DOCTORAL THESIS

Data-driven Asset Management and Condition Assessment Methodology for Transmission Overhead Lines

Henri Manninen

TALLINNA TEHNIKAÜLIKOOL
TALLINN UNIVERSITY OF TECHNOLOGY
TALLINN 2022

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
4/2022

Data-driven Asset Management and Condition Assessment Methodology for Transmission Overhead Lines

HENRI MANNINEN

TALTECH 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 20 December 2021

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

Co-supervisor: Dr. Mart Landsberg,
Head of Grid Maintenance Department,
Elering AS
Tallinn, Estonia

Opponents: Dr. Armando Rodrigo Mor,
Universitat Politècnica de València,
Valencia, Spain

Dr. Kostas Kopsidas,
The University of Manchester,
Manchester, The United Kingdom

Defence of the thesis: 08 March 2022, Tallinn

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.

Henri Manninen

signature

Financial support was provided by the Estonian Ministry of Education and Research (The Kristjan Jaak Scholarship for study period abroad) and Estonian Students Fund in USA, inc. (Scholarship for Full-time Students). The data is provided by Elering AS.

elering
CONNECTING ENERGIES

ISSN 2585-6898 (publication)
ISBN 978-9949-83-789-2 (publication)
ISSN 2585-6901 (PDF)
ISBN 978-9949-83-791-5 (PDF)
Printed by Koopia Niini & Rauam

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
4/2022

Ülekandevõrgu õhuliinide andmepõhine varahalduse ja seisundi hindamise metoodika

HENRI MANNINEN

TALTECH KIRJASTUS

Contents

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 H. Manninen, J. Kilter, and M. Landsberg, “Advanced condition monitoring method for high voltage overhead lines based on visual inspection,” in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, Aug 2018

II H. Manninen, J. Kilter, and M. Landsberg, “Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices,” in *2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, pp. 1–6, June 2019

III C. J. Ramlal, H. Manninen, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, “Toward automated utility pole condition monitoring: A deep learning approach,” in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 255–259, 2020

IV H. Manninen, J. Kilter, and M. Landsberg, “Health index prediction of overhead transmission lines: A machine learning approach,” *IEEE Transactions on Power Delivery*, vol. 37, no. 1, pp. 50–58, 2022

V H. Manninen, C. J. Ramlal, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, “Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques,” *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106726, 2021

VI H. Manninen, J. Kilter, and M. Landsberg, “A holistic risk-based maintenance methodology for transmission overhead lines using tower specific health indices and value of loss load,” *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107767, 2022

Author's Contributions to the Publications

I In I, Henri Manninen is the main author. He had a major role in writing the manuscript and presented it at IEEE PES General Meeting 2018. He proposed the idea, developed the methodology, analysed the results and wrote the manuscript.

II In II, Henri Manninen is the main author. He had a major role in writing the manuscript and presented it at 2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM). He proposed the idea, developed the methodology, analysed the results and wrote the manuscript.

III In III, Henri Manninen participated in writing the manuscript and was dealing with all issues related to high voltage equipment including defect detection.

IV In IV, Henri Manninen is the main author. He proposed the idea, developed the methodology, analysed the results and had a major role in writing the manuscript.

V In V, Henri Manninen is the main author. He proposed the idea, developed the methodology, prepared the figures and had a major role in writing the manuscript.

VI In VI, Henri Manninen is the main author. He proposed the idea, developed the methodology, analysed the results and had a major role in writing the manuscript.

Abbreviations

A-GPS	Assisted Ground Positioning System
BiFPN	Bi-directional Feature Pyramid Network
CBM	Condition-Based Maintenance
CEER	Council of European Energy Regulators
CENS	Cost of Energy Not Supplied
CIGRE	International Council on Large Electric Systems (in French: Conseil International des Grands Réseaux Électriques)
CM	Corrective Maintenance
CoF	Consequences of Failure
CSPNet	Cross Stage Partial Network
CSV	Comma-separated values
DSO	Distribution System Operator
DT	Decision Trees
ENTSO-E	European Network of Transmission System Operators for Electricity
EPSG	European Petroleum Survey Group
FMEA	Failure Mode and Effects Analysis
FN	False Negative
FP	False Positive
GB	Gradient Boosting
GIS	Geographical Information System
GPS	The Global Positioning System
h	Hour
HI	Health Index
IEC	The International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IoU	Intersection-over-Union
ISO	International Organization for Standardization
k	Kilo
km	Kilometer
KMe	Kaplan Meier
KNN	K-Nearest Neighbor
kV	Kilovolt
LAS	LiDAR Aerial Survey
lat	Latitude coordinate
LCM	Lifecycle Management
LiDAR	Light Detection and Ranging
lon	Longitude coordinate
LR	Logistic Regression
M	Million
m	Meter
MP	Megapixel
MW	Megawatt
MWh	Megawatt in hour
NACE	The Statistical Classification of Economic Activities in the European Community
NA	Nelson Aalen

NBC	Naive-Bayes Classifier
NMS	Non-Max Suppression
NN	Multi-Layer Perceptron from Neural Networks
OHL	Overhead Line
PANet	The Prototype Alignment Network
PAS	Publicly Available Specification
pH	A scale used to specify the acidity or basicity
PM	Predictive Maintenance
PoF	Probability of Failure
PR	Precision-Recall
r	Radius
RBM	Risk-Based Maintenance
R-CNN	Region-based Convolutional Neural Network
RF	Random Forest
ROC	Receiver Operator Characteristic
RoI	Region of Interest
ROW	Right-Of-Way
RPN	Region Proposal Network
SGDM	Stochastic Gradient with Momentum
SGDR	Stochastic Gradient Descent with Warm Restarts
SMOTE	Synthetic Minority Oversampling TEchnique
SSD	The Single Shot Detector
SVM	Support Vector Machine
TBM	Time-Based Maintenance
TP	True Positive
TSO	Transmission System Operator
UAV	Unmanned Aerial Vehicle
UID	Unique Identificator
VOLL	Value of Lost Load
YOLOv2	The You Only Look Once version 2
YOLOv5	The You Only Look Once version 5

Introduction

> ... all models are approximations. Essentially, all models are wrong, but some are useful. However, the approximate nature of the model must always be borne in mind....
>
> *George G. E. Box* [1]

Motivation and background

The reliability of electricity transmission has improved significantly since the beginning of the electrical era in the 19th century and most of today's economic sectors rely on the quality of supply where even short interruptions can cause major economic losses. There have been severe economic consequences caused by blackouts in transmission systems where the costs associated with them have reached billions. The North American blackout in August 2003 affected nearly 50 million people with a total economic cost of eight to ten billion dollars [2]. In the same year there were two consecutive blackouts in Europe. The first started on September 23 in Denmark resulting nearly four million customers without electricity. Another blackout started on September 28 in Italy which left most of the country without electricity. Both blackouts started with a regular failure in the grid but cascaded into blackouts due to malfunctions of the relay-protection systems. A large amount of investments and additional training has been done to prevent similar blackouts such as described in [3] but those situations vividly present how a single transmission system element can cause a catastrophic failure of the system in general.

Transmission systems in Europe have mostly been built after World War II with an expected lifetime of around 60 years. As a large amount of assets are reaching their life expectancy, transmission system operators (TSOs) will be facing a significant wave of asset replacements in the next decade. According to the ENTSO-E [4] it is expected that European TSOs will have to invest around 53 billion euros before 2030 to maintain the current level of security of supply. It is noticeable that nearly 80% of the expenses will be for the refurbishment of overhead lines (OHLs). This is leading TSOs to develop cost-effective methodologies for asset management decision-making. The old age of transmission assets does not always refer to critical asset condition and more sophisticated methodologies should be used to increase their cost-effectiveness. In some cases there are assets that are reaching the end of their projected lifetime, but do not exhibit significant degradation. Many TSOs are moving towards a life cycle management (LCM) system of OHLs to minimize long-term operating costs and to maximize the useful lifetime of assets by using more advanced condition monitoring methods that would allow them to move to sophisticated approaches such as condition- or risk-based maintenance. The key factor in effective asset management relies in precise investments in assets that are in the most critical condition or affect the network reliability the most. In addition to condition assessment of technically sophisticated assets, traditional consumption and production patterns have also changed noticeably since their installation due to the growth of renewable production units in electrical systems. Therefore, it is also essential to take into account the economic consequences associated with possible failures.

The most difficult and also the most critical tasks in improved asset-management decision-making are seemingly the simplest ones. The determination of actual assets technical condition and the risks associated with the possible failure of individual assets is a sophisticated process due to the peculiarities of transmission OHLs. A transmission OHL is a complex group of assets consisting of a large amount of individual components that may span hundreds of kilometers and cross harsh terrain. That makes the manual inspection of OHLs' components relatively costly compared to the cost of a single component. The key element in an unambiguous and cost-effective asset-management decision-making methodology is the determination of the optimal moment to replace or repair the asset just before the failure occurs. For that, the technical condition of OHLs must be monitored to prevent failures from happening, but also to avoid investments in assets with good technical condition.

Hypothesis

The thesis is structured to analyze the following hypotheses and to overcome the most common problems among TSOs regarding asset management of transmission OHLs.

- Classical time-based maintenance (TBM) is not cost effective for OHLs and can be improved significantly.
- Age is an inaccurate parameter for assets technical condition assessment.
- The unambiguity of traditional visual inspection results can be improved without additional economic costs.
- It is possible to predict the Health Index (HI) of assets with high accuracy without any additional measurements.
- It is possible to automate visual inspections of OHLs cost effectively using deep learning techniques.
- Object detection enables to detect defects from images of the same quality as the visual inspections.

Main objectives and tasks of the thesis

The main purpose of this thesis is to develop a data-driven asset management decision-making methodology for transmission OHLs to improve the efficiency of current approaches where assets are usually replaced once they reach the end of their expected lifetime. Traditionally, transmission OHLs are assets with a lack of condition information due to their high reliability. Transmission OHLs can reach hundreds of kilometers in length and cross difficult terrain that makes condition assessment expensive. Therefore, it was quickly concluded that a reliable condition assessment methodology of OHLs is required to achieve the objectives. This leads to the development of condition assessment methodologies for OHLs that are cost effective and produce reliable results. The majority of this thesis focuses on the general principles of an unambiguous OHL condition assessment with the aim of using state-of-art machine learning and deep learning technologies to achieve cost saving of OHL inspections. The backbone of the proposed methodology is a comprehensive deterioration analysis of OHL and its components where every detectable visual sign matches a certain period of its life stage. That leads to the development of component-specific condition indicators that can be used by asset managers to ease the assessment process and reduce the human factor.

