines on the same tower considering complex meteorological conditions. IEEE Access 2018;6:67427–33.
[12] Sollerkvist FJ, Maxwell A, Rouden K, Ohnstad TM. Evaluation, verification and operational supervision of corona losses in Sweden. IEEE Trans Power Deliv 2007;22(2):1210–7.
[13] Vasilenko N, Gadzhiev M, Galiaskarov I, Zhgun K, Korobka V, Ryabchenko V, Sharov Y. Corona losses reduction of OHL 500 kV of omsk electric power system based on signal processing of PMU. In: CIGRE Paris session. 2020.
[14] Pavičić I, Holjevac N, Ivanković I, Brnobić D. Model for 400 kV transmission line power loss assessment using the PMU measurements. Energies 2021;14(17):5562.
[15] Tuttelberg K, Kilter J. Estimation of transmission loss components from phasor measurements. Int J Electr Power Energy Syst 2018;98:62–71.
[16] Estonia Environmental Agency. https://www.ilmateenistus.ee/.
[17] Tuttelberg K, Löper M, Kilter J. Correcting systematic errors in corona losses measured with phasor measurement units. IEEE Trans Power Deliv 2019;34(6):2275–7.
[18] IEEE standard for calculating the current-temperature relationship of bare overhead conductors. IEEE Standard 738-2012, 2012, p. 1–67.
[19] Witten I, Frank E, Hall M, Pal C. Data mining: Practical machine learning tools and techniques. 4th ed.. Elsevier; 2016.
[20] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. 2nd ed.. Springer; 2009.
[21] Freeman C, Kulić D, Basir O. Feature-selected tree-based classification. IEEE Trans Cybern 2013;43(6):1990–2004.
[22] Kursa MB, Rudnicki WR. Feature selection with the boruta package. J Stat Softw 2010;36:1–13.
[23] Vergara JR, Estévez PA. A review of feature selection methods based on mutual information. Neural Comput Appl 2014;24(1):175–86.
[24] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016, p. 785–94.
[25] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: Machine learning in python. J Mach Learn Res 2011;12:2825–30.
[26] Poona NK, Ismail R. Reducing hyperspectral data dimensionality using random forest based wrappers. In: IEEE international geoscience and remote sensing symposium-IGARSS. 2013, p. 1470–3.
[27] Kursa MB, Jankowski A, Rudnicki WR. Boruta–a system for feature selection. Fund Inform 2010;101(4):271–85.
[28] Gupta PK, Tuttelberg K, Kilter J. Assessment of corona loss performance on aging transmission lines using PMU measurements. Energy Reports 2023;9:215–219.
[29] Tuttelberg K, Kilter J. Uncertainty propagation in PMU-based transmission line monitoring. IET Gener Transm Distrib 2018;12(3):745–55.
[30] Tuttelberg K, Kilter J. Real-time estimation of transmission losses from PMU measurements. In: IEEE eindhoven powertech. IEEE; 2015, p. 1–5.
[31] Sobhani M, Campbell A, Sangamwar S, Li C, Hong T. Combining weather stations for electric load forecasting. Energies 2019;12(8):1510.
[32] Waskom ML. Seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021.
[33] Reid G, Vermeulen H. Effects of conductor temperature on corona inception. In: 2014 49th international universities power engineering conference (UPEC). 2014, p. 1–5.
[34] Khalifa M, Morris R. A laboratory study of the effects of wind on DC corona. IEEE Trans Power Appar Syst 1967;(3):290–8.
[35] Maruvada PS, Dallaire R, Heroux P, Rivest N. Long-term statistical study of the corona electric field and ion-current performance of a±900-kV bipolar HVDC transmission line configuration. IEEE Trans Power Appar Syst 1984;(1):76–83.
[36] Gallo C, Germanos J, Courtney J. The effect of humidity and temperature variations on the behavior of wire-to-plane coronas. Appl Opt 1969;8(101):111–9.

Appendix 4

IV

P. K. Gupta, K. Tuttelberg, and J. Kilter, “Assessment of corona loss performance on aging transmission lines using PMU measurements,” *Energy Reports*, vol. 9, pp. 215–219, 2023

Energy Reports 9 (2023) 215–219

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Assessment of corona loss performance on aging transmission lines using PMU measurements☆

Pradeep Kumar Gupta *, Kaur Tuttelberg, Jako Kilter

Tallinn University of Technology, Tallinn 19086, Estonia

ARTICLE INFO

Keywords:
Corona losses
Phasor measurement unit (PMU)
Transmission system operator
Conductor surface
Real-time monitoring

ABSTRACT

This paper presents an assessment of corona loss performance for two overhead transmission lines of different ages using synchronized phasor measurements. In general, it is believed that the aged overhead transmission lines have constrained performance compared to the newer ones as they have been exposed to different environmental conditions for longer period of time. Therefore, in view of availability of synchronized measurements it is possible and of relevance to assess how the corona losses are influenced by line age and what are the main influencing factors. The results of the analysis indicate that there are higher corona losses associated with older overhead transmission lines and respective information is directly applicable for transmission system operators in their system operational planning in order to optimize the purchase of losses from the market.

1. Introduction

With the growing increase in power demand, many high-voltage long transmission lines are built with advanced configurations in order to improve power transfer efficiency. The power transmission over long distances with high efficiency can be accomplished by reducing the losses in the power system network. In current transmission networks the average age of overhead lines is mostly high and as a consequence the performance of these aged overhead lines and conductor surfaces are also deteriorated compared to the newer overhead transmission lines. The operational losses in overhead transmission lines are influenced by the changing rate of deterioration and therefore it is of interest of the transmission system operators to understand the main influencing factors of overhead line performance and in what extent the age of the overhead line should be considered. Moreover, as the operational losses are to be bought directly from the electricity market then the transmission system operator has also a direct economical interest in this.

There are two main types of active power losses – Joule (or resistive) and corona losses. The resistive losses are predictive and can be identified depending on the loading of the line, whereas corona losses are highly dependent on weather conditions and have complex nature. Corona loss not only depends on weather conditions but is also influenced by the conductor's surface field in HVAC systems. It occurs when the corona threshold is exceeded with increasing voltage.

The first empirical equation to calculate the corona loss was introduced by Peek in 1911 (Peek, 1911). Later in 1933, Peterson empirical formula was proposed (Carroll and Cozzens, 1993; Peterson, 1993), to consider low power losses and conductor irregularities. However, both of these empirical methods have limitation when performing corona loss calculation in good weather conditions. Therefore, additional modifications and alternative approaches were needed. Some of the recent fundamental studies related to corona loss calculation including modifications to empirical equations are presented in the literature (Solovyev et al., 2018; Pan et al., 2019; Llamo-Laborí and Santos-Fuentefria, 2022). In recent years, with the development in power system measurement technology and increased availability of PMU based measurement data (Tuttelberg and Kilter, 2018; Vasilenko et al., 2020; Pavičić et al., 2021), a significant improvement in corona loss estimation have been obtained.

The condition of the conductor surface changes in time and is dependent on the environmental conditions where the overhead line is located. Deterioration of conductor surface is one of the major factors that have an influence on the level and intensity of corona losses. Therefore, to estimate the degree of roughness of the conductor, Peek introduces the formula of the roughness factor which is denoted by m (Peek, 1920). The roughness factor variation for a new and aged transmission lines has been discussed in (Maruvada, 2000), and corona discharge cross sectional difference has been discussed in (Bian et al.,

☆ 2023 8th International Conference on Sustainable and Renewable Energy Engineering (ICSREE 2023) 11–13 May, Nice, France.

* Corresponding author.
E-mail address: pradeep.gupta@taltech.ee (P.K. Gupta).

https://doi.org/10.1016/j.egyr.2023.09.160
Received 4 September 2023; Accepted 25 September 2023
Available online 3 October 2023

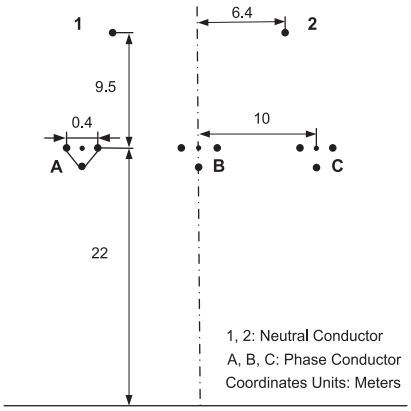


Fig. 1. Transmission line tower configuration–new line.

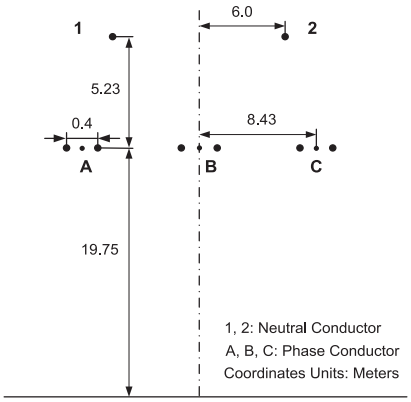


Fig. 2. Transmission line tower configuration–old line.

2010). Although the roughness factor is an important factor, the author in (Rao et al., 2022) presented a study where corona losses in the case of a used conductor was found high without discussing the roughness factor and relative air density factor. Based on the literature it can be concluded that there is a dependence between the level of corona losses and transmission line parameters and weather conditions. With the help of additional measurement data more profound analysis and conclusions can be made.

In view of this, the contribution of this paper is as follows:

- This paper presents a study where the corona loss performance is analyzed in case of aging overhead transmission lines using phasor measurements. To date, no research studies have comprehensively compared and evaluated the corona loss performance of transmission lines with different age in real operating conditions using actual network operational measurements.
- Design and configuration details of both overhead transmission lines are presented to analyze the factors influencing corona losses. Based on this information it is possible to assess whether the corona loss magnitude is as expected considering overhead lines operation.
- The results enable to understand the influence of conductor surface status (conductor temperature is estimated using a nonlinear heat-balance equation from IEEE Std 738-2012) and air temperature levels on corona losses.

Table 1
Design parameters of the transmission lines.

Description	New Transmission Line	Old Transmission Line
Line Length	210 (km)	170 (km)
Conductor Type	ACSR-402	GOST 839-59
Conductor Radius	13.8 (mm)	14.0 (mm)
Bundle Spacing	40 (cm)	40 (cm)
Conductors Per Bundle	3	2
Bundle Conductor GMR	12 (cm)	6.60 (cm)

2. Description of overhead transmission line configuration

Design and configuration of the overhead transmission lines are important and should be checked and improved constantly to enable optimal power transfer considering various climate condition and performance criterions. In this paper, results based on the assessment of two 330 kV overhead transmission lines located in Estonia are presented. Overhead transmission lines under consideration have different configuration and they were built in 1959 and in 2006. In the paper these are referenced as old conductor and new conductor, respectively. Direction of the newer line is from west to east with length of 210 km and the older line is from north to south with a length of 170 km.

2.1. Tower configuration

Tower configuration for both overhead transmission lines under consideration is shown in Figs. 1 and 2. To reduce the losses, bundled conductors are used (old line – two conductors per bundle, newline – three conductors per bundle). Higher number of conductors per bundle is used mainly to reduce line losses. The conductor type for old and new line is GOST 839-59 and ACSR-402, respectively.

2.2. GMD, GMR and bundled conductor

In (Parizad et al., 2011), different options of GMD, GMR, and the number of conductors per bundle were assessed to analyze corona loss. In general, by bundling the conductors the self GMD of the conductors is increased which thereby increases the critical disruptive voltage and hence reduces corona loss. Therefore, corona loss is lowered by using bundled configurations instead of single configuration. The GMR values and the number of conductors per bundle are shown in Table 1 for both transmission lines.

2.3. Conductor radius

The corona loss intensity is decreased when the conductor radius is larger which results in a higher surface area leading to lower surface field intensity. In the case of bundled conductor transmission line the effective radius is larger resulting in lower corona losses. The conductor radius for both transmission lines is given in Table 1.

2.4. Conductor spacing

The effect of conductor spacing on corona loss is inversely proportional, i.e. the corona loss is decreased when the space between the conductors increases. In our case, both transmission lines have the same sub-conductor spacing as shown in Figs. 1 and 2, but the horizontal spacing between conductors is 10 m for the new transmission line and 8.43 m for the old transmission line.

3. Comparison of corona loss performance for two transmission lines of different ages

The present study compares the corona loss magnitude of two aging overhead transmission lines. Estimated level of corona losses for both

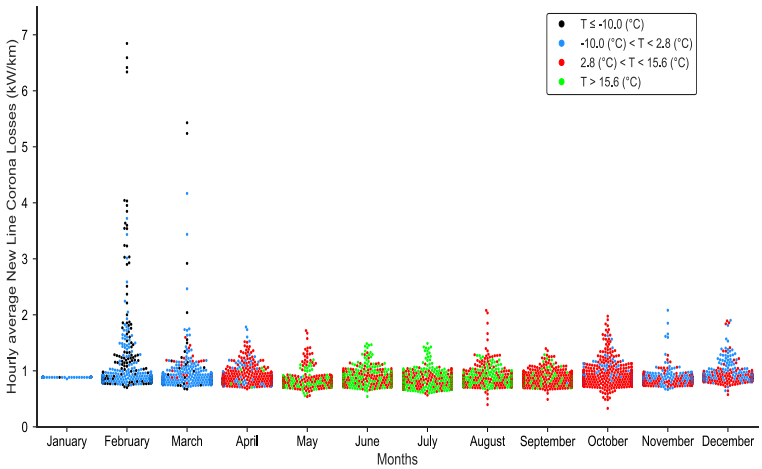


Fig. 3. Monthly comparison of hourly corona loss for New line.