The backbone of the proposed condition assessment methodology enabled to determine the HI of each OHL tower and its component separately, but there was still a significant workload placed on maintenance personnel to visit OHL towers individually. Almost a third of the thesis focuses on the automatic condition assessment of OHLs using object detection approaches to overcome the burden of manual inspections. First, a single and the most critical defect was selected to be detected from close-up images. As it proved to be successful, more advanced methods were developed to acquire detailed information from the images. A methodology to detect all visual indicators from close-up images that can also be detected from manual inspections was developed for reinforced concrete poles and structures. As both of those methods were relying on good quality close-up images, a third and the most sophisticated methodology was proposed to detect all critical defects of OHL towers using super-high-resolution images taken from high altitude produced as by-product of right of way (ROW) inspections. Using object detection approaches enabled to determine the HI of each OHL tower and its component separately, but there was still an issue with gathering condition information about assets which were not assessed previously or where the data was lost. To overcome this concern, a novel asset HI prediction methodology was proposed that allowed to include assets without previous HI data in the decision-making.

A novel risk-based asset management methodology was developed to reflect the realistic determinations of probability of failure (PoF) and consequences of failure (CoF) in the grid. PoF is strongly affected by the technical condition of assets, and therefore, reliable information from condition assessment is used as an input data. CoF determination is principally based on economic consequences that can be represented by using the value of lost load (VOLL). VOLL is determined individually for each asset in the grid to reflect the realistic situation and to provide transparency on a single asset level. The proposed methodology is developed to explain investments and maintenance budget allocation by determining the most critical assets in the grid and comparing their risk against the cost of renewal. Mathematical optimisation is used to further increase the efficiency of the methodology by determining the optimal combination of assets to focus on in terms of limited budget.

Contribution of the thesis and dissemination

Theoretical novelty of the work

- A novel HI determination methodology is proposed for transmission OHLs based on advanced visual inspections with mobile applications using predefined visual indicators.

- New condition indicators are developed to achieve unambiguous results and decrease the subjectivity of traditional visual inspections done by foot patrols.

- Machine learning models are developed to predict the HI of missing assets and model the aging of assets.

- Object detection models based on deep learning are used in OHLs condition assessment process to enhance the efficiency of OHL inspections by automatically detecting defects from images of OHL towers.

- PoF of assets is determined for each voltage level and HI class separately by using survival analysis.

- VOLL together with outage combination and estimated duration determination is used to determine the consequences of failure.

- A holistic decision-making model is created to support risk-based decisions.
- Results are mathematically optimised in terms of limited budget.

Practical originality of the work
This thesis proposes solutions for the following widespread issues among TSOs:

- Traditional visual inspections using foot patrols are usually subjective due to the human factor and a new unambiguous methodology is proposed.
- In addition to labor-intensive visual inspections two novel and cost-effective approaches are proposed.
- An HI prediction model is created to predict OHL tower HI without additional measurements.
- HI determination using object detection models is used to detect condition indicators from close-up and fly-by photos taken during OHL inspections.
- PoF of a single tower is calculated according to its technical condition.
- Realistic assessment of CoF.
- Risk is calculated on comparable parameters.
- The most critical assets in the grid are determined.
- The proposed asset management decision-making methodology outperformed all widespread approaches.

Thesis outline

This thesis is divided into four main chapters to give a complete overview of the proposed methodology from the data collection to the decision-making. Chapter 1 focuses on the asset management decision-making using risk determination. Chapter 2 focuses on the condition assessment of transmission OHLs using visual inspection and HI prediction for assets without condition information. Chapter 3 proposes three approaches for the automatic HI determination of OHLs using object detection techniques and Chapter 4 presents a case study supporting the use of real data from the actual transmission network. The thesis is concluded in the final part with recommendations for further works.

1 Asset Management of Transmission Overhead Lines

Asset management of transmission OHLs is a sophisticated task as there are currently no common approaches to assess the timeframe and urgency of the replacements. Various methodologies are used among TSOs. Traditionally, the most widespread approach among TSOs is to replace assets once they reach the end of their expected lifetime. It is easy to implement, but may lead to overinvestments as the expected lifetime of assets is usually selected with considerable safety margins to prevent failures. In addition to economic inefficiency, using traditional maintenance approaches would generate a significant wave of replacements. This could be dispersed or postponed with more advanced decision-making methodologies as many of them were constructed in a small timeframe after World War II.

This chapter focuses on the framework of an asset-management decision-making methodology that enables to increase the cost-efficiency and determines the most critical elements in the grid by using a sophisticated, but transparent approach. The methodology is based on comprehensive PoF and CoF determination using asset HI and VOLL. Condition assessment of OHLs is explained in detail in Chapter 2 and Chapter 3, where different approaches are proposed to acquire accurate results cost effectively. The main framework of this chapter is based on the publication **VI** and the VOLL determination is based on the publication **II**.

1.1 General Overview

1.1.1 Asset Management of Transmission Overhead Lines

The simplest asset management decision-making principle in electricity systems is presented in Fig. 1. It is impossible to reach the ideal quality of supply even in terms of unlimited costs as random failures tend to happen. An increase in the quality of supply causes higher costs due to excessive maintenance and lowering the quality of supply causes an increase in failure-related costs. A simplified optimum can be drawn where costs to maintain certain level of quality of supply and costs associated with failures are balanced. However, determining that exact point in a real system is an extremely sophisticated task as there is a large number of individual assets in the grid that can fail and cause an outage to customers.

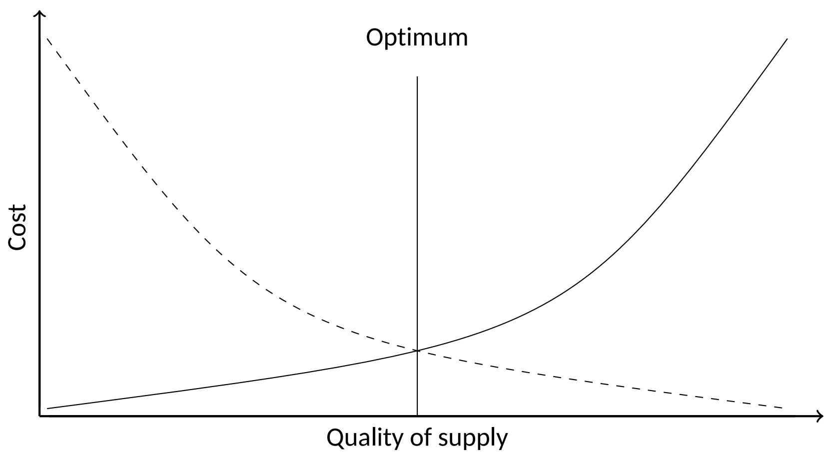


Figure 1: The simplified optimization principle where the solid line is presenting costs related to maintenance and the dashed line shows costs related to failures. [5]

The general principles of asset management are described in standard series ISO 55000 [6–8] and PAS 55 [9, 10]. They give generalized guidelines to implement an effective asset management in the companies, but both lack details, especially in condition assessment and decision-making methodologies. PAS 55 focuses on the asset management process of physical assets and is developed for the oil and petrol industries where ISO 55000 is a further development of PAS 55 for all asset types. For optimal decision-making it is essential to accurately determine the actual need for maintenance and possible consequences after a failure occurs. To overcome this issue, it is essential to develop methodologies that take into account the real condition of assets using precise condition assessment of OHLs. There has been an increase in the condition assessment methodologies for transmission assets in the last few decades, but the majority of these methods focus on substation equipment as they are the most expensive assets in the grid. Some examples of these methodologies include power transformers [11–13], large electrical machines [14], cables [15] and circuit breakers [16] where condition assessment is done to determine the actual technical condition and, therefore, to prevent failures.

1.1.2 Lifecycle Management

Transmission OHLs are complex assets to maintain as they cover large distances, consist of large number of individual towers, thousands of kilometers of conductors and have a lifespan of more than 60 years. OHLs are not just individual towers or wires that require attention to ensure the safety and reliability of OHLs throughout their lifecycle. LCM is used to optimize the total cost of maintenance, investments and condition assessment costs throughout the complete period of in-use OHLs. Fig. 2 gives a brief overview of OHL LCM works that can be divided into three main component classes. The first class consists of everything related to the ROW such as vegetation management or safety margins to structures or crossings with other infrastructures. The second class is related to line objects such as conductors and grounding wires. The third consists of individual support structures such as towers.

In the past decade it has become more widespread to use light detection and ranging (LiDAR) [18] technology to acquire point-clouds of OHLs due to the high accuracy and relatively low cost per inspection kilometer. Point-clouds consist of up to billions of precise geographical coordinates (X,Y,Z) where each of these points represents a single point in 3D space with corresponding categories that distinguish assets, ground, buildings, vegetation, roads, etc. It is recommended to use a LiDAR Aerial Survey (LAS) [19] format as it is an industry standard file format defined by the American Society of Photogrammetry and Remote Sensing that includes a system of point classification. Using aerial vehicles to collect point-clouds with LiDAR enables, after data processing, to acquire the precise distance from one object to another to cover almost all ROW LCM activities. Condition assessment for LCM of ROW mainly consists of vegetation analysis and detection of structures in the ROW. The ROW by itself does not affect the technical condition of OHLs but it must be maintained to eliminate outage because of vegetation flash-overs. Preventing vegetation-related outages also improves the safety of the OHLs as flash-overs may cause fires or even dangerous step-voltage for humans and animals. In addition to vegetation analysis LiDAR data is also used to determine the minimal ground clearances of each span or crossing with infrastructures by modeling conductor thermal behaviour for allowed conductor temperature ranges as explained in [20]. This ensures the required safety margins for agricultural machines, structures in the ROW and crossings with roads by measuring distance to the conductor. Fig. 3 presents an example of vegetation analysis using collected LiDAR point-clouds after data processing where purple dots represent dangerous trees

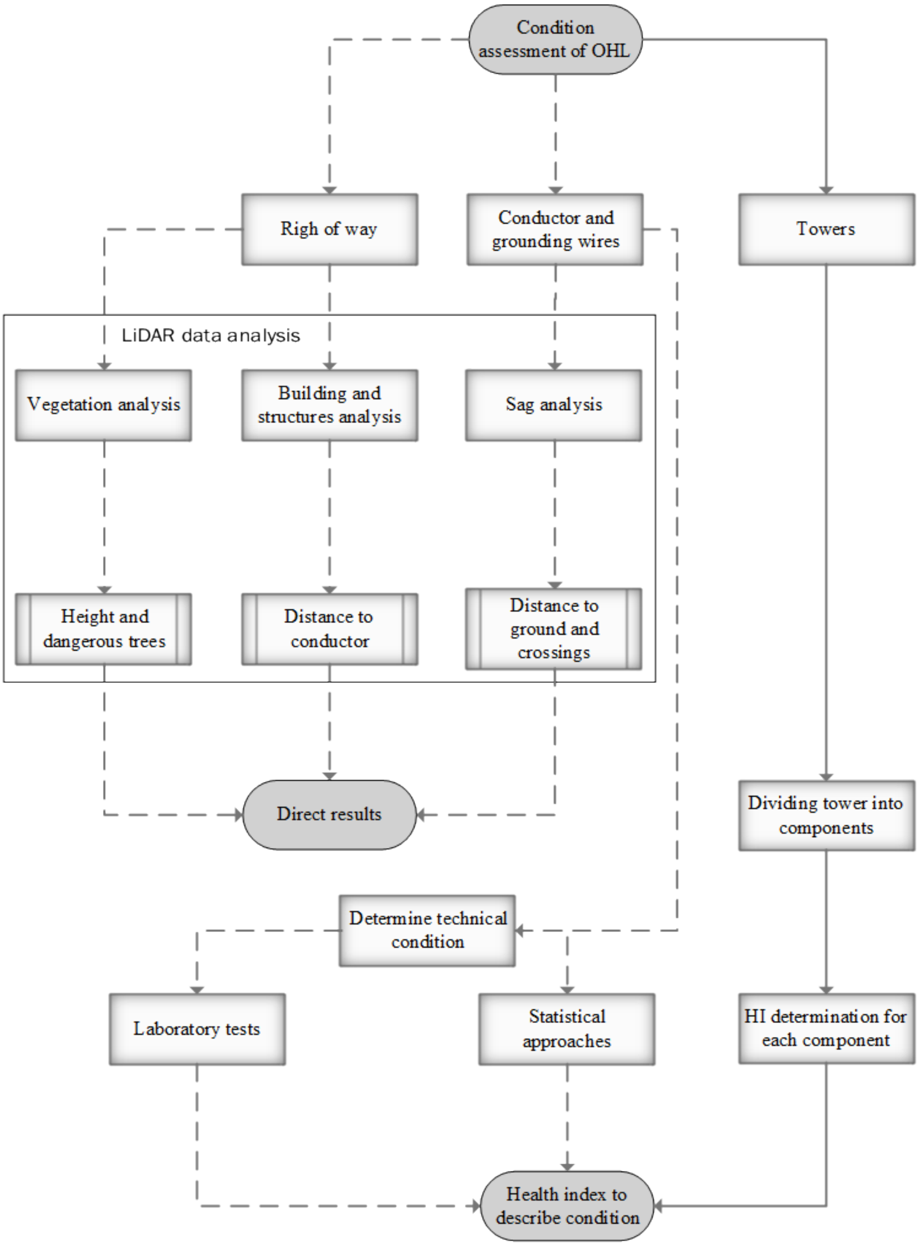


Figure 2: Principle scheme of OHL condition assessment where solid lines present condition assessment process that is covered in the thesis and dotted lines that are not covered in the thesis. Rectangles represent process steps and round shapes the start and end of the process. [17]

and other colors represent the vegetation height classes of polygons. Dangerous trees refers to trees that might cause a flash-over as they are tall enough to reach conductors or danger zones once they fall. Darker green colors represent vegetation up to 2 meters and dark orange vegetation with a height up to 15 meters. This approach is widely used as it allows to determine vegetation height accurately using distance measurements and provides excellent input to vegetation management. Condition assessment of conductors, grounding wires and towers is described in Section 2.1.2.

Figure 3: An example of OHLs ROW vegetation analysis using processed LiDAR data. Black dots represents dangerous trees and green to orange vegetation height classes of polygons from low to high.

1.1.3 Maintenance strategies

The most widespread approaches of maintenance are corrective maintenance (CM), and preventive maintenance (PM) including TBM, CBM, and RBM. A comprehensive comparison of different maintenance strategies is conducted in [21]. The most widespread approach among TSOs, due to its simplicity, is TBM where assets are usually replaced once they reach the end of their expected lifetime. It is easy to implement, but may lead to overinvestment and does not provide condition information about assets. The RBM methodology is proposed as the most cost-effective technique, but is also the most sophisticated to implement. Asset management decisions using RBM are usually made on the basis of risk matrices [22] that combine two main factors associated with risk, PoF and CoF, into decisions. PoF represents the technical condition of assets and CoF everything that occurs once the asset fails. The use of risk matrices is widespread among TSOs as it allows to justify investments and enables simplified visualization of PoF and CoF, but the main disadvantage to this approach is the lack of transparency. A thoughtful review of using risk matrices is done in [23] where it is concluded that many companies use risk matrices without ensuring their efficiency on improved decision-making. As decisions are done by using multiple levels of weighting factors on aggregation of assets or results then there is a high probability of losing the transparency of the methodology.

The simplified selection process of an asset specific maintenance strategy is presented in Fig. 4. The selection process starts with the failure mode effect analysis (FMEA) for each asset to determine its criticality and the most crucial failure modes. FMEA was first applied in the aircraft industry in the 1960s and was used to determine all possible failure modes of individual components of the Boeing 747 as the technical complexity of 747s overwhelmed maintenance engineers. This approach provides the opportunity to focus on the most critical failure modes for further analysis and therefore focuses first on failures modes that affect the asset and the grid in general the most.

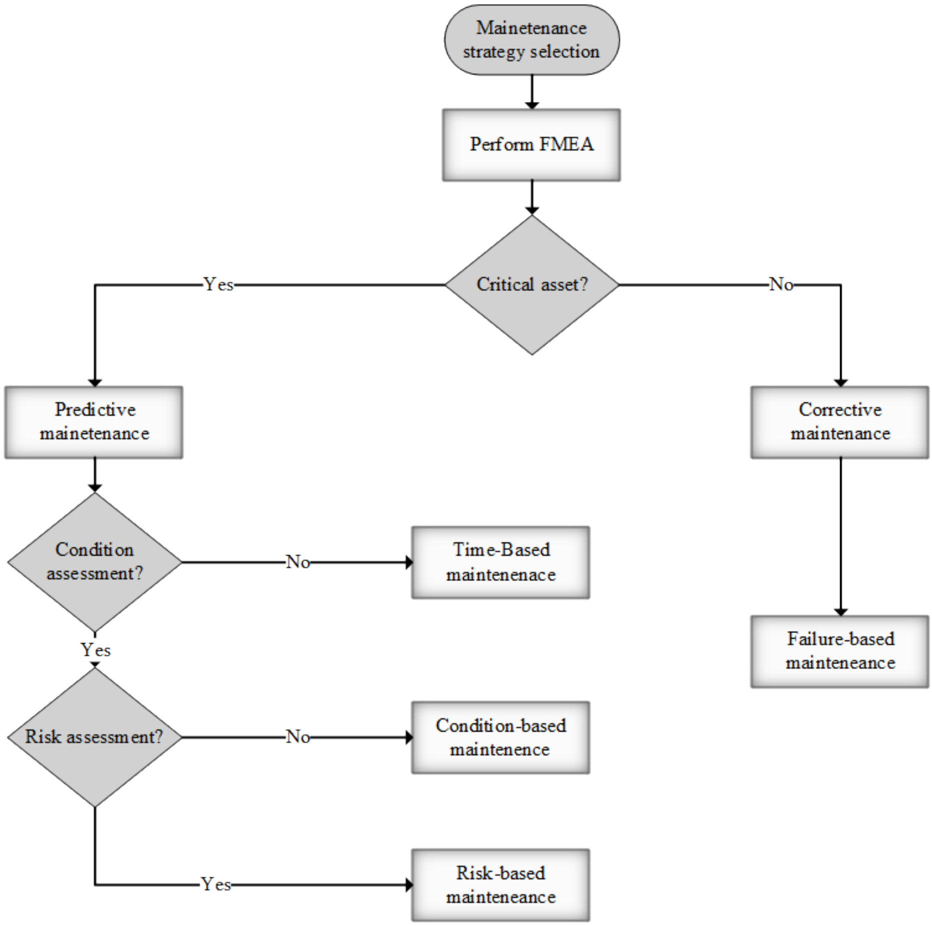


Figure 4: The simplified maintenance strategy selection process. Rectangles represent process steps, rhombuses decision points and round shapes the start and end of the process.

Maintenance strategies can be divided into two main branches, PM and CM, according to mitigation of the impact of operational failures. CM is the simplest approach where assets are replaced as they fail and is mainly used by distribution systems as risks associated with failures are lower than in the transmission system. PM is the most common maintenance strategy among TSOs and it usually focuses on replacing or maintaining assets based on fixed intervals or operations. PM is usually implemented for TSOs as a TBM where assets will be refurbished or replaced after a certain amount of time. CBM is also a part of PM, but it relies on the assessment of the asset's technical condition, and maintenance decisions are done according to the asset's actual condition, not assumptions. RBM is formed on the basis of an asset's importance in the system and assessment of its technical condition. It can be considered as a further development of a CBM strategy where investments are not only performed on assets with a bad technical condition but also those that will have a significant impact on the grid once they fail. Based on [21, 24–31] the RBM methodology is proposed as the most cost-effective for expensive high-voltage equipment and its principles are further used in the thesis.