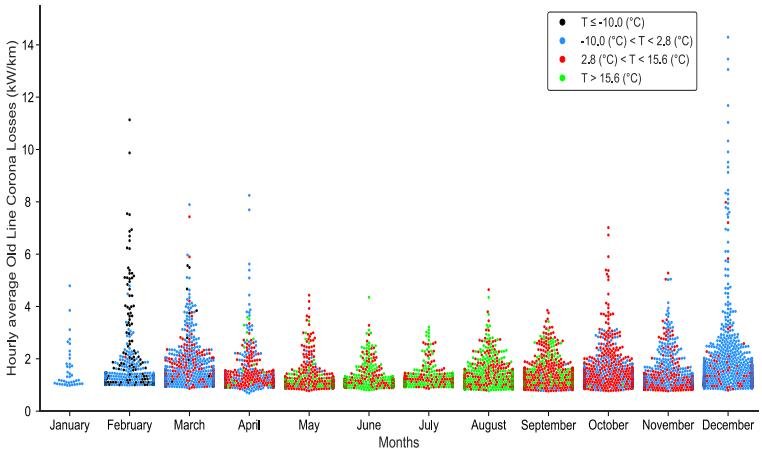


Fig. 4. Monthly comparison of hourly corona loss for Old line.

transmission lines were based on the principles given in (Tuttelberg and Kilter, 2018). To reduce systematic errors a correction method presented in (Tuttelberg et al., 2019) was implemented. The corona loss solution obtained from PMU measurements can be used while taking into account the actual weather conditions. The empirical methods are determined mainly based on experimental set-up and the requirements for conventional SCADA systems includes both hardware and software, which do not have phase angle measurements due to the synchronization factors and its dependence on various other factors associated with technical difficulties. However, in general, due to the complex dependence of corona loss and weather conditions including the design of the conductor and conductor surface properties, the conventional empirical methods are difficult to apply in real operational transmission lines in the power system.

To determine the unknown parameters of the empirical formulas, experimental verification is needed in order to fully understand the characteristics. This also makes it difficult to perform experimental measurements on operating overhead transmission lines for corona loss measurement that required sophisticated instruments and equipment. Therefore, the estimation of corona loss based on PMUs measurement will help to reduce the complexity and detect changes in real-time scenarios (Tuttelberg and Kilter, 2015). It should be noted that the observable overhead transmission lines have different directions and they are experiencing the change in weather conditions slightly different. For the assessment, weather information from multiple nearby weather stations were used.

3.1. Comparison of corona losses for different time periods

This section discusses the change in corona loss on both transmission lines for nearly two years (2018–19). The corona loss analysis is performed using data from February 2018 up to December 2019 (limited data for January 2018 was available). From the results it is seen that during winter, especially in the month of December, February and March, higher corona losses are observed as shown in Figs. 3 and 4. Corona loss information for both lines show higher values in case of an older transmission line. Moreover, the results presented indicate also clear seasonal variation of corona losses where the corona losses are higher mostly during winter and lower during summer months.

Results of the study support the assumption that the aged

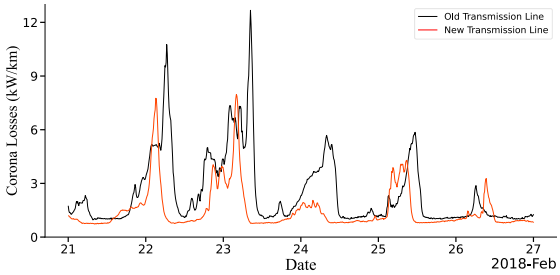


Fig. 5. Comparison of corona loss over the span of one week (Feb).

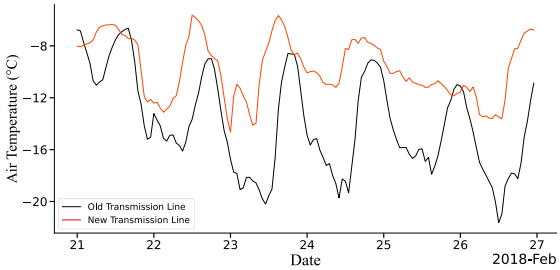


Fig. 6. Comparison of air temperature over the span of one week (Feb).

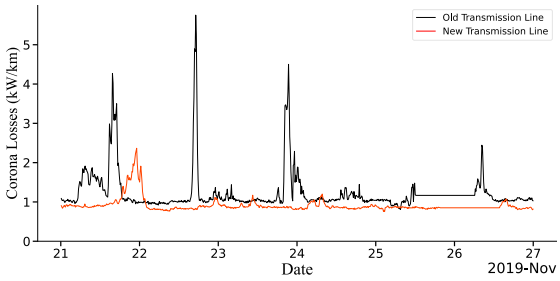


Fig. 7. Comparison of corona loss over the span of one week (Nov).

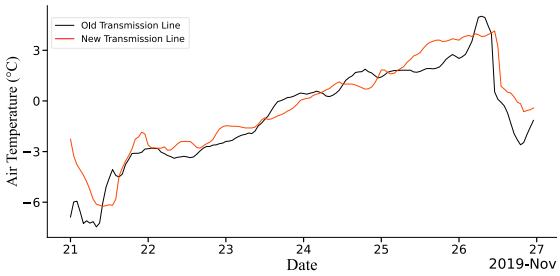


Fig. 8. Comparison of air temperature over the span of one week (Nov).

transmission line (old transmission line) exhibits more corona losses as compared to the new transmission line for different time cycles as shown in Fig. 3 and in Fig. 4. It can also be concluded that the aging affects the performance and available transmission capacity of the overhead transmission lines.

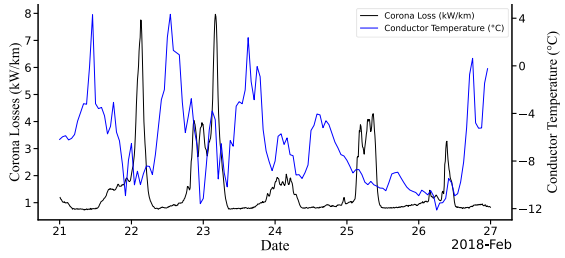


Fig. 9. Dependence of corona loss and conductor temperature for New overhead line.

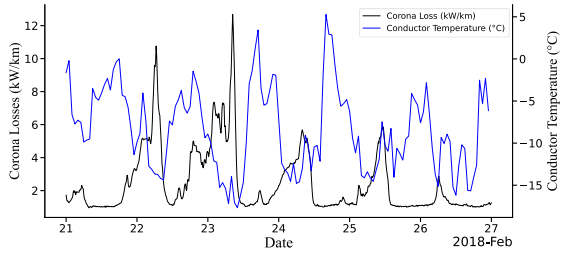


Fig. 10. Dependence of corona loss and conductor temperature for Old overhead line.

3.2. Weekly comparison of corona losses

To obtain a better understanding on corona losses the weekly values of corona losses are plotted in Fig. 5 and in Fig. 7 for the months of February and November, respectively. Results are presented for one week using 10 min time-step phasor measurements. Based on the results presented in Fig. 5 and in Fig. 7 it can be concluded that the old transmission lines reflect high corona losses.

To illustrate the dependence between air temperature and the level of corona losses, a respective temperature variations from the same time period (Figs. 6 and 8, respectively) are enclosed for reference. According to Peek, the air density correction factor, which is a function of pressure and temperature, is inversely proportional to corona loss. In our case, it is found that the air temperature values are primarily below −5 ℃ and dropped up to −22.5 ℃ which coincides with the significantly higher level of corona loss in the month of February. A similar observation was seen for November 2019 where the rise in temperature reduces the occurrence of high corona loss. However, in the case of November 2019, the old transmission line shows higher corona loss even though the air temperature behavior is the same for both lines.

3.3. Influence of conductor temperature on corona losses

The results of previous studies show how the conductor surface affects corona loss (Maruvada, 2000). The physical characteristics of the conductor surface are influenced by its temperature. In order to show the clear dependence between conductor temperature and corona loss, Figs. 9 and 10 demonstrate that when the conductor temperature is decreased corona loss is significantly increased. Here, the conductor temperature is approximated using a nonlinear heat-balance equation (IEEE Std 738-2012) which is dependent on various other parameters such as wind speed, azimuth, air temperature, heat gain, resistance, and magnitude of electrical current, etc (IEEE Standard, 2012).

The results presented indicate that the corona losses of high-voltage transmission lines are dependent on various factors. Here, corona loss of higher magnitude is observed when the conductor temperature is lower,

especially during winter. To support the conclusions, Figs. 9 and 10 show the maximum corona losses when the conductor temperature is below −10 ℃ for new transmission line and −15 ℃ for old transmission line. Here, the loading factor also plays an important role that influences conductor temperature, which affects the formation of snow or icing conditions on the conductor surface, resulting in a lower magnitude of corona loss.

4. Conclusion

The paper presents an analysis and comparison of corona losses monitored on two transmission lines. The results indicate that corona losses tend to be higher on the older transmission line. This is assumed to be mostly due to aging of the conductors (nearly 50 years difference) and a less optimized design (two conductors per bundle vs three). While it was not possible to compare the lines in exactly equal conditions, significant differences were observed in similar conditions. Similar analyses could be used in other networks with sufficient monitoring capability to compare the corona loss performance of transmission lines. The results can be useful for better overview of losses, improved forecasting of losses, considered as a factor in network planning decisions, etc. Further studies on more transmission lines would help to understand the effects and differences in more detail.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

The Estonian TSO, Elering, provided measurement data for this study. Weather data from the Estonian Environmental Agency has been used.

References

Bian, X., Zhao, X., Cao, J., Gu, L., Wang, L., Guan, Z., 2010. Aging effects of conductor surface conditions on ac corona discharge. In: Proceedings of the Annual Report Conference on Electrical Insulation and Dielectric Phenomena. IEEE, pp. 1–4.

Carroll, J.S., Cozzens, B., 1993. Corona loss measurements for the design of transmission lines to operate at voltages between 220 kV and 330 kV. Trans. Am. Inst. Electr. Eng. vol. 52 (1), 55–62.

IEEE Standard, 2012. IEEE Standard for Calculating the Current-Temperature Relationship of Bare Overhead Conductors, 738–2012, pp. 1–67.

Llamo-Laborí, H.S., Santos-Fuentefria, A., 2022. A new method to calculate corona losses for active conductors considering real transmission like unbalance. J. Eng.

Maruvada, P.S., 2000. Corona Performance of High-voltage Transmission Lines. Research Studies Press Baldock, UK.

Pan, W., Li, Y., Chen, X., 2019. Calculation method of corona loss in EHV/UHV system based on distributed parameter characteristic. IEEJ Trans. Electr. Electron. Eng. vol. 14 (5), 730–734.

Parizad, A., Dehghan, S., Saboori, H., Kazemi, A., 2011. Transmission network augmentation planning considering the impact of corona power loss. In: Proceedings of the IEEE Trondheim Power Tech, pp. 1–6.

Pavičić, I., Holjevac, N., Ivanković, I., Brnobić, D., 2021. Model for 400 kV transmission line power loss assessment using the PMU measurements. Energies vol. 14 (17), 5562.

Peek, F.W., 1911. The law of corona and the dielectric strength of air. Proc. Am. Inst. Electr. Eng. vol. 30 (7), 1485–1561.

Peek, F.W., 1920. Dielectric Phenomena On High Voltage Engineering. McGraw-Hill Book Company, Incorporated.

Peterson, W.S., 1993. Discussion. Trans. Am. Inst. Electr. Eng. vol. 52 (1), 62–63.

Rao, K.D., Rao, G.V., Ramesh, G., Govardhanachari, K., 2022. Comparison of used conductor for corona, radio interference voltage (RIV) characteristics and audible noise values with unused conductor. Power Electronics and High Voltage in Smart Grid. Springer.

Solovyev, Y., Gutman, I., Byrkjedal, O., Thorsteinsson, B.H., 2018. Practical methodology for calculation of corona losses induced by hoar-frost using operational measurements. CIGRE Sci. Eng. vol. 10.

Tuttelberg, K., Kilter, J., 2015. Real-time estimation of transmission losses from PMU measurements. In: Proceedings of the IEEE Eindhoven PowerTech. IEEE, 29 June 2015, pp. 1–5.

Tuttelberg, K., Kilter, J., 2018. Estimation of transmission loss components from phasor measurements. Int. J. Electr. Power Energy Syst. vol. 98, 62–71.

Tuttelberg, K., Löper, M., Kilter, J., 2019. Correcting systematic errors in corona losses measured with phasor measurement units. IEEE Trans. Power Deliv. 34 (6), 2275–2277.