1.2 Asset Management using a Risk-based Maintenance Framework

The simplest principle for decision-making to determine whether to replace the asset or not is by using (1). To use that approach, a well-explained and transparent risk assessment methodology is required for reliable results. The RBM approach provides an additional parameter for decision-making that enables decisions to be made based on the cost of an asset's calculated risk. That also enables to detect individual assets in the grid that influence the performance of grid the most.

$$Risk > Cost \tag{1}$$

According to [21], risk in a transmission system can be stated in its simplest form by using (2), where PoF is the probability of failure and CoF the consequences of the failure.

$$Risk = PoF \times CoF \tag{2}$$

It is an extremely simple equation, but in order to implement it in decision-making, its usage becomes sophisticated as PoF and CoF values are complicated to determine. That leads to risk assessment which is heavily dependent on the input data where errors in the first steps affect the final result drastically.

A general overview of the proposed asset management decision-making methodology is presented as a flowchart in Fig. 5. It is a data-driven methodology that relies on the PoF and CoF determination where condition assessment and VOLL results are input parameters for risk assessment. As decisions are based on the data, an additional layer to increase the data quality is applied by using an outlier detection model to detect invalid condition assessment information before risk assessment. Once PoF and CoF are calculated for each tower it is possible to move to decision-making. Decision-making uses the risk of each tower as an input and enables mathematical optimisation to achieve the best results according to the cost of maintenance works and limitations such as budget or planned outages.

1.2.1 Risk Assessment

Risk of each tower is calculated based on the PoF and CoF of each asset as presented in Fig. 6. Risk assessment starts with full-scale grid calculations where all possible combinations that will cause an outage or limitation in the transmission grid, for consumers or producers, will be saved. Then VOLL for each affected substation will be calculated based on the consumption profiles in parallel with PoF determination for each tower. Risk calculation for each tower ends with a list of all possible combinations of failures where a single OHL tower participates. All those combinations will be compared and the the combination with the highest risk will be selected as a tower risk.

As a single failure in transmission grid does not usually cause any outages due to the N-1 requirement in grid planning, the number of circuits sharing the same OHL tower must be checked. For example, in some cases there are two circuits connecting the same substation sharing the same individual tower. In the case of a tower failure, there will be an outage in the substation even though the classical N-1 requirement is satisfied. Once there is a single OHL that will cause an outage in the substation, all towers of the OHL will acquire the VOLL of the substation with individual estimated outage duration based on the tower parameters as described in more detail in Section 1.4. In cases where there is more than

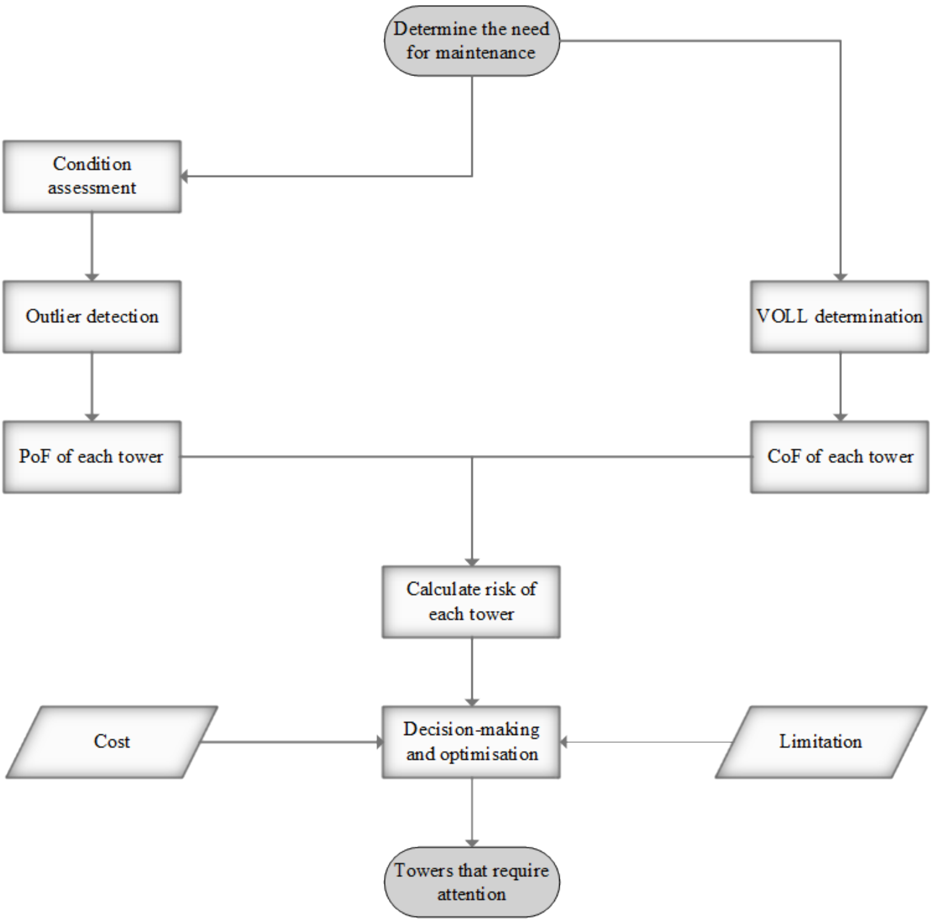


Figure 5: Principle scheme of an asset management decision-making methodology. Rectangles represent process steps, parallelograms data and round shapes the start and end of the process. [32]

one OHL in the outage combination, it is checked whether there are towers that share the circuits of both OHLs. If that requirement is satisfied, all towers that share those OHL circuits are saved with substation VOLL using the individual estimated outage duration for each tower. In cases where more than a single tower involved in the combination of outage, the PoF of each tower will be calculated by using the probabilities for the selected tower and the maximum PoF value of other OHLs in the combination according to (3). VOLL of that tower will be calculated on the basis of the substation VOLL that will be affected by the failure and the estimated outage duration is selected for the tower that was under selection. Due to the large amount of different outage combinations, a single OHL tower can participate in more than one combination and, therefore, the risk of an individual tower in all possible combinations is calculated and saved. The maximum risk among all combinations for each tower is selected for further sections to find the worst-case scenarios for each asset. For the purpose of simplicity, up to two individual towers are used simultaneously to find possible outage combinations. Increasing the number of towers in a single outage combination will increase the computational complexity significantly but will marginally improve the results as probabilities are unpretentious values.

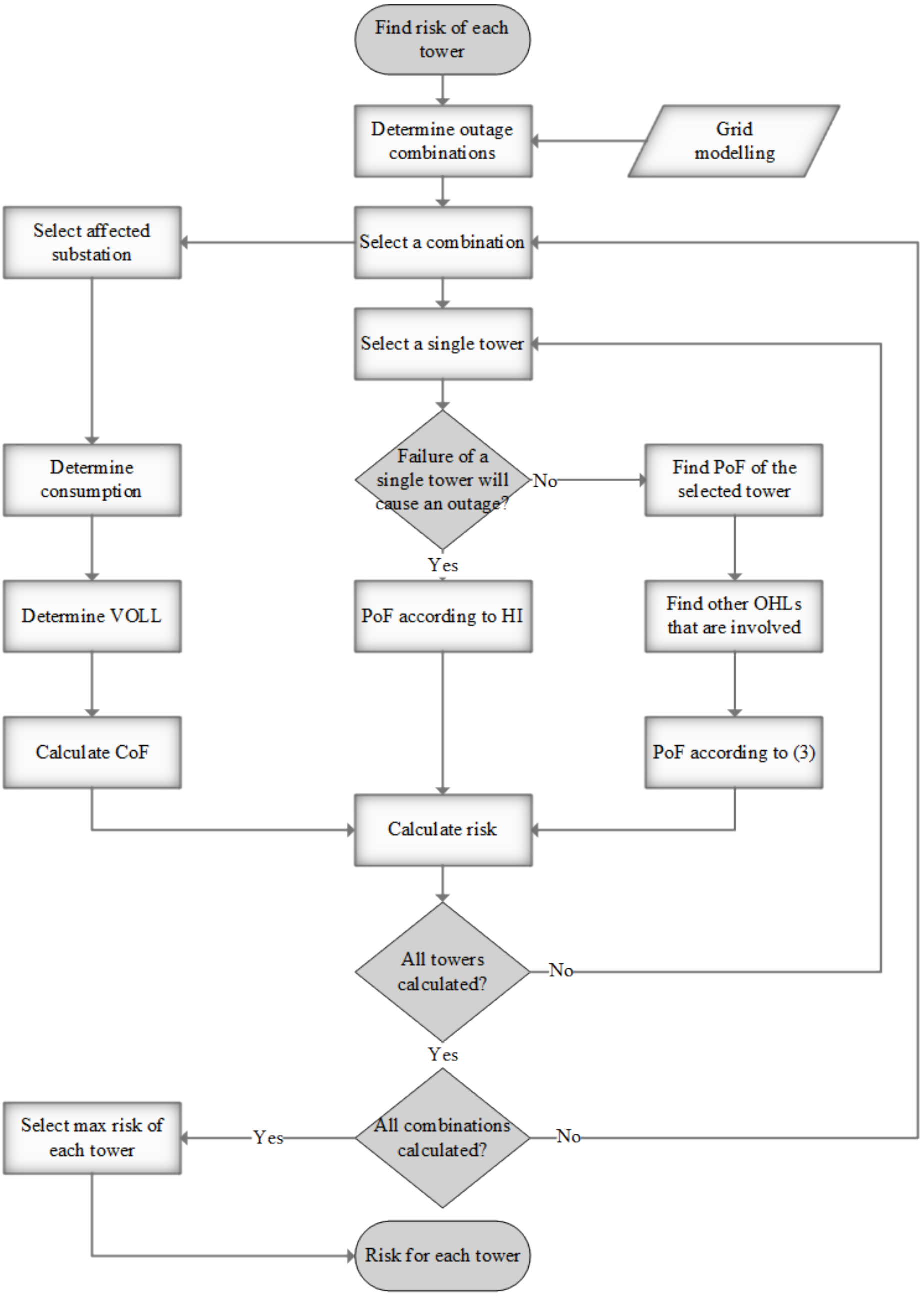


Figure 6: The risk determination flowchart to find risk for each tower in the grid. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process. [32]

$$PoF_{JP} = PoF_{Selected} \times \max_{i=1}^{n}(PoF_{HI_i}) \tag{3}$$

where PoF_{JP} is the joint probability of failure of the combination, $PoF_{Selected}$ is the PoF of the selected tower according to its HI, n is the number of towers in the combination and PoF_{HI_i} is the PoF of each tower in the combination according to the HI.