Vasilenko, N., Gadzhiev, M., Galiaskarov, I., Zhgun, K., Korobka, V., Ryabchenko, V., Sharov, Y., 2020. Corona losses reduction of OHL 500 kV of Omsk electric power system based on signal processing of PMU. CIGRE Paris Session.

Appendix 5

V

P. K. Gupta, K. Tuttelberg, and J. Kilter, “The weather impact on corona losses of 330 kV aging transmission lines,” in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, *Dubrovnik*, *Croatia*, pp. 1–5, 2024

2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE) | 979-8-3503-9042-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISGTEUROPE62998.2024.10863024

The weather impact on corona losses of 330 kV aging transmission lines

Pradeep Kumar Gupta, Kaur Tuttelberg, and Jako Kilter, *Senior Member, IEEE*

***Abstract*—The main objective of this work is to investigate the characteristics of different weather parameters on corona losses for aging overhead transmission lines using synchronized phasor measurements. The study is conducted for a long-term assessment using two years of data based on PMU measurements collected from Estonian TSO and weather data from the Estonian Environmental Agency. This research gives insight into information about weather factors influencing corona loss for two different aged transmission lines using a statistical graphical study. It is generally believed that aged transmission lines are more sensitive as they are exposed to different environmental conditions, leading to the deterioration of the line performances. Therefore, it would seem appropriate to investigate the factors influencing corona loss in an overhead transmission line. These results can be the basis for transmission system operators to reduce the power loss purchase and improve corona loss management under complex weather conditions to increase power grid efficiency.**

***Index Terms*—Aging transmission lines, Corona loss, Real-time monitoring, Transmission system operator, Weather dependency.**

I. Introduction

The corona phenomenon generally occurs on all types of overhead transmission lines, however they become more visible in the case of high voltages. Many researchers from over a century of study on corona losses have proposed different methodologies to reduce corona loss in an overhead transmission line by using bundled conductors, corona rings, increasing the conductor diameter, eliminating the sharp edges, etc. The corona discharge phenomenon is caused by air ionization surrounding a high voltage-carrying conductor. When air is ionized during a corona, oxygen is converted to ozone, a highly reactive gas that corrodes and chemically destroys insulating materials [1]. Some of the important factors that change the magnitude of corona effects include a higher voltage, conductor surface status, spacing between conductors and weather conditions.

The recent fast growth stage and construction of a large number of high-voltage lines across the world lead the corona losses to be more noticeable. The relationship between corona loss and weather parameters increases with the development of high-voltage lines and the integration of renewable energy across the world. The difficulty of corona loss performance analysis depends on weather conditions. However, the detailed relationship between them is complex and should be verified over a longer time period. The longevity of insulating materials on overhead transmission lines depends on the ability of the material used to withstand the destructive effects of the environmental conditions, such as fast-changing temperature, humidity, emission of harmful gases corrosion, high wind accumulation of dust and other environmental factors [1].

Due to the limitation of the empirical method, many previous studies have conducted experiments based on corona cages and artificial climate chambers to examine the influence of climatic conditions on corona loss. Numerous research studies focusing on different weather conditions have developed practices to facilitate the issue of conventional empirical methods. Many researchers have demonstrated the importance of weather influence on corona losses using experimental methods (indoor and outdoor) with corona cage and climate chamber. Under different environmental conditions, such as icing [2], hoar frost [3], sandy and dusty [4], fog [5] and rain [6] conditions, corona loss behavior was observed.

Based on weather parameters research, corona loss assessment work has already been conducted by [7], [8] to explore, understand and illustrate the relevant dependencies. However, it didn't consider the aging factor and also failed to rank their importance. In a recently published article, the authors discuss the statistical analysis of weather dependency and ranking of important weather parameters that influence corona loss of a 330 kV overhead transmission line [9]. Despite this, further studies are needed to explore the weather dependency for aging transmission lines to understand the complex correlation.

The discussion of aged sand cast substation connectors for corona loss assessment was examined and presented in [10]. However, different aspects of the environment were a limiting factor. Similarly, in the case of used and unused conductors, the assessment of corona loss performance was discussed in [11]. In [12], the losses have been characterized in fair weather conditions by different contamination levels of a conductor. Nevertheless, concerning aging transmission lines, the corona loss performance using PMU measurement was introduced in [13] where monthly evaluation along with geometrical configuration and temperature dependency were analyzed. There has been limited research studies related to weather dependency for aging transmission lines using PMU measurement. Therefore, specific studies using operational data will help to determine the performance of aging transmission lines.

P. K. Gupta, K. Tuttelberg, and J. Kilter are with the Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Tallinn, Estonia (e-mail: pradeep.gupta@taltech.ee; kaur.tuttelberg@taltech.ee; jako.kilter@taltech.ee).

979-8-3503-9042-1/24/$31.00 ©2024 IEEE

The proposed paper presents a study to compare the dependency of different weather parameters on the corona loss for aging overhead transmission lines using actual network PMU measurement. The influence of corona loss magnitude is examined with weather parameters to account for potential variations depending on the aging of the overhead transmission line. This paper briefly examines weather dependency using statistical graphical analysis and the percentile distribution.

The construction of overhead transmission lines over different geographical regions makes them vulnerable to different extreme adverse weather conditions over time, which more often influences the corona loss in an overhead transmission line. The potential outcomes are identifying the distribution of corona losses as expected or a normal level of overhead line operation for complex weather conditions. This study supports the issues of identifying the most important weather parameters that affect the behavior of corona losses for two different aged transmission lines.

II. Estimation of Corona Loss Background

To estimate corona losses in the overhead transmission lines using synchronised phasor measurements units are referred from [14] where the PMU data are collected from both ends of the lines. Additionally, the correction method for PMU measurement is used to avoid systematic errors [15]. Uncertainty aspects of PMU measurements are discussed in [16]. Further, the weather parameters data are collected from the nearby weather station, a combination of 4-weather stations data refer to uniform distribution of weather conditions. The time-step for collected weather data is 10 minutes and estimated corona loss from PMU data are merged accordingly. The analysis uses two years (2018-2019) of data collected from the Estonia region. Each of the 330 kV overhead transmission lines are of different configurations and were constructed in 1959 and 2006. The older line stretches from north to south and is 170 km, whereas the new line direction is from west to east and is 210 km. The overall transmission lines configuration, along with tower configuration and how it impacts the corona loss performance, were studied in [13].

A. Corona Loss Performance Influence by Weather Parameters

1) Air Temperature: The dependence pattern for varying air temperature and corona loss is shown in Fig. 1 and Fig. 2. The analysis shows that the corona loss reflects its highest values when the temperature predominantly decreases during winter for both lines. Here, it is interesting to see how the varying air temperatures influence the corona loss on aging transmission lines. Fig. 1 demonstrates the performance of corona loss in a new transmission line which is significantly high during winter ($T < -5\ ^{\circ}C$) almost four times as compared to summer ($T > 5\ ^{\circ}C$). However, in the case of the old transmission line, the corona loss exhibits a rather complicated dependence on air temperature, as shown in Fig. 2, where higher corona loss is observed between $2\ ^{\circ}C$ and $-20\ ^{\circ}C$.

The hypothesis for high corona loss during winter is due to the formation of ice crystals and the collection of snow on the conductor's surface [7]. Here, we also investigated percentile levels to understand the corona loss distribution. This study explored the effects of air temperature at the typical distribution of corona loss using percentile levels – 95th percentile and 75th percentile. Percentile data plots show that the corona loss is generally instigated at a lower temperature.

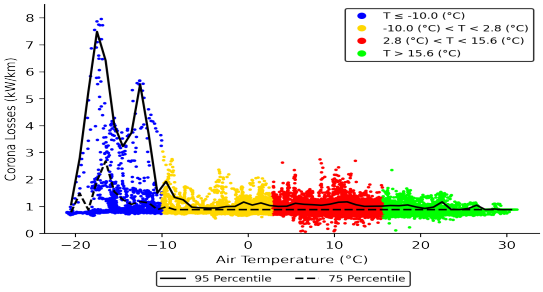


Fig. 1. New line air temperature dependence on corona losses.

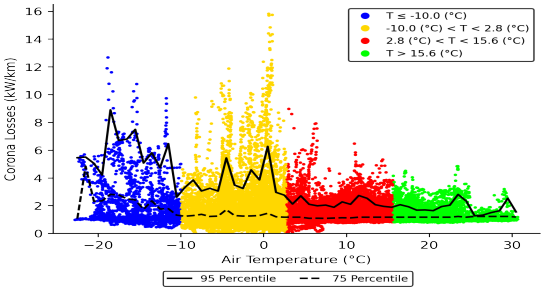


Fig. 2. Old line air temperature dependence on corona losses.

2) Air Pressure: The dependency between air pressure and corona loss under varying temperatures is shown in Fig. 3 and Fig. 4 for both lines. When looking at the relationship between air pressure and corona losses then seasonality has to be considered as well. In this type of climate, in winter high air pressure often coincides with cold air temperatures but in summer high air temperatures instead. Low air pressure in winter usually means milder conditions, while in summer it means rainier or damper weather. The expected increase in corona loss is clearly visible for a newer line when the pressure is increased during winter with a greater percentile distribution. Yet, in the case of older lines, the effect is not clearly distinguishable and corona loss peaks are sometimes visible at lower pressure.

The fact that we know from Peek's empirical equation is that at sufficiently lower pressure, the corona loss should be high. The reason for the higher corona loss value is that the difference between the operating voltage and disruptive voltage is increased, which is why the hilly region or low-pressure

area's corona loss is highest. However, our analysis reflects such dependency for the older lines and contrasting scenarios for newer lines due to winter where corona loss exhibits higher values at lower temperatures and high pressures.

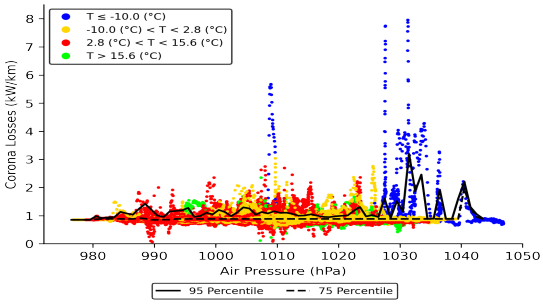


Fig. 3. New line air pressure dependence on corona losses under varying air temperature.

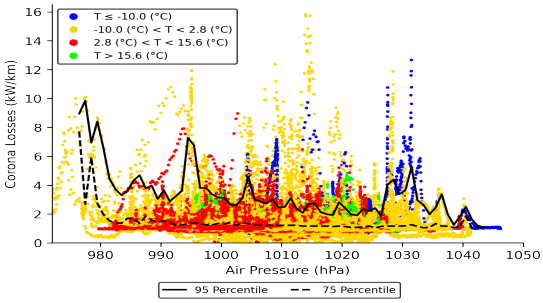


Fig. 4. Old line air pressure dependence on corona losses under varying air temperature.

3) Wind speed: To illustrate the relationship between corona loss concerning wind speed is shown in Fig. 5 and Fig. 6 for both lines. Generally, at lower wind speeds, a spark initiates the conducting path due to the instant ionization of air, which is maintained, whereas at higher speeds, due to ion mobility, the corona discharge phenomenon is disrupted leading to lower values of corona loss. The overall analysis for both transmission lines shows a decreasing pattern of corona loss values for higher wind speeds. Thus, the general dependence of the two cases is similar and will be of the same type, as shown in the figures. A similar trend was followed with minor differences among the percentile curves. The hypothesis can be viewed in terms of peak current density decreasing or having less influence on conductor temperature. The most common reason for low corona loss at higher wind speeds is the melting of snow or clearing of the conductor surface and increased ion mobility.

4) Wind Direction: In the case of wind direction, the results shown in Fig. 7 and Fig. 8 suggest a weak correlation between

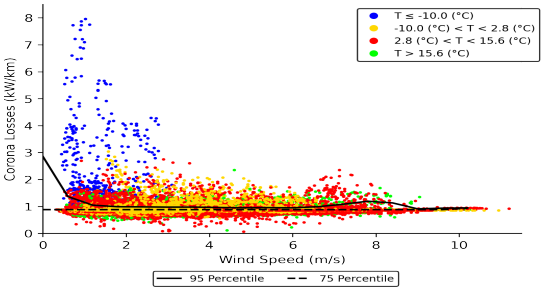


Fig. 5. New line wind speed dependence on corona losses under varying air temperature.

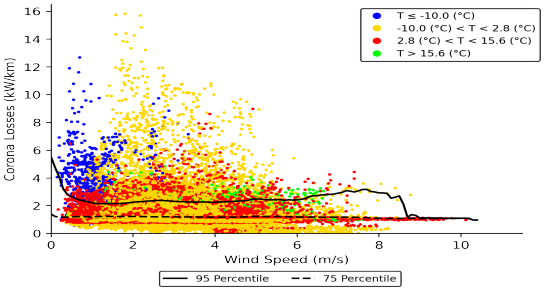


Fig. 6. Old line wind speed dependence on corona losses under varying air temperature.

wind direction and corona loss due to an even distribution of points but possible correlations of lower temperature with higher corona loss were exhibited for both the transmission lines. However, they show a variational difference in the magnitude for a certain angle due to the directions of lines (New line – west-east; Old line – north-south). The overall analyses of wind direction indicate a much weaker dependence even with the distribution percentile line.