1.2.2 Increasing Input Data Quality

Even the most sophisticated models may lead to inaccurate results if there are incorrect values in the raw data. To minimize errors in the data, an outlier detection model is used. It is also possible to detect and fix all anomalies manually, but as soon as the amount of data increases, it will become overwhelming. A more effective approach is to use an outlier detection model that is based on unsupervised machine learning algorithms and detects incorrect values in the data automatically, even at big data level. As seen in Fig. 5, the outlier detection model is implemented right after condition assessment to minimize the risk of incorrect values in further steps by detecting suspicious values in the data. As the outlier detection model detects all assets that have the largest deviations in technical features compared to the HI, this does not always mean that the data is incorrect. For example, in some cases there might only be few-year-old towers with mechanical defects from heavy machinery that results in the end-of-life criteria of towers represented with the maximum HI value. That is not correlated with aging, but an outlier detection model highlights those towers as an anomaly in the dataset. Highlighted values can be double checked by experts to determine whether the value is incorrect or just abnormally different. This minimizes the data validation workload of experts by only double-checking a short list of highlighted assets, not the full dataset.

Outlier detection sometimes called anomaly detection algorithms are well described in [33]. Anomaly scores are used to decide whether the data point is an anomaly or not, as illustrated in Fig. 7 where different anomaly score levels are presented around data points. Decisions are usually made based on following rules:

- If anomaly score is close to 1 it indicates an anomaly.
- If anomaly score is close to 0 it indicates a normal data point.
- If anomaly scores for all data points are close to 0.5 it indicates that there are no anomalies in the data.

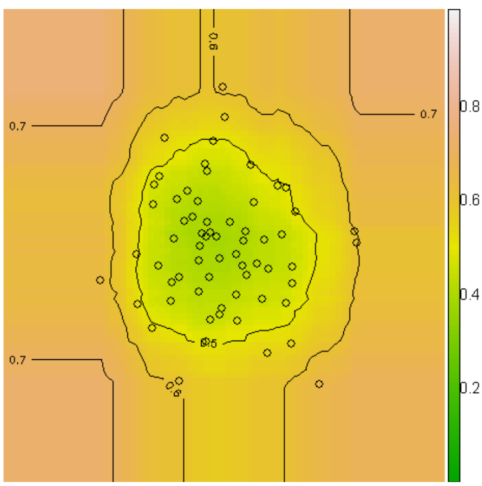


Figure 7: An example of the anomaly score boundaries using outlier detection models. [34]

One of the most accurate and widely used outlier detection algorithms is the Isolation Forest [34], that is built on an ensemble of decision trees for a given dataset. It is explicitly developed to isolate anomalies in the data instead of profiling normal points. It is ideal for high-volume datasets due to the low memory requirement and it works well even when there are no anomalies present in the training set. Fig. 8 presents a simple example of anomaly detection using the unsupervised machine learning algorithm Isolation Forest where straight black lines describe random partitions generated by the model. Anomaly score for Isolation Forest defined is by (4).

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{4}$$

where $h(x)$ is the path length of data point x, $c(n)$ is the average path length of unsuccessful search in a Binary Search Tree and n is the number instances in the dataset.

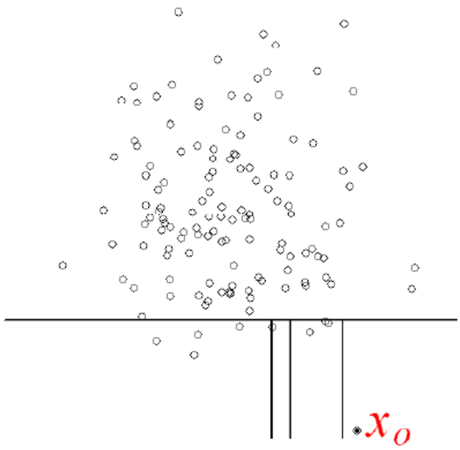


Figure 8: An example of the anomaly point detection (X_0) using Isolation Forest. Straight black lines describe random partitions by the model and circles individual data points. [34]

For successful outline detection, information about all assets that have technical condition information and HI data will be used to train the model and the output is a list of assets that might have data quality issues. An outlier detection model is composed and integrated into the full framework using the scikit-learn [35] toolbox in Python. Technical features to describe OHL towers and their parameters in the thesis are:

- Number of circuits (1 to 4).
- Tower type (suspension or tension).
- Voltage level (110 kV and 330 kV).
- Tower material (reinforced concrete, zinced, painted and uncoated steel lattice tower).
- Manufacturer (6 manufacturers).
- Tower configuration (214 configurations).
- OHL direction changes (angle of deviation).

- Existence of a perch guard (yes or no).
- Age (1 to 67 years).

1.2.3 Optimisation

To maximize the efficiency of the proposed RBM decision-making methodology under budget limitations, mathematical optimisation is used. Budget and replacement optimisation is described as a simple knapsack problem [36] that is mathematically expressed as (5) where $Cost_i$ is the cost of replacement and $Risk_i$ risk of i-th element. The output of decision-making optimisation is a list of assets that require maintenance.

$$maximize \sum_{i=1}^{n} Cost_i * Risk_i <= Budget \tag{5}$$

Solving a small knapsack is relatively simple, as after determining all possible combinations, the best one can be chosen. Solving a large knapsack however, becomes extremely computationally expensive as the number of combinations to find the best solution grows exponentially. To overcome this issue, different approaches are used to solve complex knapsacks using more effective methods than simple brute-force, where all possible combinations are tested. In [36] different approaches are proposed to solve complex knapsack problems using more effective methods than simple brute-force when all possible combinations are tested to find the best solution. Dynamic programming from [37] is used to solve the knapsack problem in the thesis. Linear programming [30] can be used to choose further optimisation algorithms that are computationally effective.

1.3 Probability of Failure

The framework of PoF determination for each individual tower is presented in Fig. 9. Inputs for PoF determination are HI data determined in Chapter 2, critical replacements and historical failures. Cumulative hazard functions are calculated for each predefined asset category using survival analysis [38]. Different voltage levels are differentiated to acquire more reliable results due to different reliability requirements for 110 kV and 330 kV OHLs as described in OHL standard for Estonian National Normative Aspects [39] where 330 kV OHLs are at reliability level 3 and 110 kV OHLs at reliability level 2. Different reliability levels are used by designers to select weather parameters such as wind and ice load that the OHL has to withstand. A higher level of reliability reflects more durable OHLs.

1.3.1 Historical failures

Failure data is used to describe the historic performance of various assets. TSOs have traditionally used PM where a large number of assets will never reach their end-of-life condition or fail in service. When scaling historical failures to TSO level then it is also essential to take into account not only failures but also critical defects and just-before-failure replacement of assets. For the use of historical data on a comparable basis then standardized failure reports should be used to record asset HI before failure or critical replacement. There are examples of failure reports by CIGRE for OHLs [24] or IEC 62271 [40] for substation switchgear. In addition to a comparison of different failures on a common basis, standardized failure reports also enable asset technical condition to be registered at all. As HI presents the technical condition of the asset, it also enables asset-specific PoF for each HI class to be found by using survival analysis and composing cumulative hazard functions.

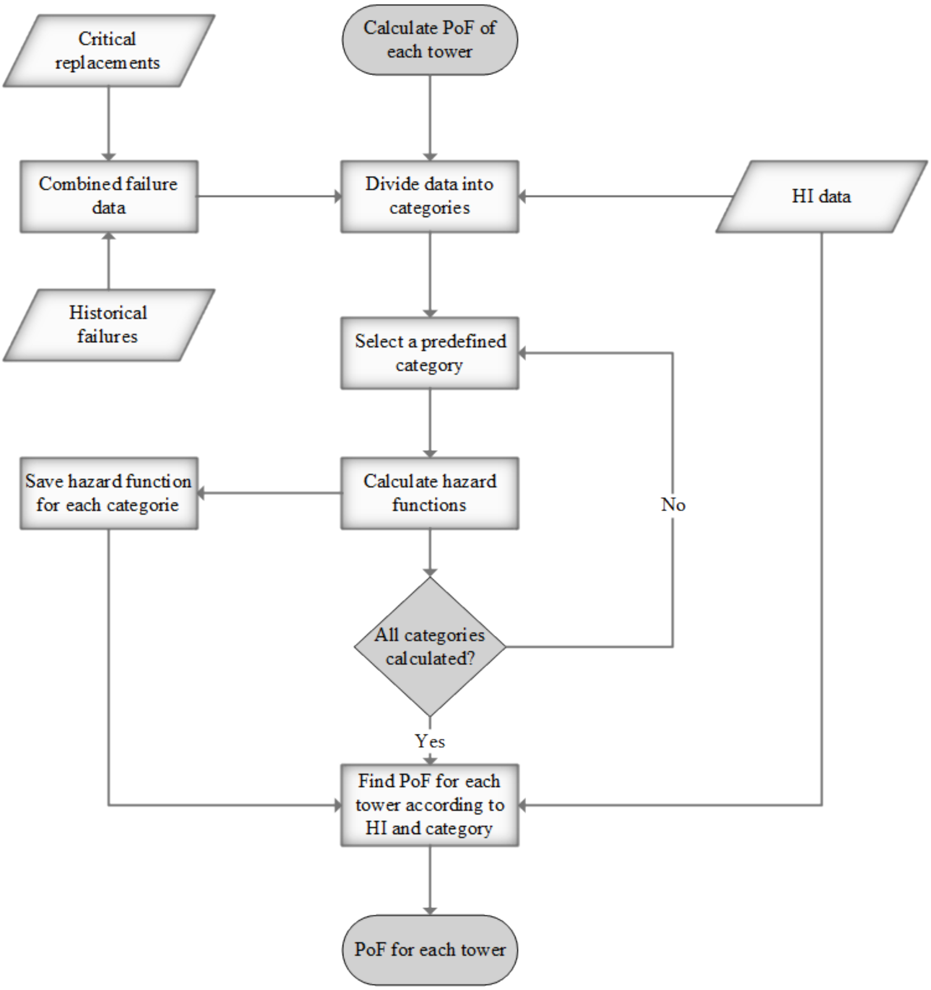


Figure 9: The flowchart of PoF determination for each tower from Fig. 6. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process. [32]