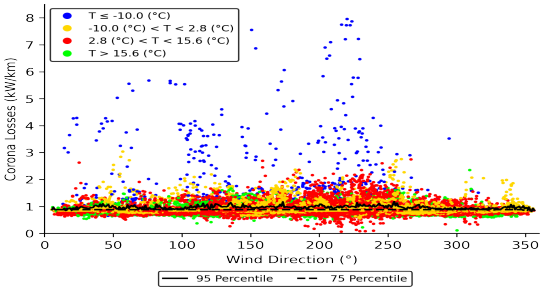


Fig. 7. New line wind direction (0-360 degree) dependence on corona losses under varying air temperature.

5) Air Density: Fig. 9 and Fig. 10 show the correlation plot of the air density correction factor on corona loss for both

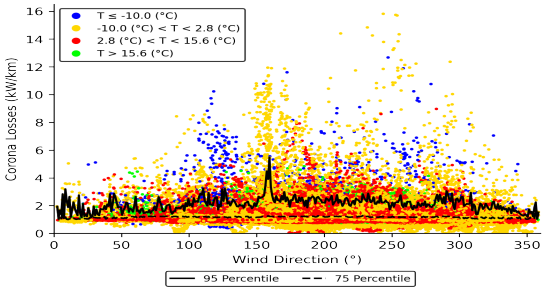


Fig. 8. Old line wind direction (0-360 degree) dependence on corona losses under varying air temperature.

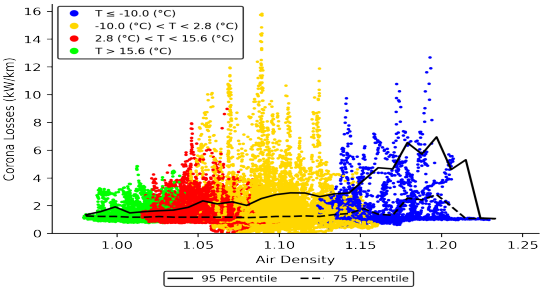


Fig. 10. Old line air density correction factor dependence on corona losses under varying air temperature.

lines. It should be noted that this analysis shows strong inter-dependence for both lines. The clear dependence is seen from the new line where high corona loss is identified for higher δ. The dependence for older line δ is quite complex, volatile and swiftly changes with weather conditions, as shown in Fig. 10. Peek's empirical analysis provides information about delta dependence on corona losses, offering limited support to this analysis. Although this study shows a high correlation between them at higher air density, the percentile values give more information about the dominant nature of higher corona loss data division. During bad weather conditions like rain, snow, and hailstorms, air density decreases, which in turn decreases critical disruptive voltage and increases the corona loss in the transmission lines. This is due to the fact that the accumulation of snow on the conductor surface forms sharp edges of icing, which increases the electric field intensity.

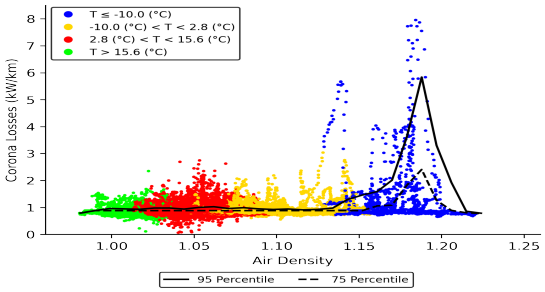


Fig. 9. New line air density correction factor dependence on corona losses under varying air temperature.

6) Visibility: The dependency between visibility and corona loss is shown in Fig. 11 and Fig. 12 for both lines. Our analyses show that many data points lie below 20 km, reflecting higher corona losses. The old transmission line is mostly affected with varied ranges of visibility and the impact on the new line is partial, where higher corona loss is only observed at lower temperatures.

The main hypothesis behind this is due to winter and less visibility resulting in the accumulation of moisture and humid nature on the conductor, leading to an increase in corona losses. In support of this, lower visibility leads to higher corona loss which is also referenced in [7], where the long-range visibility impact on corona losses is discussed. The corresponding percentile distribution of corona loss for the older line reflects high corona loss at lower visibility, although the same is unclear for a newer line.

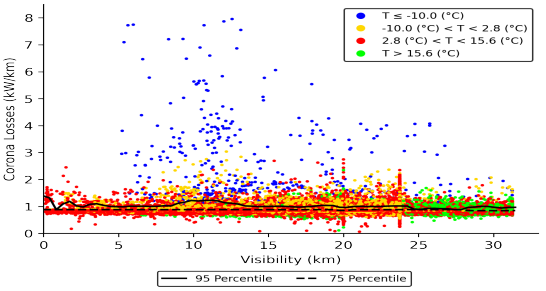


Fig. 11. New line visibility dependence on corona losses under varying air temperature.

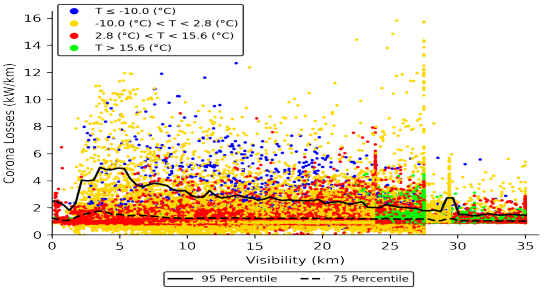


Fig. 12. Old line visibility dependence on corona losses under varying air temperature.

7) Precipitation: Fig. 13 and Fig. 14 show a typical case of precipitation for both lines. These analyses evaluate the

precipitation behavior where corona loss is mostly high for lower precipitation values. The results of the graphical analysis show that precipitation depends on different phenomena. In the case of weak/strong rainy conditions and during snow precipitation level changes, the corona loss impact potentially grows from weak to strong. Here, corona losses are strongly affected and reach the highest value at the lower precipitation value. The different categories of precipitation considering weak/strong rain or snow should also be discussed to understand the behavioral correlation pattern.

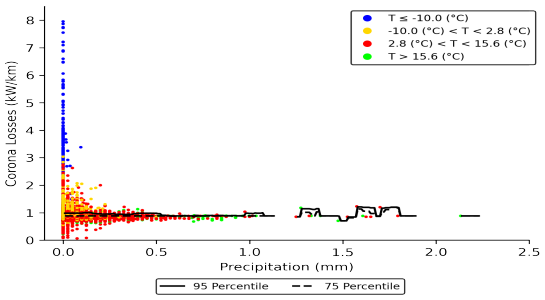


Fig. 13. New line precipitation dependence on corona losses under varying air temperature.

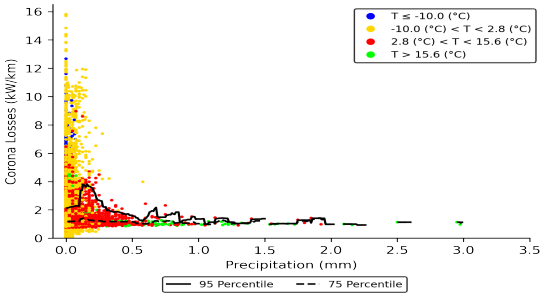


Fig. 14. Old line precipitation dependence on corona losses under varying air temperature.

In summary, the study analyzed in this section shows the behavior and dependency of different weather parameters on corona losses. Corona losses are strongest under low-temperature environments, lower precipitation levels and wind speed for both lines. However, in the case of high pressure and air density, the old lines have complex relations, whereas new lines show directly proportional relations. The study highlights the importance of different weather parameters affecting corona losses and are significantly influenced by air density, air pressure and air temperature.

III. Conclusion

The result indicates that the magnitude of corona loss differs significantly according to the age of overhead transmission lines. The statistical graphical study shows that the dependency of weather parameter influences on corona loss for a new line is less sensitive than the older transmission line. This study supports the issues of identifying weather-related resilience on corona loss for two different aged transmission lines, evaluating the effect of different weather parameters. These results help TSO to understand the distribution of corona losses for different aged lines determining which lines to reconstruct and adapt to future weather conditions.

Acknowledgments

The Estonian TSO, Elering provided the measurement data. Weather data given by the Estonian Environmental Agency.

References

[1] B. H. Chudnovsky, *Electrical Power Transmission and Distribution: Aging and Life Extension Techniques*. CRC Press, 2013.

[2] F. Yin, M. Farzaneh, and X. Jiang, "Laboratory investigation of AC corona loss and corona onset voltage on a conductor under icing conditions," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 3, pp. 1862–1871, 2016.

[3] K. Lahti, M. Lahtinen, and K. Nousiainen, "Transmission line corona losses under hoar frost conditions," *IEEE Transactions on Power Delivery*, vol. 12, no. 2, pp. 928–933, 1997.

[4] L. Yunpeng, H. Feng, L. Chen, L. Yongshuang, T. Jian, C. Shaoshuai, L. Daran, C. Sijia, and H. Shilong, "Corona loss of the bundle conductors on EHV/UHV AC power lines under sandy and dusty conditions in high-altitude areas," *Journal of Electrostatics*, vol. 107, p. 103476, 2020.

[5] S. Ghosh, N. Ahmed, and S. Banerjee, "Impact of weather (fog) on corona loss and its geographical variation within eastern region," in *2018 20th National Power Systems Conference (NPSC)*. IEEE, 2018, pp. 1–6.

[6] H. Kirkham, "The influence of rain rate on transmission line corona performance," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 1, pp. 420–430, 1981.

[7] F. J. Sollerkvist, A. Maxwell, K. Rouden, and T. M. Ohnstad, "Evaluation, verification and operational supervision of corona losses in sweden," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 1210–1217, 2007.

[8] J. C. Matthews, "The effect of weather on corona ion emission from ac high voltage power lines," *Atmospheric Research*, vol. 113, pp. 68–79, 2012.

[9] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109537, 2024.

[10] J.-R. Riba, S. Bogarra, Á. Gómez-Pau, and M. Moreno-Eguilaz, "Experimental study of the corona performance of aged sand-cast substation connectors," *Energies*, vol. 13, no. 11, p. 2785, 2020.

[11] K. D. Rao, G. V. Rao, G. Ramesh, and K. Govardhanachari, "Comparison of used conductor for corona, radio interference voltage (RIV) characteristics and audible noise values with unused conductor," in *Power Electronics and High Voltage in Smart Grid*. Springer, 2022, pp. 79–91.

[12] E. E. Mombello, G. Ratta, H. D. Suárez, and F. O. Torres, "Corona loss characteristics of contaminated conductors in fair weather," *Electric Power Systems Research*, vol. 59, no. 1, pp. 21–29, 2001.

[13] P.K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023.

[14] K. Tuttelberg and J. Kilter, "Estimation of transmission loss components from phasor measurements," *International Journal of Electrical Power & Energy Systems*, vol. 98, pp. 62–71, 2018.

[15] K. Tuttelberg, M. Löper, and J. Kilter, "Correcting systematic errors in corona losses measured with phasor measurement units," *IEEE Transactions on Power Delivery*, vol. 34, no. 6, pp. 2275–2277, 2019.

[16] K. Tuttelberg and J. Kilter, "Uncertainty propagation in pmu-based transmission line monitoring," *IET Generation, Transmission & Distribution*, vol. 12, no. 3, pp. 745–755, 2018.

Appendix 6

VI

P. K. Gupta, K. Tuttelberg, and J. Kilter, “Forecasting corona losses on high voltage transmission lines using machine learning,” *IEEE Transactions on Power Delivery*, vol. 40, no. 5, pp. 2696–2705, 2025

Forecasting Corona Losses on High Voltage Transmission Lines Using Machine Learning

Pradeep Kumar Gupta, Kaur Tuttelberg, and Jako Kilter, *Senior Member, IEEE*

***Abstract*—This paper presents the application of machine learning in forecasting corona losses on high voltage overhead transmission lines. Forecasting of corona losses is challenging due to a highly-complex relationship between weather conditions and corona losses. To develop a forecasting model for corona losses, two years of weather data from different weather stations along the transmission lines and respective PMU measurement data from both ends of the line were considered. Machine learning algorithm-based forecasting models, e.g., XGBoost and ensemble random forest (ERF) regression models with multiple weather inputs, were used. The forecasting scenario was divided into four frames, i.e., different time-step based forecasting, seasonal forecasting, combined transmission lines corona losses forecasting for larger grid networks, and the effect of reduced features (weather parameters) on the accuracy of predicted corona losses. The best case performance was an error in the range of ± 0.5 kW/km for 98% of the data points and a root mean squared error of 0.16 kW/km. Precise forecasting of corona losses can assist Transmission System Operators in cost savings and increasing the reliability of the power system.**

***Index Terms*—Corona losses, Ensemble Random Forest Predictor (ERFP), Forecasting, Machine Learning, Phasor Measurement Unit (PMU), Random Forest (RF), Real-Time Monitoring, Transmission System Operator (TSO).**

I. Introduction

Corona discharge phenomenon mainly occurs on high-voltage overhead transmission lines, and in favourable conditions can cause excessive losses in the network. To balance the day-ahead energy market, the transmission system operator must accurately forecast system losses. Generally, any imbalance costs have to be covered by the TSO with additional purchases on the intra-day imbalance market, which increases energy costs. Forecasting corona losses is not straightforward and depends on various interrelated factors. While there are various methodologies available in the literature, they have limitations when it comes to considering complex weather conditions with different weather parameters. These methods have usually been intended either for longer term planning or analyzing specific scenarios, not operational forecasting.