1.3.2 Cumulative Hazard Functions

Cumulative hazard functions are used to determine the mortality of individuals for certain time intervals based on historical data as described in detail in [38]. Using a cumulative hazard function provides the opportunity to visually examine distributional model assumptions for reliability data of assets and to have a similar interpretation as probability plots. This thesis uses survival analysis to determine cumulative hazard functions for selected asset categories over HI values based on their historical performance and registered technical condition. Failures are rare in TSO's grid due to preventive maintenance approaches and, therefore, using parametric functions may lead to over- or underestimations using wrongly selected parameters. Non-parametric methods are used instead of parametric or semi-parametric functions because the assumptions made on parametric functions are not always justified by the data and determining parameters on a small dataset may distort the results. In cases where there is a sufficient amount of failure data such as in a distribution system's (DSO) grid, Weibull, exponential model, gamma or the Cox proportional hazard models can be more detailed and accurate compared to non-parametric models. Both Kaplan Meier (KMe) and Nelson Aalen estimators (NAe) can be used to estimate the cumulative hazard function, but according to [41] NAe is slightly superior in terms of increasing failure rates. Cumulative hazard functions are composed for each selected asset category. The NAe is used to directly estimate the cumulative hazard function given by (6). An example of composed cumulative hazard functions for 110 kV and 330 kV OHLs is presented in Fig. 58 in Chapter 4.

$$H_{NA}(t) = \sum_{t_i \leq t} \frac{d_i}{n_i} = \sum h_{NA,i}(t_i) \tag{6}$$

where n_i corresponds to the number of towers present at time t_i and d_i is the number of failures at time t_i.

In order to perform a survival analysis, it is essential to record the time to event. But sometimes this is not possible because of the limitations in the real world and, therefore, only partial information about time to event is available. In that case, censoring is used and in the context of asset management, it might happen for several reasons. For instance, it may be that asset was replaced before there was a failure or the asset HI is not known. Such situations require data to be excluded about those assets for more accurate results. The methodology used here uses one-year inspection data about OHL towers as an input combined with historical failure data. As there is only data about those assets that still exist and age and HI are known, censoring is not used for that data, but for further years it is recommended to use censoring when assets are taken out of service.

1.4 Consequences of Failure

The consequences of the failure in the methodology are based on the VOLL and replacement costs of towers after failure. These can be expressed financially and are measurable. In the literature there are various other parameters that are commonly used in CoF determination, but they are ignored as they are based on non-comparable values and each utility has its own risk mitigation strategies. Such parameters can be CoF in safety, environmental issues or loss of reputation. For example, if one TSO includes loss of human life in CoF with great value and the other TSO does not, then the results are significantly different. These additional parameters could be included in further studies to increase the accuracy of the methodology, but as the main aim is to present a complete methodology to combine HI and VOLL for decision-making then they are ignored here. CoF of an outage can be calculated according to (7), but only direct cost and VOLL parameters are used in the thesis due to the simplifications and therefore $\sum CoF_i = 0$.

$$CoF = VOLL + CoF_{Direct} + \sum CoF_i \tag{7}$$

where CoF_{Direct} are costs associated directly with the outage such as cost of repair works, CoF_i is consequences related to safety, environment, and publicity for the company or even political pressure.

1.4.1 Value of Lost Load

Value of lost load [42] is used as a monetary indicator to express the cost associated with an interruption of electricity supply. VOLL is determined through multi-step approaches that usually start with dividing consumers into predefined categories and assessing cost of energy not supplied (CENS) for each customer sector separately. VOLL can be calculated using (8), where $CENS(t)_i$ is cost of energy not supplied for the sector i at estimated outage duration t.

$$VOLL_{substation} = \frac{\sum CENS(t)_i \times load_i}{\sum load_i} \times t \tag{8}$$

VOLL reflects the total cost of electricity outage for a single substation based on the price of the consumer-specific energy units, consumption, and estimated duration of the possible outage. It should be noted that it is impossible to predict failure occurrences with an exact timeframe and therefore average consumption of substations is used in this thesis. VOLL determination also covers the selection of consumer categories, finding CENS values for each category and assumption of the estimated duration of the outage.

Substation CENS calculations are usually performed based on multi-step approaches that are discussed more in [26, 43–47]. There are several studies about CENS determination where a comprehensive comparison is done in [26, 42]. It states that there is a large gap between the cost of two most widely used approaches, macroeconomic and willingness-to-pay. It is also concluded that by using different methods, economic environments or consumer categories it is possible to receive up to tens of times different CENS results for the same customer sectors. Simple analytical methods such as the macroeconomic approach are easy to implement, but produce the most inaccurate results due to the lack of accurate data. Detailed customer survey methods such as willingness-to-pay require considerable work to implement, but produce more reliable results as they describe each consumer separately. The main disadvantage of the willingness-to-pay method is its high cost due to the customer surveys required. CENS is time dependent and can be expressed as functions illustrated in Fig. 10. As seen in the figure it is thus essential to define the estimated outage duration as accurately as possible as it affects CENS significantly. A simplified method selection for the substation CENS determination process is illustrated in Fig. 11. As presented in Fig. 11 the method selection is mostly influenced by input data availability.

1.4.2 Classification of Consumers

It is not practical to involve every single consumer one by one into CENS determination as each customer has specific consumption, dependence of supply and financial status. Therefore, it is reasonable to classify consumers into groups that are compiled on similar basis and are comparable. To do this, it is recommended to use standardized international classification to enable the comparison of among various countries. For example, CEER has developed a guideline document [49] for CENS calculations where they recommend using NACE Rev.2 [47] classification to group consumers into categories. There is also the possibility to develop alternative groupings, depending on the country-specific factors affecting the input data or chosen methodology, but it is recommended to use an international approach for comparable results. The theory behind NACE groups and sub-categories is explained in more detail in [50].

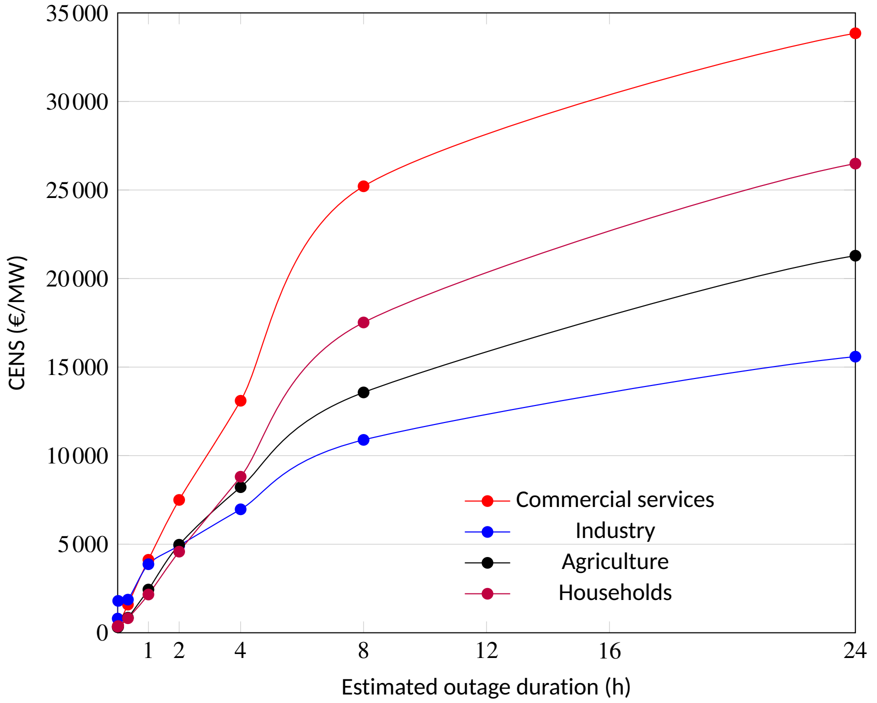


Figure 10: Cost of energy not supplied according to outage duration. Red line represents commercial services, blue industry, black agriculture and purple households. The data is from [48].

- Households
- Commercial services (without infrastructure)
- Public services (without infrastructure)
- Industry (without large customers)
- Large customers
- Infrastructure

1.4.3 Estimated Outage Duration

It is important to determine the estimated outage duration precisely as it affects VOLL results on a large scale. For an example, assuming an eight hour outage duration as a replacement time instead of 24 hours results in significant differences in VOLL values and, therefore, in risk assessment. To overcome the issue, a sophisticated methodology is proposed instead of using average values for all towers. For OHLs, the estimated outage duration is usually determined by the type of the failure and the complexity and repair time of assets. This is especially the case for transmission OHLs that cover large distances in remote areas. Outage duration is also strongly affected by the geographical location as there are areas that are impossible to reach due to the large distance from the nearest roads, but there are also towers that are next to roads. Estimated outage duration can be modified by using (9) as some towers may be located on difficult terrain where it is significantly more challenging to replace a tower. It must be noted that $C_{Terrain}$ is an