Empirical equations have predominantly been used to estimate corona losses [1], [2], but they also have limitations in considering complex weather conditions. In recent years, statistical and machine learning-based approaches have been applied to transmission loss forecasting, with potential to also be applied to corona losses. On the beneficial aspects of active power loss forecast in transmission networks using machine learning were discussed in [3]. In [4], a commercially deployed machine learning system that utilizes the CatBoost algorithm has shown promising results in day-ahead planning for expected grid loss for a Norwegian utility company. The authors in [5] employ a back propagation neural network model to compute line loss in the distribution system, and it is used to predict line loss in a certain area. Nevertheless, these approaches do not consider corona loss and are utilized at the distribution level.

In [6], a forecasting model based on a statistical approach to monitoring the hourly corona losses was proposed, with the objective of saving the imbalance costs for Bulgarian TSO. The authors of [7], presented a novel, short-term forecasting model including a fuzzy-based weight clustering method for active power loss estimation using power flow analysis for the forecasted day. To predict the transmission line losses, the authors of [8] use the heat balance equation for conductor temperature estimation, which was given as input data for the numerical weather forecast combined with GIS mapping of the line assets and the day ahead hourly forecast of the power flow. A prediction model for transmission line losses based on linear regression and exchange flow modelling applied to the Swedish National Grid was proposed in [9]. Recently in [10], the author proposes different machine learning algorithms such as ARIMAX, SARIMAX, Feedforward Neural Network, etc, to forecast day-ahead grid loss, which help to minimize the risk for both producers and consumers, facilitating a less volatile energy market by reducing the occurrence of high imbalance prices; however, neglecting the impact of the corona loss significantly undermines accuracy and should not be overlooked.

In addition, various efforts have been made to use advanced methodologies to improve transmission loss forecasting performance. In [11], machine learning algorithms such as support vector regression (SVR) were extensively used by the Croatian Transmission System Operator (HOPS) to forecast short-term active power losses with weather forecasts as input data. In [12], the author proposed a long short-term memory (LSTM) model to predict overall grid losses but this approach faces difficulties due to large fluctuations of short-duration corona loss. A more advanced method was presented in [13], which uses ensemble empirical mode decomposition (EEMD) to decompose the line loss and proposes the EEMD–LSTM–SVR prediction model. Based on the literature it can be concluded that most of the proposed approaches and analyses are made

This work was funded by the European Union and co-funded by the Estonian Ministry of Education and Science (Project No. TemTa-134).

P. K. Gupta, K. Tuttelberg, and J. Kilter are with the Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Tallinn, Estonia (e-mail: pradeep.gupta@taltech.ee; kaur.tuttelberg@taltech.ee; jako.kilter@taltech.ee).

for overall transmission line losses and some of these are even neglecting corona loss influence and there is limited information on influence of multiple weather parameters on the level and extent of corona losses.

More precise corona loss estimation can be made using synchronised phasor (PMU) measurements. In [14], [15], transmission line parameters are estimated based on PMU data and these are consequently used to estimate line corona losses. The assessment of the influence of different weather parameters on corona loss in high voltage lines using PMU measurements is discussed in [16]–[18]. In general, corona losses are difficult to measure and forecast. Even if the losses can be measured on some of the transmission lines, it is still difficult to monitor a larger region or the whole power system. Therefore, a corona loss forecasting model could help to improve overall transmission network loss forecasting accuracy, but it could also provide input to monitor corona losses on a larger area level, i.e., determine the corona losses on lines where the losses cannot be measured. To the authors' knowledge, no research studies have comprehensively evaluated the practical use of PMU measurement data and different weather data for forecasting corona losses.

In this paper, a machine learning based model for forecasting corona losses in 330 kV overhead transmission lines was developed and its performance was analysed in different scenarios. In the analysis, the following aspects were considered: i) influence of time-step between weather measurement data points, ii) seasonal trends to assess the behaviour of the forecast model during a particular season, iii) combined forecasting on two transmission lines to analyze forecasting on larger parts of a grid, and iv) training the forecast models with reduced input features to determine the most important weather parameters needed for forecasting. The aim is to determine the feasibility of predicting corona losses with machine learning methods. If sufficient accuracy is achievable, the models could be used to forecast corona losses based on weather forecasts for short time horizons—from a few hours to a few days. While the study is based on the example of 330 kV lines, there is nothing to limit applying the method at higher voltage levels than that.

This research focuses on evaluating the corona loss forecasting approach on historical data of measured corona losses and recorded weather data. There is potential to apply the model to forecast corona losses based on short-term weather predictions but is difficult to obtain the historical data of weather forecast results required to evaluate such an approach. The study presents the forecasting task on the Ensemble Random Forest machine learning method, but also presents some results with XGBoost and a simple linear regression model. Performance is measured with RMSE and MAE metrics, but also evaluated more in detail by analysing how well larger values of corona losses are predicted by the model. The study presents the techniques which were found to increase the accuracy of forecasting.

The structure of this paper is as follows: Section I features an introduction and discusses recent literature on corona loss forecasting methodologies. In Section II, the working principle and proposed methodology are described. Section III addresses the result analysis and discussion. Finally, Section IV concludes the paper.

II. Methodology

The basic flow chart shown in Fig. 1 describes the overall forecasting process and estimation model of corona loss in high-voltage transmission lines using PMU data. It is divided into two parts: the first describes the PMU data collection and estimation of different overhead line losses with error correction method along with weather data collection, and in the second part, data is preprocessed by removing outliers and missing values. Furthermore, to significantly improve the forecast model's accuracy, ERF regression with different hyperparameters was tuned for effective training and testing. Two sets of error metrics were used for model evaluations considering different case studies to determine the accuracy and reliability of metrics.

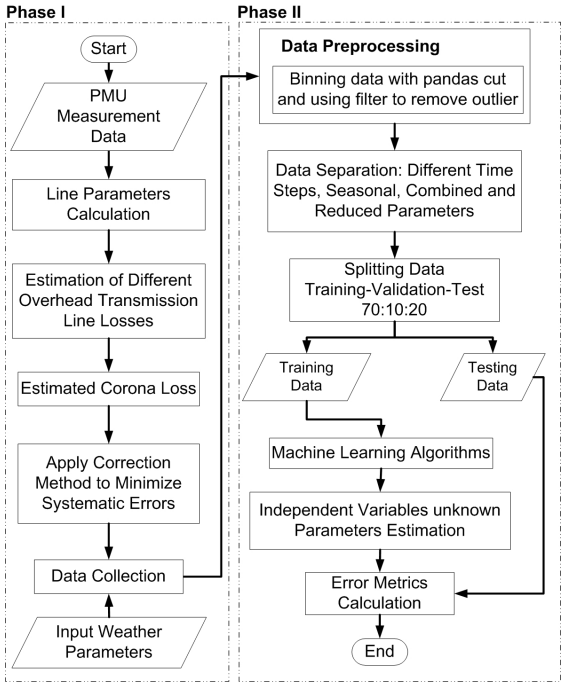


Fig. 1: Flowchart of corona loss forecasting model using machine learning.

A. Data collection and preprocessing

Actual recorded PMU data were obtained from the Estonian TSO, Elering, to test the forecasting model. The availability of PMU measurement data from both ends of the overhead transmission line was required to perform this analysis. The PMU data collection was carried out between 2018 and 2019 to assist in monitoring the corona losses of 330 kV overhead transmission lines in Estonia. This data is further being used

to measure line parameters, which help to estimate different types of losses such as Joule, corona, inductive, and capacitive losses.

Meteorological weather data was collected from Estonian weather stations, including air temperature, humidity, air pressure, precipitation, visibility, wind speed and wind direction [19]. The case studies presented use hourly data to estimate the influence of weather parameters and forecast corona losses. To avoid the scenario of a change in weather conditions concerning sending and receiving ends, data from four weather stations were collected along the direction of each overhead transmission line. The approximately 170 km (north-south) and 210 km (east-west) long transmission line corridors of the 330 kV lines, together with different locations of nearby weather stations, are illustrated in Fig. 2. The transmission line's tower configuration and design parameters are discussed in greater detail in [20]. The corona loss data was preprocessed by eliminating outliers and reducing the systematic measurement error using the error correction method [21].

To gain a comprehensive understanding of the rate of change in weather conditions along the transmission line, numerous weather parameters were selected, depending on the availability of the data from the weather station. The weather stations along the line are situated approximately 10-20 km from the line corridor. The elevation difference along the line is approximately 55 m, characterized by predominantly flat terrain. Combining different weather station data sets will enhance the detection of real climate variations along the transmission line.

Numerous techniques have been developed for merging the meteorological weather data in multiple ways from different locations, such as linear combination, exponential combination, MAPE-based combination, geometric mean combination, two-fold combination, GA-based combination, and sophisticated weighting techniques, etc. [22]. However it is very difficult to select which method would be suitable for combining the data, therefore the most common approach i.e. averaging method is used to combine data from multiple weather stations. In [16], a comparison of mean values and standard deviations for the weather parameters clearly demonstrates that the differences between averages for multiple stations and a single station are minimal, where the difference observed is very small, suggesting a consistent weather pattern along the transmission line.

B. Corona loss estimation using PMU measurement

The algorithm developed for real-time monitoring of corona losses is based on the principles where PMU measurement data is collected from both ends of the lines, which contains instantaneous values of voltage and current phasor information [15]. Further, the line parameters were estimated using the measurement data for different types of losses by distributed line equation based on the following equations:

$$\underline{\gamma} = \frac{1}{L} arcosh\left(\frac{\underline{V}(l_s)\,\underline{I}(l_s) + \underline{V}(l_r)\,\underline{I}(l_r)}{\underline{V}(l_r)\,\underline{I}(l_s) - \underline{V}(l_s)\,\underline{I}(l_r)}\right) \tag{1}$$

$$\underline{Z}_c = \frac{\underline{V}(l_s) - \underline{V}(l_r)\cosh(\underline{\gamma} l_s)}{\underline{I}(l_r)\sinh(\underline{\gamma} l_s)} \tag{2}$$

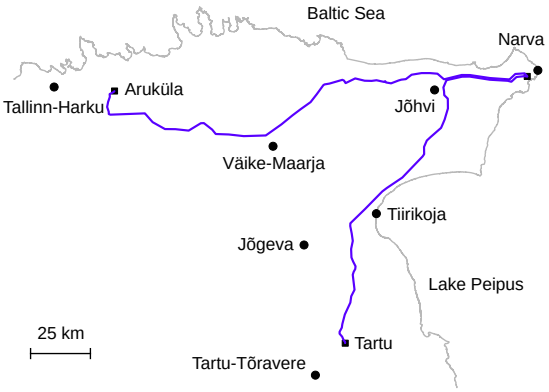


Fig. 2: Map of monitored 330 kV transmission line corridors. Lines are coloured blue, substations are denoted by square markers and weather stations by round dots.

where L is the length of the transmission line (km), $\underline{V}(l_s)$, $\underline{I}(l_s)$, $\underline{V}(l_r)$ *and* $\underline{I}(l_r)$ are the voltage and current phasor values at the sending and receiving end.

The per unit impedance and admittance are represented by $\underline{Z} = R + jX = \underline{\gamma}\,\underline{Z}_c$ $[\Omega/km]$ and $\underline{Y} = G + jB = \underline{\gamma}/\underline{Z}_c$ $[S/km]$.

The active transmission loss and reactive power balance can be represented by integrating over the line length, as shown below:

$$\Delta P = R\int_0^L |\underline{I}(l)|^2\, dl + G\int_0^L |\underline{V}(l)|^2\, dl \tag{3}$$

$$\Delta Q = X\int_0^L |\underline{I}(l)|^2\, dl - B\int_0^L |\underline{V}(l)|^2\, dl \tag{4}$$

where $\underline{V}(l)$ and $\underline{I}(l)$ are the voltage and current phasor values at distance l along the line. The dependence of the dissipation losses across the transmission line network is mainly due to resistive (R) losses, consumption of reactive power in reactance (X), conductance (G) losses (due to the leakage current and corona phenomenon) and generation of reactive power in susceptance (B).

C. Error correction method

The correction method to reduce the systematic measurement error that occurs in the PMU measurement for corona loss estimation was proposed by [21] based on the proportionality linear fit of shunt parameters to the power flow.

The corrected corona loss values are expressed as follows:

$$P_C(t) = T_C(t, z(t)) - B(x(t)) \tag{5}$$

where $B(x(t))$ is defined based on an approximation linear model.

$$B(x(t)) \cong D\,x(t) + G - F \tag{6}$$

where $x(t)$ is the assumed value of average power flow and the constant values D, G, and F are assumed for a given transmission line.

D. Machine Learning: Regression Model

In this analysis, multiple linear regression, ensemble random forest regression and extreme gradient boost regression models were used.

1) Multiple linear regression model: The multiple linear regression model uses multiple inputs to understand the dependent and independent variable relationship. By examining the weather conditions using multiple weather parameters in a region — input variables or features — we will develop a model that forecasts corona losses, which is the target variable. In a multiple linear regression model, a bias constant is applied to the weighted sum of the input variables to estimate the target variable predicted value, expressed by the equation below:

$$Y = \alpha_0 + \omega_1 X_1 + \omega_2 X_2 + \ ..... \ + \omega_n X_n \qquad (7)$$

where Y is the corona loss (dependent or target variable), X_1,X_2,...,X_n are the selected weather parameters collected from four weather stations between both transmission lines (independent variable), ω_1, ω_2,...,ω_n are the weighted slope coefficient and α_0 is the bias constant or intercept.

Multiple linear regression was selected for comparison mainly for two reasons:

- To understand the linear relationship and the strength associated between a dependent variable and independent variables.
- It is the most widely used method implemented easily to understand the rate of change.

2) Ensemble random forest regression:

Breiman in [23] proposes the concept of the random forests approach, which uses a series of decision trees that grow in randomly chosen data subspaces to form a constructive prediction ensemble, where each tree collection is created by selecting a group of inputs at each node. For developing a regression tree with greater reliability, a subset of predictive factors is taken into account to be examined at each node, increasing the prospects for weak correlation variables to assist in the development of a regression tree. An ensemble random forest (ERF) regression is a meta-estimator that is implemented with the aim of improving the overall model performance by aggregating multiple random forest estimators with different hyperparameters. The prediction model of multiple random forest estimators generates many trees with randomised subsets of predictors and uses averaging trees to improve the accuracy and overfitting.

The random forest method is widely used for different purposes to solve real-world problems such as partial discharge localisation [24], power quality assessment [25] and in various other power systems applications. Statistical methods were used in a few of the references, but the extensive use of the machine learning model (ERF) for corona loss forecasting using PMU data has not been evaluated. The network architecture of the ERF model is shown in Fig. 3. In order to handle corona loss forecasting, this work uses the ensemble system and random forest theory. To combine conceptually distinct machine learning regressors, i.e., to ensemble random forest regression with different hyperparameters, the VotingRegressor concept is used to predict the mean output value [26]. The benefit of

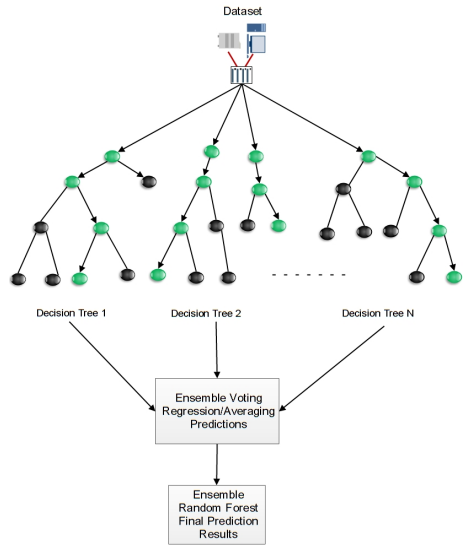


Fig. 3: Network architecture of Ensemble Random Forest Regression.

using ensemble systems is shown in issues involving complex input data for a problem such as corona loss forecasting highly dependent on complex weather conditions. The ERF models can manage nonlinear functions efficiently and are useful for real-time forecasting. Also, to counteract the flaws of a group of random forest regression models that perform equally well, VotingRegressor can be helpful for balancing. The majority vote system is used to combine the final forecasted value.

The most important parameters of the random forest model are the number of trees in the forest (forest size), the number of predictors available parameters (features) which are used and the maximum depth of the tree nodes (leaf size) required. The hyperparameters tuned for the ensemble random forest model are listed in Table I. Following that, the combination of three RF algorithms of different hyperparameters was integrated to create a voting regressor algorithm that voted for the best predictor among them to fit the final trained ERF regressor. Further, on the testing set, the trained ERF final predictions were carried out.

TABLE I: Hyperparameters tuned for ERF Regressor

Hyperparameters	Experimental Chosen Values
Number of Trees (n_estimators)	100, 500, 1000
Maximum Depth of Tree (max_depth)	20, 30, 50

3) Extreme Gradient Boost Regression (XGBR): The XGBR technique is based on a decision tree model, which is an efficient ensemble learning approach. A gradient-boosted decision tree variation created specifically for speed and performance is termed as XGBR. When it comes to creating supervised regression models to address complex and nonlinear relationships between input and the target variables,

XGBR generates superior results. It also facilitates parallel processing, which makes it effective for handling big datasets and cutting down on training time.

The different case studies were analyzed using the Extreme Gradient Boost Regression model, with three-fold cross-validation, which results in most of the parameters set as default, resulting in a good performance, i.e., n-estimators=100, learning rate ($\eta = 0.1$), max depth=6, min child weight=1, tree method='auto', n jobs=1, sample by level node and tree.

E. Error metrics for evaluations

The two sets of error metrics are chosen to identify and understand the behaviour of the forecasting model and how much error differs for an individual result. RMSE is most widely used in the case of regression; accordingly, the model performance is evaluated. However, MAE is widely used in industries and is therefore considered to check the suitability of different metrics to recommend and evaluate how well various metrics capture error distributions.

1) RMSE: The root mean square error (RMSE) is regularly employed in machine learning models for evaluation purposes. Sometimes RMSE is not a good indicator of average model performance and might be a misleading indicator of average error, therefore MAE and other metrics should also be checked. RMSE is widely used if the outliers are to be detected.

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y'_i - y_i)^2} \tag{8}$$

2) MAE: It measures the average size of the mistakes in a collection of predictions without considering their direction. It is measured as the average absolute difference between the predicted and actual values and is used to assess the effectiveness of a regression model.

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y'_i - y_i| \tag{9}$$

Here, m is the total number of non-missing data points, w_i is the weights multiplier, y' is the forecasted value and y is the actual true value of the dataset.

III. Results and Analysis

In this section, the experimental results analysis is discussed for the operational forecasting model performed to forecast corona losses using a machine learning model depending on various scenarios listed below. The performance was analyzed with models trained on data collected over two years. The dataset is divided into training, validation, and testing sets in roughly a 70%, 10%, and 20% randomized split where validation data is needed. However, in the case of seasonal analysis where separate validation in model building is not needed, 80% of the data is used for training and 20% for testing. The dependent variable in the regression model is corona loss estimated from PMU measurements and the independent variables are input weather parameters collected from four weather stations for each transmission line, including air temperature, air pressure, humidity, precipitation, visibility, wind direction and wind speed.

A. Varying Time Step Length

This section outlines the comparison results that investigated the impact and role of different time steps of the operational forecasting model performed to forecast corona losses using different machine learning models trained with shorter time steps of estimated corona loss and interpolated weather data in different strategies detailed below.

Generally, weather data is available with hourly resolution and the most natural way to create the dataset is by taking weather data as it is and downsampling the estimated corona loss data to an hourly time step. This results in a dataset that has slightly over 15 000 data points after the data has been cleaned of missing values and other issues. On the training data separated from that dataset, three regression models were trained – linear regression, ensemble random forest, and XGBoost regression model. The performance of these models is analysed based on residuals of predictions on the validation data and their histograms, shown in Fig. 4.

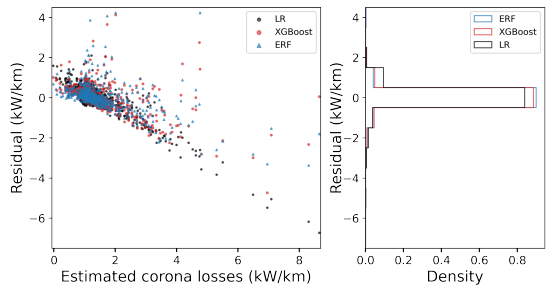


Fig. 4: Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the hourly dataset.

The models trained on the hourly dataset reveal three main findings. Firstly, the majority of estimated corona loss values are in the range of baseline corona losses which are present in normal conditions and all models have relatively small deviation in the predictions. Secondly, in the range of higher corona losses—which is the main point of interest—the predictions are poor and the number of observations is small. Thirdly, the linear regression model is the least capable of predicting losses in the higher corona loss range while more advanced regression models show better performance. Based on these observations, we can see that the number of high corona loss observations has to be larger. Furthermore, when analysing model performance, it is justifiable to concentrate on observations of higher corona losses.

However, hourly estimated corona loss data loses the capability to detect smaller variations and misses the greater detail of higher corona loss. While high precision measurement data will potentially improve the performance of the forecasting model. This study's outcome suggests that a data-rich environment with a precise larger sample size has a role in obtaining good forecasting performance. In order to increase the datasets and expose the models to a larger amount of training data, we

propose a hypothesis that the amount of weather data can be expanded by interpolation with shorter time steps to improve performance to detect smaller spikes and higher corona losses.

Weather phenomena are normally relatively slowly evolving and changes in physical quantities like ambient temperature, air pressure, and humidity occur gradually within an hour. It can be assumed that the changes are sufficiently slow and smooth so that the values for these quantities can be interpolated within an hour down to a reasonably short period of time. As such, we propose to interpolate the weather parameters to time steps of 10, 5, and 2 minutes, a linear interpolation is performed that is efficient to implement and involves filling the empty values based on the known values with a linear fit. Larger datasets are obtained when using the interpolated weather data and estimated corona loss value for each corresponding time point.

Similarly to hourly data, regression models were fitted to the 10 min, 5 min, and 2 min time-step datasets. The industry standard error metric mean absolute error (MAE) for predictions on validation data with each of the models trained on each of the datasets is presented in Fig. 5. While the ensemble random forest models present smaller errors and consistently strong performance with the extended datasets, then, overall results are not conclusive about whether any improvements were made. However, as we observed in the analysis of models trained on hourly data, it is important to take a more detailed look at corona losses in the range of higher values.

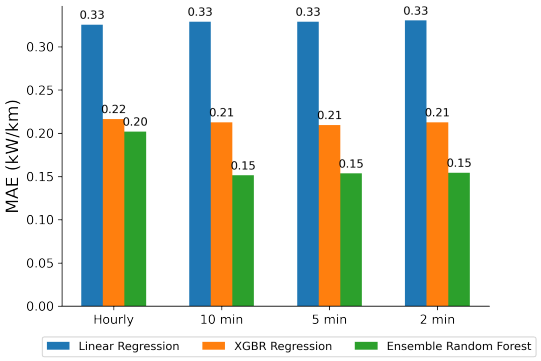


Fig. 5: Evaluation of error metric MAE of forecasted corona loss values using machine learning models for different time steps.

We analyse the ERF models trained on the datasets with different time-steps, but only look at the results of predicting losses where the target loss values were high. We take the 90th percentile value of estimated corona losses and look at all data points above that value (i.e. top 10% of losses). The residuals of predicting losses on the validation data and their histograms are presented in Fig. 6. These results indicate that in the area of interest, i.e. higher corona losses, the models trained on datasets extended by interpolating weather data perform considerably better. The distribution of residuals shows that a larger number of predictions are closer to the reference value of estimated corona losses. The residual plot generally represents the prediction error based on the difference between the observed value of the target variable and the predicted value. The histogram also shows that our error is mostly distributed around 0, indicating a better fitted model for decreasing time-step models.

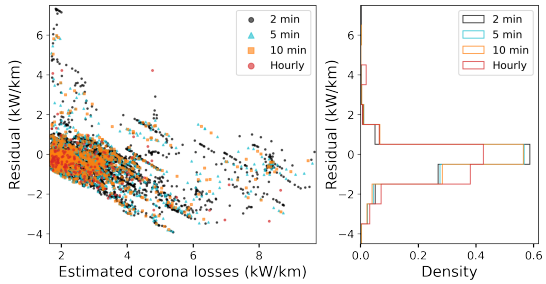


Fig. 6: Residuals and their histograms comparing the performance of the ERF model's forecasting capability with different time steps. Highest 10% of losses are presented.

In order to compare the performance of different models, we analyse the results of predicting losses on the validation data with models trained on the 2 min dataset. The residuals of predictions and their histograms are presented in Fig. 7 for the linear regression, ERF, and XGBoost models. As was already demonstrated by previous results, the linear regression model is not capable of predicting any of the higher loss values – the residuals show that the model predicts an almost constant value. Both the XGBoost and ERF models can predict most of the higher corona loss events, but the accuracy of the ERF model is clearly better as indicated by the histogram. On the test set with the full range of corona losses, the 2 min ERF model predicts results within ± 0.5 kW/km 93% of the time.

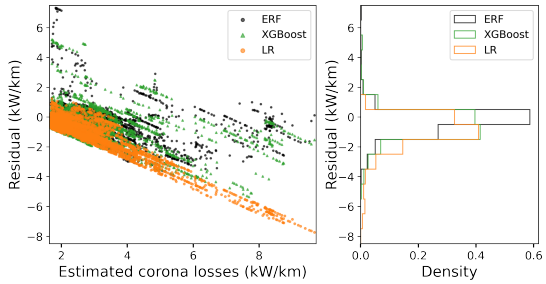


Fig. 7: Residuals and their histograms comparing the performance of different forecasting models—Linear Regression (LR), XGBR Regression (XGBoost), Ensemble Random Forest (ERF)—using the 2 minutes data set. Top 10% of losses.