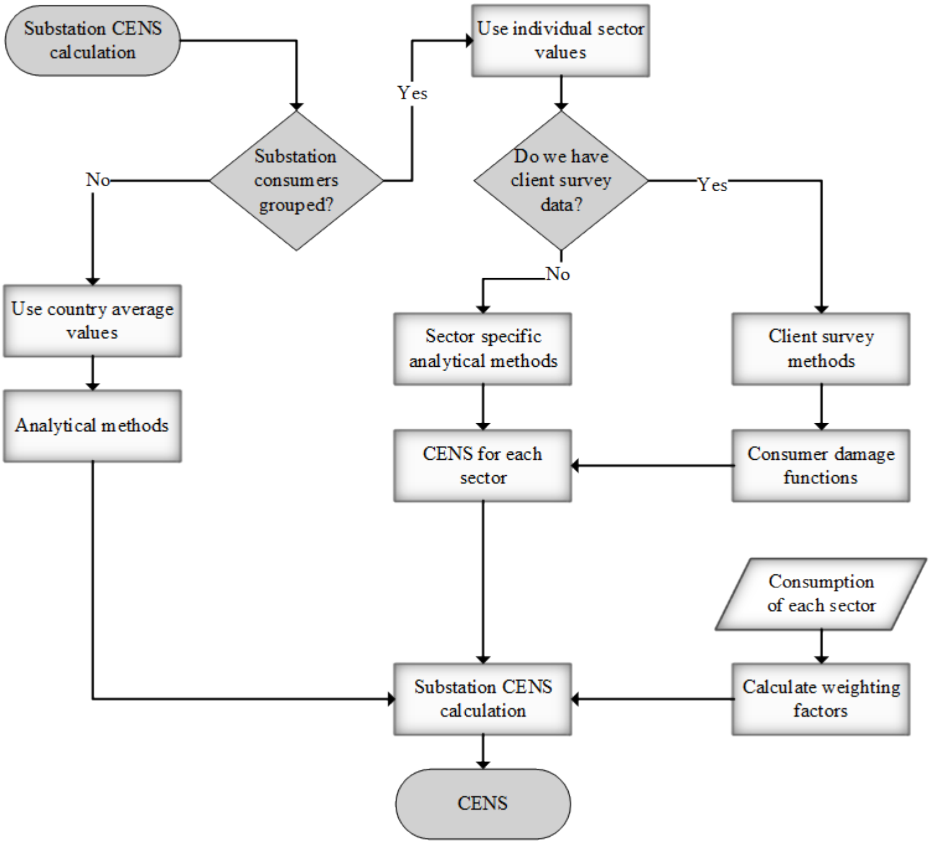


Figure 11: The flowchart of substation CENS determination according to the input data. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process.

empirical constant that should be selected according to a country's geography. But as Estonia is relatively flat, is not influenced by the terrain type as strongly as countries that are mountainous. Estonia also has a high infrastructure coverage all around the country and, therefore, the distance from the nearest road in combination with tower type is used as an indicator of possible outage duration, where $C_{Terrain}$ = 1.

$$T_{outage} = C_{Terrain} \times T_{estimated} \tag{9}$$

where $C_{Terrain}$ is constant that takes into the account terrain type and $T_{estimated}$ the estimated outage duration according to Table 2.

Table 2 presents estimated repair times of OHL tower failures that are developed empirically on the best practice from the Estonian TSO. It must be noted that estimated outage duration times are heavily influenced by the TSO, which defines those times due to country-specific geographical features and availability of repair personnel or spare parts. Table 2 is based on the knowledge where the full tower failure is expected as a worst-case scenario. In addition to the distance from the nearest road, different voltage levels and materials are also distinguished as they influence the complexity of repair works.

The acquisition of a minimal distance from an OHL tower to the nearest road is solved through the use of the tower's geographical location and road information from Open-

Table 2: Estimated outage duration principles for OHL towers based on the tower type, voltage level and distance from the nearest road. [32]

Distance (m)	**Time to replace (h)**			
	110 kV		**330 kV**	
	Steel	**Concrete**	**Steel**	**Concrete**
<100	12	8	16	12
101 - 1 000	24	12	24	24
1 001 - 10 000	36	24	48	36
>10 000	72	72	72	72

StreetMap [51]. For calculation simplicity, four different proximity zones according to Table 2 around OHL towers are used to determine the individual distance from the nearest road to the tower. These proximity zones are generated based on the best practice and empirical values from the Estonian grid and can be selected based on country-specific features. The first proximity zone has a radius of 100 m around the tower, the second has 1 000 m, the third 10 000 m and the fourth 10 000 m or more. To determine a single proximity zone for each tower, the following steps are used:

1. Select a single OHL tower with its geographical location.
2. Make a query with a proximity zone from Table 2 around the tower to acquire road data.
3. Check if there is a road for vehicles in the zone.
4. If no then select larger radius for the zone and make a new query.
5. If there is a road in the zone then return road ID with distance to the nearest road.

The methodology is explained in Fig. 12, where an OHL (marked with dashed line) with two selected towers (*Tower X* and *Tower Y*) and two roads (*Road1* and *Road2*) is presented. It must be noted that, two towers are selected from the OHL for the simplified example and the span between two towers is not 20 000 m. Proximity zones are marked with colours that represent the following: red - <100 m, orange - 100 m to 1 000 m, yellow - 1 001 m to 10 000 m. The *Tower X*, *Road*1 enters into that tower's orange proximity zones and *Road2* is not the in aforementioned zones. For *Tower Y* there is *Road2* in the yellow proximity zone and *Road1* does not enter into the aforementioned zones. That simplified example shows that it is possible to determine distance according to proximity zones where for the *Tower X* the distance is 1 000 m and for *Tower Y* 10 000 meters.

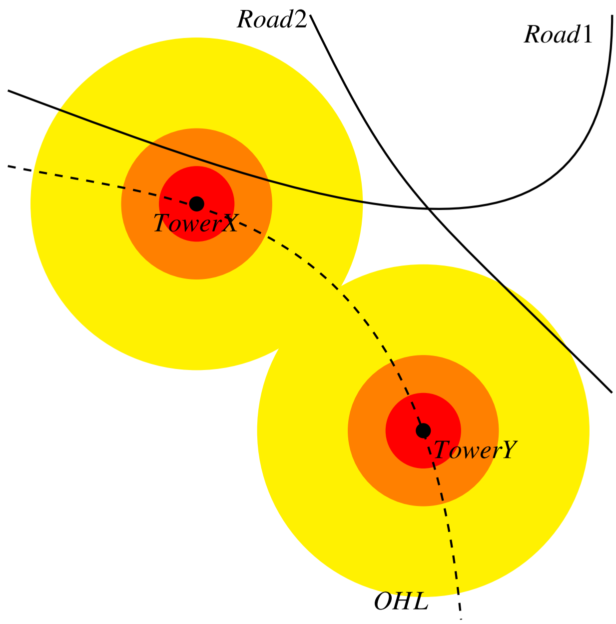


Figure 12: The simplified example of the OHL tower's distance determination from the nearest roads using predefined proximity zones. The OHL is marked with dashed line with two individual towers TowerX and TowerY. Red represents a proximity zone with a radius of 100 m, orange 1 000 m and yellow a zone with a radius of 10 000 m. Two roads, Road1 and Road2, are colliding proximity zones. The radii of the proximity zones is not in a linear scale to provide a clearer example.

2 Condition Assessment of Transmission Overhead Lines

Condition assessment of high-voltage equipment is a challenging task due to the technical complexity of individual assets and even more challenging for the OHLs as they cover large areas with limited physical access. They also consist of tens of thousand of individual towers and components with only a few measurable or comparable parameters. Sophisticated condition assessment with physical measurements is not as widespread for transmission OHLs as it is for substation equipment because OHLs are considered a priori to have high reliability and a long expected lifespan. Condition assessment of OHLs still predominantly relies on manual visual inspections that are costly in such a geographically expansive system.

This chapter presents novel condition assessment methodologies to increase the efficiency of visual inspections and to predict HI of assets that do not have any condition information. Automatic condition assessment using image recognition methodologies is discussed separately in Chapter 3. Condition assessment using advanced visual inspections is based on publication **I** and HI prediction using machine learning methods is based on publication **IV**.

2.1 General Overview

2.1.1 Condition Assessment

Condition assessment of transmission OHLs is usually performed using periodic visual inspection as described in [52] by foot patrols that provide information about the technical condition of OHLs in the grid. Foot patrols physically visit each tower in the grid and record all detected defects. The downside of using foot patrols is that they usually only note down critical defects and do not focus on the condition assessment of assets. Also, there is a large workload involved in assessing each tower individually in a TSO's grid and a human factor that causes individuals to see the same thing differently based on the previous experience of the assessor. In addition to traditional visual inspection, there are several technical publications about OHL condition assessment such as [24, 25, 53–61] which focus only on the determination of critical defects or failures. Others, [62, 63], have developed specific, predetermined criteria lists for each component of OHL to determine the technical condition of assets by using the HI. A range of studies [64–68] describe the condition assessment of wooden poles. It must be noted that wooden poles are not used in the Estonian transmission grid, but they provide a brief overview of condition monitoring mechanisms using invasive and noninvasive methods of structural health monitoring of OHL towers.

Visual inspection is only the first indication to detect critical defects and the technical condition of OHLs as it is usually carried out by ground personnel that do not have expert knowledge in the field of OHL structural integrity and aging phenomena. For a more precise condition assessment, multiple level inspections should be carried out as done in [69]. A potential decision process to determine the actual condition of an OHL tower is divided into four levels, which starts with the traditional visual inspection, moves on to mobile and laboratory measurements and ends with full-scale tests. All levels differ from each other in the level of detail and the cost that increased rapidly. However, in that four-level approach, further levels after visual inspection are rarely used due to the high cost of testing. As the visual inspection is the first level of inspection and it is the most cost-effective method compared to other methods, visual inspections should give reliable and unambiguous results before further steps. There are also statistical approaches to transmission OHLs' condition assessment using bathtub curve [70] or statistical approaches as done in [71]. The bathtub curve was developed for United Airlines in 1978 [70] as a decision-making indicator by using reliability engineering. Paper [71] focuses on statistical trends in data mixed with expert opinions, environmental factors and the weighting of different parameters to determine the HI of OHLs. Both approaches are easy to implement, but they assume that there is a strong correlation between an asset's age and its failure rates, especially bathtub curve. CIGRE has conducted a study on substation switchgear [16] that illustrates a situation whereby the majority of failures for high-voltage assets are random. A study of OHL conductors [72] also concludes that the estimated remaining life of OHL conductors varies significantly when comparing approaches based on age and measurements. That leads to a focus on the condition assessment methodologies whereby the condition of OHLs is received on reliable data not just assumptions.