Even though the time step could be decreased further, there is a limit to the efficacy of this. As can be seen from the results with two-, five-, and ten-minute interpolated data, the increases in accuracy become diminishing. Also, the computational time for training the models increases with every decrease in time

step as the dataset increases. In this study, the two-minute time step was found to be suitable for good accuracy and reasonable model training times. However, for smaller datasets, an even shorter time step might be more suitable.

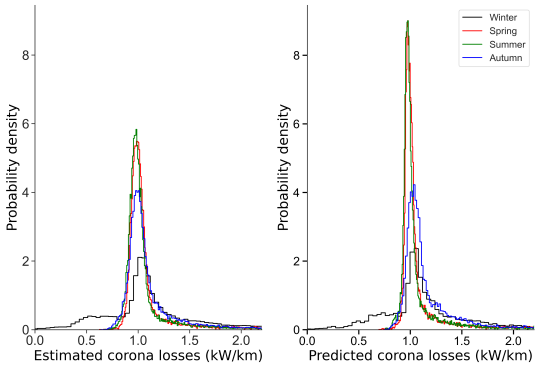
B. Forecasting With Seasonal Models

Generally, classical seasonal patterns of increasing corona loss are observed during winter and decreasing patterns in summer. Due to weather uncertainty, the corona loss might increase or decrease depending upon various other weather parameters. Atmospheric variability significantly influences corona loss on several time scales, from small-scale turbulence to day-ahead weather or seasonal anomalies. Therefore, to explore the correlation between different seasonal trends, it is vital to analyse the performance of different forecasting models for seasonal variation to understand their benefits and drawbacks. The input weather parameters are the same as in case A.

For the analysis, the monitoring data was split into four seasons – winter, spring, summer, and autumn – based on the timestamps. Each of the seasons was taken as a separate dataset and a regression model trained for each. This was done in order to compare if the prediction models capture the seasonal differences. The results are presented in Fig. 8 where approximated probability density functions have been plotted — corona losses estimated from measurements are presented in the left subplot and predicted losses in the right subplot. In spring and summer the majority of data points for corona losses represent low baseline losses and the probabilities are higher in these ranges. In autumn and winter the tail of the distribution is longer and higher, showing that higher corona loss data points are more frequent. The prediction probability densities have higher peaks due to predicting closer to the average baseline corona loss more frequently. Overall, the seasonal characteristics are very similar between estimated corona losses and the predictions of the seasonal regression models.

The seasonal corona loss forecasting models have been further analysed by plotting the residuals and their histograms for the highest 10% of corona losses in Fig. 9. Compared to the models trained on data for the whole year, the seasonal models perform differently. The residuals are more evenly distributed around zero and showing less of a slope compared to the results in Fig. 6. The error metrics of the seasonal models are presented in Table II. When combining the overall results of all of the seasonal models, the prediction error on test data is within ± 0.5 kW/km 98% of the time, which is better than the single model trained on the whole dataset. It can be seen that if the training data is collected over a period of several years, it can be beneficial to train seasonal models instead of or in addition to a model trained on the complete dataset. Predictions from such seasonal models can give more accurate results.

Seasonal models would make it possible to learn more specific behaviors, but the data must be gathered separately and the computational time is by far reduced in this training technique. The results analysis shows an advantage, effectiveness and the desirable accuracy of the machine learning



(a) Corona estimated using PMU data (b) Ensemble random forest model

Fig. 8: Measured corona loss (a) and forecasted corona loss (b) using the ERF model, for different seasons of the Estonian transmission network with a two-minute time step of interpolated weather parameters inputs.

TABLE II: RMSE and MAE comparison of corona loss forecasts for different seasons using a two-minute time step model.

Seasonal Period	ERF Model Error (kW/km)	
	RMSE	MAE
Winter	0.247	0.117
Spring	0.097	0.047
Summer	0.071	0.046
Autumn	0.163	0.090
Combined Annual	0.157	0.074

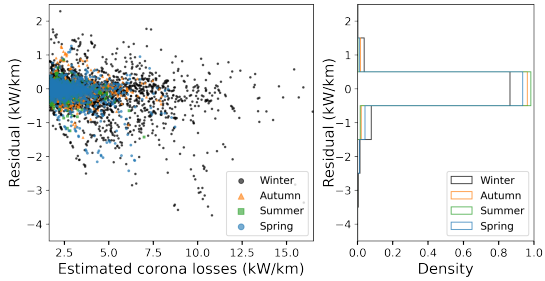


Fig. 9: Residuals and their histograms for the four seasonal ERF models with the two-minute time step for interpolated weather data. Highest 10% of losses are presented.

algorithm for forecasting corona losses using multiple weather parameters. As a result, we advise using more weather-related parameters in corona loss forecasting systems wherever possible, as these indicators can assist the model in generating better accurate forecasts. Overall, the suggested ERF model are more flexible in addressing the effects of seasonal behaviour.

C. Combined Transmission line Corona Loss Forecasting

In most cases a practical application of corona loss forecasting is concerned with the transmission losses of an entire power system or some area of a system. While a single

transmission line can represent a wider area, it would be beneficial to consider multiple transmission lines. Ideally the lines would cover a wider geographical area and have multiple weather stations in the vicinity. With data about estimated corona losses and weather monitoring for multiple lines, it is possible to better represent the overall corona losses of the system or an area.

For the purpose of this study, we have obtained the relevant data for two transmission lines. In order to evaluate the forecasting of losses for a larger area than a single line, we have trained a regression model for the combined losses of the two lines based on the estimated loss data and weather data from seven different weather stations. Similarly to the previous steps, we have used two minute interpolated weather data and trained an ERF model.

For the same dataset corresponding to a higher 10% of losses, residuals of the predictions and their histograms are plotted for the combined and single transmission lines as shown in Fig. 10 following a normal distribution using the ERF model. It shows a similar pattern corresponding to single-line for higher corona losses.

On the whole test set, the model trained for both of the transmission lines gives results within ± 0.5 kW/km 95% of the time, compared to 93% when trained on the data for the single transmission line. The error metrics given in Table III are also lower for the case of the two transmission lines. The differences in accuracy are small and the available data too limited to prove that using data from multiple lines clearly increases forecasting accuracy. However, small improvements were seen and using data from multiple transmission lines remains a potential technique to improve forecasting accuracy over a larger area.

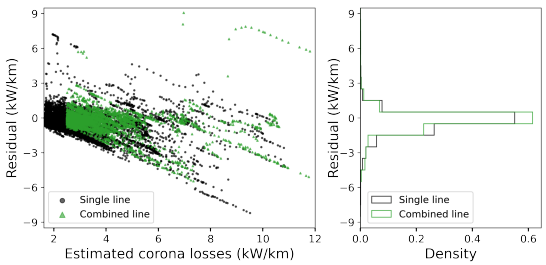


Fig. 10: Residuals and their histograms comparing the ERF models for the single line and the combination of two lines. Two-minute time step for weather parameters was used and the top 10% of corona losses were observed.

TABLE III: RMSE and MAE comparison of corona loss forecasts for a single line and combination of two lines considering all weather parameters.

Observed Lines	ERF Model Error (kW/km)	
	RMSE	MAE
Single line	0.401	0.154
Combined lines	0.384	0.147

D. Forecasting Performance With Reduced Features

This section analyses the effects of feature reduction, i.e. using fewer input weather parameters, on the accuracy of corona loss predictions. In [16], the author discusses the influence of different weather parameters using statistical graphical methods and feature importance methods. This paper seeks to extend research and analyse the forecasting results using the same set of the most important weather parameters reported previously, i.e., air pressure, air temperature and humidity. The minimum weather parameters data requirements for a good forecast of the corona loss are studied. For example, in some practical cases, some of the measured weather parameters might not be available in certain locations or might be missing for some time periods. As a result, observing how well the model performs when we need to utilize fewer inputs is interesting.

1) Single transmission line estimated corona loss (two-minute time step) with reduced features:

The feature reduction is analyzed by training the ERF model using different sets of input features. The sets of features, starting from the full set, are as follows

- RP_1 – Air Temperature, Air Pressure, Humidity, Precipitation, Visibility, Wind Direction and Wind Speed
- RP_2 – Air Temperature, Air Pressure, Humidity, Precipitation, Wind Direction and Wind Speed
- RP_3 – Air Temperature, Air Pressure, Humidity, and Precipitation
- RP_4 – Air Temperature, Air Pressure and Precipitation
- RP_5 – Air Temperature, Air Pressure and Humidity.

For the corresponding five models, the residuals of the predictions and their histograms for the top 10% of losses have been plotted in Fig. 11 and the error metrics for the whole test dataset have been presented in Fig. 12.

The results show almost identical performance for the first two sets of features (no features removed and only visibility removed). Removing wind direction and wind speed causes a slightly bigger reduction in accuracy than removing visibility but overall performance does not degrade significantly. Further reducing the input fetures down to three causes another decrease in accuracy, but the models still maintain some ability to predict corona losses. Overall, it can be seen that air temperature, air pressure, humidity, and precipitation are the most important weather parameters to utilize, but wind speed and direction can improve the accuracy too. The minimal set of temperature, pressure, and humidity can be measured with simple sensors and the model can still predict the majority of higher corona losses, even though the accuracy is smaller.

2) Combined two transmission lines corona losses (two-minute time steps) with reduced features:

To further illustrate the effects of using reduced input features, the ERF model for the combination of two transmission lines was also trained for the minimal set of input features—air temperature, air pressure, and humidity. Fig. 13 presents the residuals and their histograms for the ERF models trained on two transmission lines with the full set of features and reduced features. The top 10% of corona losses are considered like in

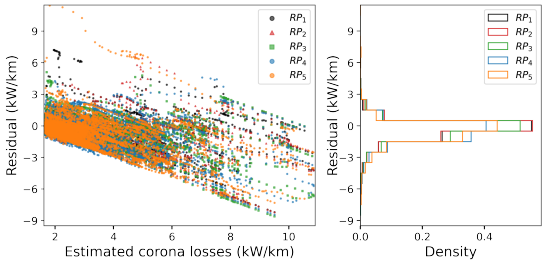


Fig. 11: Residuals and their histograms comparing the ERF models trained with different sets of input features. Two-minute time step for weather parameters was used and the top 10% of corona losses were observed.

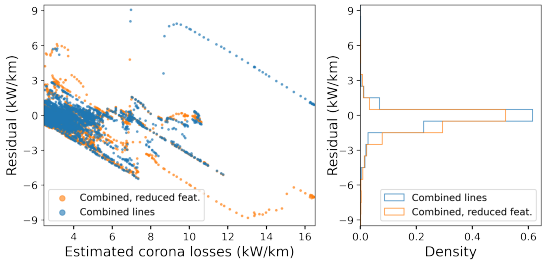


Fig. 13: Residuals and their histograms comparing the ERF models trained with all input features and reduced features on the combined two transmission lines. Two-minute time step was used and the top 10% of corona losses were observed.

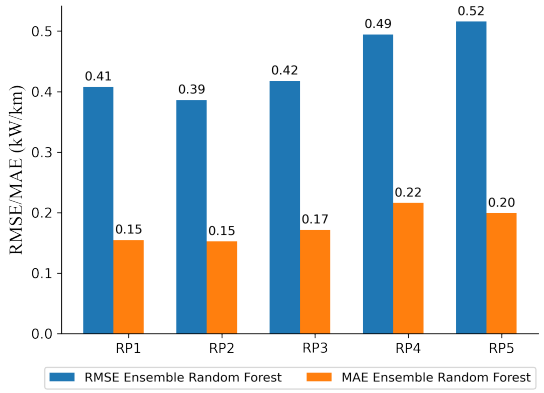


Fig. 12: RMSE and MAE of ERF model for forecasting corona loss in the Estonian transmission network using single-line data with different sets of input features.

the previous cases. From the histograms it can be seen that the prediction accuracy decreased when a smaller set of input features were used. The same is visible from the error metrics presented in Table III. While there is a visible reduction in accuracy, it does not make the prediction model unusable. Even if not all of the considered weather parameters are being monitored, training a forecasting model is still feasible.

TABLE IV: RMSE and MAE comparison of corona loss forecasts for the combination of two transmission lines considering all weather data and reduced features.

Feature Set	ERF Model Error (kW/km)	
	RMSE	MAE
RP_1	0.384	0.147
RP_5	0.489	0.193

- RP_1 – Air Temperature, Air Pressure, Humidity, Precipitation, Visibility, Wind Direction and Wind Speed.
- RP_5 – Air Temperature, Air Pressure and Humidity.

IV. Conclusion

In this study, the objective was to forecast the corona loss in high voltage transmission lines based on various weather parameters collected from different weather stations along the transmission lines. The study was based on the example of 330 kV lines, but the proposed method could be applied the same way at other voltage levels where significant corona losses occur. Three machine learning regression approaches were employed for the forecasting model, i.e., ensemble random forest, multi-linear regression and XGBoost regression model. During testing, the comparative evaluation of different regression machine learning models exhibited varying levels of performance in the forecasted value of corona loss for different case studies. In conclusion, the ERF model outperformed all other models during the evaluation to forecast corona losses. The effectiveness of the ERF model was justified by using different error metrics and residual plots.

Despite the limitations and complex non-linear nature of varying weather conditions, the obtained results show good accuracy of forecast and are interesting for further research. These findings help to demonstrate the applicability of machine learning in forecasting corona losses in the electrical network which are neglected in most of the case studies. The study demonstrate potential for short term—a few hours to a few days ahead—forecasting of corona losses. Such a short-term projection of corona loss can be beneficial for improved operation of the electrical network.

To determine the best performance of the machine learning regression model, an investigation of the following four aspects was carried out—using data with different time resolution, applying seasonal models, combining the corona losses of two transmission lines, and reduced sets of weather parameters. The first case demonstrated that interpolating weather data can be used to increase the number of corona loss data points used in the training of models, which also improves the accuracy of the models. The second case showed that using a combination of seasonal models instead of one model can increase the accuracy of forecasts. The third case involving more than one transmission line did not clearly prove significant accuracy improvements, but showed potential for future research. The last case of feature reduction illustrated the possibility of using less input data and the resulting effects on accuracy.

The primary conclusion from the findings is that corona

loss forecasting with sufficient accuracy can be implemented in practice using machine learning. The best case performance was an error in the range of ± 0.5 kW/km for 98% of the test data points and a root mean squared error of 0.16 kW/km. This method can potentially be expanded into a comprehensive in-process monitoring system using short-term weather forecasting as the input. Future work could improve the accuracy of forecasting with more advanced machine learning techniques and possibly look into the interdependence of sequential time points, e.g. by applying time series methods. The applicability and accuracy of the method on larger areas using measurement data from multiple transmission lines should be studied further.

ACKNOWLEDGMENTS

Information on the transmission line and respective PMU data was provided by Estonian TSO, Elering. Weather data was received from the Estonian Environmental Agency.

REFERENCES

[1] F. W. Peek, "The law of corona and the dielectric strength of air," *Proceedings of the American Institute of Electrical Engineers*, vol. 30, no. 7, pp. 1485–1561, 1911.

[2] J. S. Carroll and B. Cozzens, "Corona loss measurements for the design of transmission lines to operate at voltages between 220 kv. and 330 kv." *Transactions of the American Institute of Electrical Engineers*, vol. 52, no. 1, pp. 55–62, 1933.

[3] F. Pandžić, I. Sudić, T. Capuder, and I. Pavičić, "On the practical aspects of machine learning based active power loss forecasting in transmission networks," *IET Generation, Transmission & Distribution*, vol. 18, no. 14, pp. 2452–2463, 2024.

[4] N. Dalal, M. Mølna, M. Herrem, M. Røen, and O. E. Gundersen, "Day-ahead forecasting of losses in the distribution network," *AI Magazine*, vol. 42, no. 2, pp. 38–49, 2021.

[5] H. Huang, "Line loss prediction of distribution network based on bp neural network," *Scientific Programming*, vol. 2022, no. 1, p. 6105316, 2022.

[6] S. I. Sulakov, "Forecasting hourly corona losses applying statistical approach," in *19th International Symposium on Electrical Apparatus and Technologies (SIELA)*, 2016, pp. 1–4.

[7] M. Rejc and M. Pantos, "Short-term transmission-loss forecast for the slovenian transmission power system based on a fuzzy-logic decision approach," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1511–1521, 2011.

[8] I. Baran, T. Leonida, and D. Sidea, "Using numerical weather forecast to predict power losses on transmission lines," in *2013 4th International Symposium on Electrical and Electronics Engineering (ISEEE)*. IEEE, 2013, pp. 1–8.

[9] J. Sahlin, R. Eriksson, M. T. Ali, and M. Ghandhari, "Transmission line loss prediction based on linear regression and exchange flow modelling," in *2017 IEEE Manchester PowerTech*. IEEE, 2017, pp. 1–6.

[10] A. Söderlind, "Day-ahead grid loss forecasting: A study of linear and non-linear models when modelling electrical grid losses," 2022.

[11] I. Sudić, M. Mesar, B. Franc, T. Capuder, T. Ivanković, K. Pavić, and I. Pavić, "Short-term transmission system losses forecast based on supervised machine learning," in *2020 International Conference on Smart Systems and Technologies (SST)*. IEEE, 2020, pp. 199–204.

[12] J. Tulensalo, J. Seppänen, and A. Ilin, "An LSTM model for power grid loss prediction," *Electric Power Systems Research*, vol. 189, p. 106823, 2020.

[13] C. Ding, Y. Zhou, Q. Ding, and Z. Wang, "Loss prediction of ultrahigh voltage transmission lines based on EEMD–LSTM–SVR algorithm," *Frontiers in Energy Research*, vol. 10, 2022.

[14] M. Asprou and E. Kyriakides, "Identification and estimation of erroneous transmission line parameters using PMU measurements," *IEEE Transactions on Power Delivery*, vol. 32, no. 6, pp. 2510–2519, 2017.

[15] K. Tuttelberg and J. Kilter, "Estimation of transmission loss components from phasor measurements," *International Journal of Electrical Power & Energy Systems*, vol. 98, pp. 62 – 71, 2018.

[16] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power & Energy Systems*, vol. 155, p. 109537, 2024.

[17] F. J. Sollerkvist, A. Maxwell, K. Rouden, and T. M. Ohnstad, "Evaluation, verification and operational supervision of corona losses in Sweden," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 1210–1217, 2007.

[18] P. K. Gupta, K. Tuttelberg, and J. Kilter, "The weather impact on corona losses of 330 kv aging transmission lines," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*. IEEE, 2024, pp. 1–5.

[19] Estonia Environmental Agency, "https://www.ilmateenistus.ee/."

[20] P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023.

[21] K. Tuttelberg, M. Löper, and J. Kilter, "Correcting systematic errors in corona losses measured with phasor measurement units," *IEEE Transactions on Power Delivery*, vol. 34, no. 6, pp. 2275–2277, 2019.

[22] M. Sobhani, A. Campbell, S. Sangamwar, C. Li, and T. Hong, "Combining weather stations for electric load forecasting," *Energies*, vol. 12, no. 8, p. 1510, 2019.

[23] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[24] E. T. Iorkyase, C. Tachtatzis, I. A. Glover, P. Lazaridis, D. Upton, B. Saeed, and R. C. Atkinson, "Improving RF-based partial discharge localization via machine learning ensemble method," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1478–1489, 2019.

[25] M. Venkateswara Reddy and R. Sodhi, "A modified s-transform and random forests-based power quality assessment framework," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 78–89, 2018.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

Curriculum Vitae

1. Personal data

Name	Pradeep Kumar Gupta
E-mail	pguptakumar12@gmail.com

2. Education

2021–2026	Tallinn University of Technology, School of Engineering, Electrical Power Engineering and Mechatronics, PhD studies
2017–2019	National Institute of Technology Nagaland, Electrical Power Engineering, Department of Electrical and Electronics Engineering, M.Tech
2012–2016	National Institute of Technology Sikkim, Department of Electrical and Electronics Engineering, B.Tech

3. Professional employment

2019–2020 Grinntech Motors Services Pvt. Ltd, Software Engineer (Training Period Intern)

4. Defended theses

- 2019, Human Learning Based Optimization Algorithms to Solve Economic Load Dispatch Problem in Power Systems, M.Tech, supervisor Prof. Dushmanta Kumar Das, National Institute of Technology Nagaland, Department of Electrical and Electronics Engineering.
- 2016, Comparative Power Quality Analysis of Right and Left Shunt UPQC Topologies, supervisor Prof. Pradeep Kumar, National Institute of Technology Sikkim, Department of Electrical and Electronics Engineering.

5. Scientific work related to the thesis
Journal Papers

1. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023
2. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power and Energy Systems*, vol. 155, p. 109537, 2024
3. P. K. Gupta, J.-R. R. Ruiz, P. Casals-Torrens, J. T. i Garvín, K. Tuttelberg, and J. Kilter, "Sensitivity analysis of corona discharges measuring instruments using different electrodes and high voltage supplies," *IEEE Sensors Letters*, vol. 8, no. 12, 2024
4. P. K. Gupta, J.-R. R. Ruiz, K. Tuttelberg, and J. Kilter, "Quantification of corona discharge intensity applied to sphere-plane configurations," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2025
5. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Forecasting corona losses on high voltage transmission lines using machine learning," *IEEE Transactions on Power Delivery*, vol. 40, no. 5, pp. 2696–2705, 2025

Conference presentations

1. P. K. Gupta, K. Tuttelberg, and J. Kilter, "The weather impact on corona losses of 330 kV aging transmission lines," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, *Dubrovnik*, *Croatia*, pp. 1–5, 2024

6. Scientific work not related to the thesis

1. P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "An adaptive class topper optimization for scheduling thermal generation with cubic fuel cost function," in *2025 15th International Conference on Power, Energy, and Electrical Engineering (CPEEE)*, pp. 195–199, IEEE, 2025

2. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Adaptive chaotic class topper optimization to solve non-convex chped problems in power systems," in *2024 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 1–6, IEEE, 2024

3. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Class topper optimization approach to solve non-convex combined heat and power economic dispatch problems in power systems," in *2024 IEEE International Conference on Power and Energy (PECon)*, pp. 179–184, IEEE, 2024

4. P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "Ameliorated quantum class topper optimization to solve economic load dispatch problems," in *2024 3rd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)*, pp. 1–5, IEEE, 2024

Elulookirjeldus

1. Isikuandmed

Nimi	Pradeep Kumar Gupta
E-mail	pguptakumar12@gmail.com

2. Haridus

2021–2026	Tallinna Tehnikaülikool, Inseneriteaduskond, Elektroenergeetika ja mehhatroonika, doktoriõpe
2017–2019	Nagalandi Riiklik Tehnoloogiainstituut, Elektroenergeetika, Elektro- ja Elektroonikatehnika osakond, M.Tech
2012–2016	Sikkimi Riiklik Tehnoloogiainstituut, Elektri- ja Elektroonikatehnika osakond, B.Tech

3. Professionaalne töökoht

2019–2020	Grinntech Motors Services Pvt. Ltd, tarkvaraarendaja (praktikant)

4. Kaitstud lõputööd

- 2019, Inimese õppimisel põhinevad optimeerimisalgoritmid energiasüsteemide majandusliku koormuse jaotamise probleemi lahendamiseks, M.Tech, juhendaja Prof. Dushmanta Kumar Das, Nagalandi Riiklik Tehnoloogiainstituut, Elektroenergeetika, Elektro-ja Elektroonikatehnika osakond.
- 2016, Parempoolse ja vasakpoolse šundi UPQC topoloogiate võrdlev võimsuskvaliteedi analüüs, B.Tech, juhendaja Pradeep Kumar, Sikkimi Riiklik Tehnoloogiainstituut, Elektri-ja Elektroonikatehnika osakond.

5. Väitekirjaga seotud teadustöö
Ajakirjaartiklid

1. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Assessment of corona loss performance on aging transmission lines using PMU measurements," *Energy Reports*, vol. 9, pp. 215–219, 2023
2. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Weather dependency of corona losses on 330 kV overhead transmission lines," *International Journal of Electrical Power and Energy Systems*, vol. 155, p. 109537, 2024
3. P. K. Gupta, J.-R. R. Ruiz, P. Casals-Torrens, J. T. i Garvín, K. Tuttelberg, and J. Kilter, "Sensitivity analysis of corona discharges measuring instruments using different electrodes and high voltage supplies," *IEEE Sensors Letters*, vol. 8, no. 12, 2024
4. P. K. Gupta, J.-R. R. Ruiz, K. Tuttelberg, and J. Kilter, "Quantification of corona discharge intensity applied to sphere-plane configurations," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2025
5. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Forecasting corona losses on high voltage transmission lines using machine learning," *IEEE Transactions on Power Delivery*, vol. 40, no. 5, pp. 2696–2705, 2025

Konverentsi ettekanded

1. P. K. Gupta, K. Tuttelberg, and J. Kilter, "The weather impact on corona losses of 330 kV aging transmission lines," in *2024 IEEE PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, *Dubrovnik*, *Croatia*, pp. 1–5, 2024

6. Teaduslik töö, mis ei ole väitekirjaga seotud

1. P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "An adaptive class topper optimization for scheduling thermal generation with cubic fuel cost function," in *2025 15th International Conference on Power, Energy, and Electrical Engineering (CPEEE)*, pp. 195–199, IEEE, 2025

2. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Adaptive chaotic class topper optimization to solve non-convex chped problems in power systems," in *2024 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 1–6, IEEE, 2024

3. P. K. Gupta, K. Tuttelberg, and J. Kilter, "Class topper optimization approach to solve non-convex combined heat and power economic dispatch problems in power systems," in *2024 IEEE International Conference on Power and Energy (PECon)*, pp. 179–184, IEEE, 2024

4. P. K. Gupta, A. K. Gupta, K. Tuttelberg, and J. Kilter, "Ameliorated quantum class topper optimization to solve economic load dispatch problems," in *2024 3rd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)*, pp. 1–5, IEEE, 2024

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-484-1 (PDF)