2.1.2 Health Index

The health index is a simple yet powerful indicator implemented in electricity grids to describe the technical condition of complex assets as a simple value. The HI expresses an asset's technical condition from good to bad and estimates expected remaining fault-free service life in a specified timeframe. There are multiple ways to represent the HI, with the most common being using numbers on a linear scale from 0-100 [73, 74] and 0-5 [24]. As aggregation of assets with their sub-assets loses crucial information for assets while operating with linear values, then logarithmic HI scores with a base of three or ten are also widespread. Color codes according to traffic lights [21] or letters [11] are also used in various studies. Well explained and detailed HI determination options are provided in the technical brochure for power transformers [13], where other indices such as maintenance and refurbishment are also used. However, many papers [11, 70, 73–75] are using HI to express the remaining lifetime of assets using statistical approaches such as Weibull or bathtub [70] curves. Those approaches link HI directly with PoF using asset age and historical failure rates, but should be used extremely carefully as old age does not always equate to bad technical condition. These approaches are widespread because the age of assets is almost always known where other more precise approaches require additional measurements for the HI determination. In some cases asset actual age is modified by various parameters that could be environmental or technical. The output of studies as done in [74–76], is a value that will be compared against the expected lifetime of the asset. Once the modified age is greater than the expected lifetime of the asset it indicates that the asset should be replaced. It can be concluded from the previous studies that condition assessment of transmission OHLs is not widespread and it is usually done by using manual visual inspections or statistical approaches. That leads to the issue where there is subjective or even no information at all about an OHL's technical condition and it is impossible to make precise investment or maintenance decisions due to the lack of data. To overcome this problem, it is essential to determine an OHL's technical condition precisely and cost effectively.

The thesis combines three proposed HI determination methodologies into a single model to provide precise input for PoF determination. Fig. 13 presents an overview of the combined HI determination model. It allows to determine the HI of OHL elements using visual inspections described in Section 2.2 or to use image recognition described in Chapter 3. The selection of the methodology is based on the input data. The output of both approaches is data where HI and technical data of assets is combined. Anomaly detection is used to highlight suspicious values in the data as described in Section 1.2.2. All assets that do not have corresponding HI data will be sent to asset HI prediction model, described in Section 2.3, to acquire HI values. The output of the combined model is a list of towers where all towers in the grid have corresponding HI values.

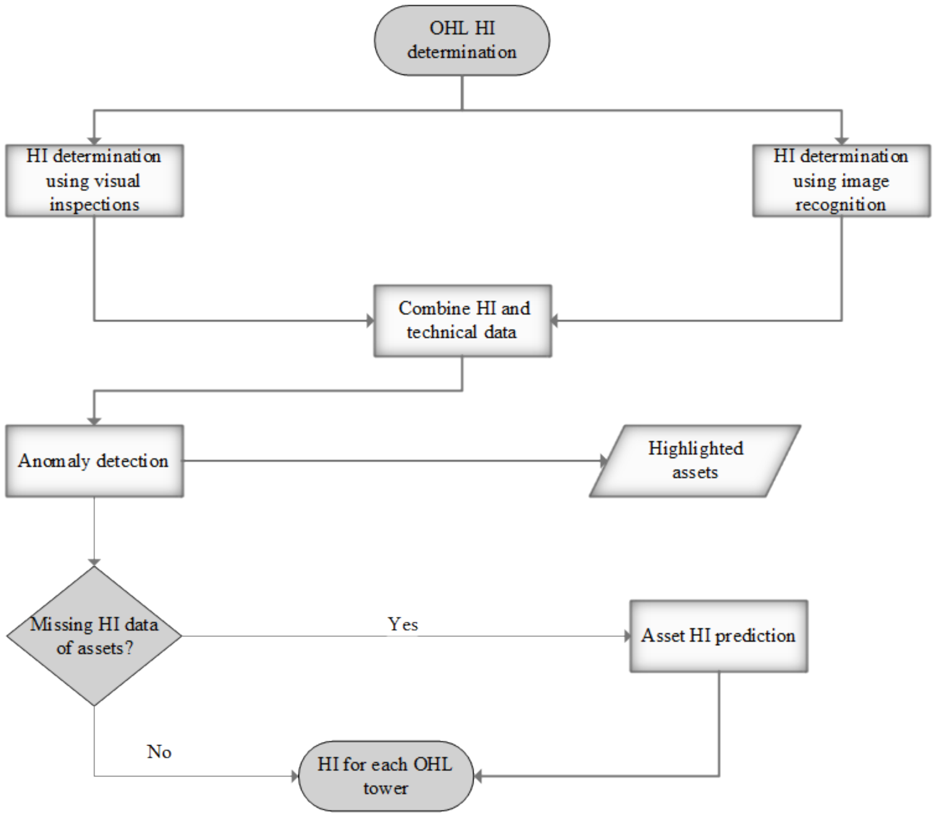


Figure 13: The flowchart of the asset health index determination using the proposed combined model. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process.

2.1.3 Dividing Overhead Line into Components

Each OHL is composed of tens or hundreds of individual towers and hundreds or thousands of kilometers of conductors. This leads to an issue where it is extremely difficult to determine the technical condition of an OHL as it is an aggregation of individual assets. To increase the accuracy of condition assessment of OHLs, each OHL is observed individually on the tower level and each tower is divided into smaller observable parts - components. Components differ from others by construction, material, purpose or cost of replacement. That enables to differentiate each OHL into components that might have different life expectancies or could be replaced separately. That also enables to assess each component individually and therefore, acquire the technical condition of each tower in greater detail compared to more general approaches. It is possible to differentiate components that are tower specific such as foundation, grounding system, crossbar, support, insulation and guy wires and components that are span specific such as grounding wire and conductors. It is also possible to list ROW as an individual OHL component but, as the condition assessment is focused on the technical aspects of OHLs, then it is not done here. The components of OHLs include the following:

- Foundation
- Support
- Crossbar
- Guy-wire
- Insulation
- Conductor
- Grounding wire
- Grounding system

OHLs can be divided into components in other ways, but it must be noted that the division increases the accuracy of condition assessment in general. The downside of the division is that it increases the workload of inspections drastically but using a smart application such as inspections using specially designed tablet applications or object detection from images should decrease the extra workload while providing improved results.

2.2 Condition Assessment Based on Advanced Visual Inspections

2.2.1 The basic principles

Traditional visual inspections of OHLs are time consuming and costly due to the long distance of transmission OHLs. Without specific assessment criteria, it can be subjective due to the fact that the evaluation results from different inspection patrols can vary between regions and assessors on a large scale. There is a strong dependence on results based on the assessor's previous experience in the field that needs to be minimized to increase the accuracy of visual inspections. Therefore, highly specific and unambiguous evaluation criteria are developed for all OHL components that not only describe defects but also give an indication of the technical condition of the selected components. The main idea behind the advanced visual inspections is to describe all possible defects of OHL components in relation to their occurrence in the asset's lifecycle. The backbone of the methodology is the unambiguous determination of an asset's HI using foot patrols to classify the OHL's condition. To do that, all OHLs are divided into components, where visual indicators will be developed for each component separately. The detection of existing defects using a standardized data entry methodology increases the results, as assessors do not have to know the background or physical processes behind the assessment but only need to find predetermined visual indicators. That enables to gather homogeneous results from inspection patrols after just a short training period without the increase of visual inspection costs.

2.2.2 Health Index Determination

The basic principle of the advanced visual inspection method for OHLs is the development of such assessment criteria where every detectable visual sign of a selected component matches a certain period of the same component's life stage. That requires the development of a HI behind every visually identifiable assessment criteria that describes a selected component's technical condition in the range of zero to five. HI values are in the range of zero to five and divided linearly from excellent technical condition to the end of its operational lifetime at HI=5. Nearly 150 different criteria were developed to define the technical condition of all OHL components including towers, foundations, insulators, grounding systems, cross-bars, guy-wires and conductors.

Reinforced-concrete structures such as foundations and tower poles are a vivid example to understand the basic changes in the material's physical properties due to aging. The aging of reinforced concrete is described thoroughly in [77]. In general, reinforced-concrete structures are very durable by construction and their potential lifetime is primarily influenced by the external environment such as the climate or corrosion rate. The main factor in the aging process of reinforced-concrete structures is concrete's surface pH level as the concrete protects steel reinforcements from corrosion. When the pH becomes lower than 8.3 then concrete loses its corrosion protective properties and with sufficient humidity and oxygen in the environment, steel starts to oxidize, i.e. rust. The rust's volume is up to ten [78] times more than the steel's volume resulting in cracks due to the concrete's inner stress. This process will accelerate in time as more humidity and air will reach the reinforcements and eventually the concrete cover will fall off. Once there are cracks and the concrete starts to fall off then it may lead to structural breakage of the pole in severe weather conditions. Of course, mechanical damage from heavy transport or agricultural equipment leads to instant changes of HI from good to poor, but as the HI is designed to show linear decrease of components' exploitation resources then mechanical damages rapidly accelerate the aging process.

Predefined visual indicators of a reinforced concrete pole correlated to HIs are presented in Table 3 and Fig. 14. A similar approach was implemented for steel lattice towers, where the HI was determined based on the presence of mechanical defects, bolt condition, rust level and cross-sectional reduction. The HI values are based on the impact to overall reliability of the pole and, therefore, HI determination is based on the maximum function of all detected defects according to (10).

$$HI = \max \{HI(x) : x = 1 \ldots n\} \tag{10}$$

where $HI(x)$ is a function that looks for the HI value of selected defects in Table 3 and n is the total number of detected defects. For example, if there are cracks and micro longitudinal cracks detected on the asset then according to Table 3, a "crack" corresponds to HI4 and "micro longitudinal cracks" to HI2. By using the maximum function in (10), the overall HI of that asset is four.

Table 3: A list of visual indicators presented in Fig. 14 and their corresponding HI values. [79]

Visual indicator	Health Index
Hole (Fig. 14a)	5
Loss of cross section $> 20\%$ (Fig. 14b)	5
Concrete is falling off (Fig. 14c)	5
Loss of cross section $< 20\%$ (Fig. 14d)	4
Crack (Fig. 14e)	4
Visible reinforcements (Fig. 14f)	3
Micro longitudinal cracks (Fig. 14g)	2
Hair-like cracks (Fig. 14h)	1
Other minor Defects (Fig. 14i)	1
No visible defects	0

The asset HI values are defined as a set of discrete HI categories from zero to five. The value itself is linked to the asset's technical condition with respect to its projected

Figure 14: An example presenting reinforced concrete pole and foundation visual indicators according to their severity as listed in Table 3. [79]

lifetime by using the linear equation (11). The HI determination indicators were developed according to the moment of occurrence in the asset's lifecycle. Therefore, the developed asset HI represents the expected remaining lifetime of asset.

$$L_{expected} = \frac{HI_{max} - HI_i}{HI_{max}} * L_{Projected} \tag{11}$$

where HI_{max} is the maximum value of the HI, HI_i is the HI of the selected asset, and $L_{Projected}$ is the projected lifetime of the asset.

To achieve unambiguous assessment results and effective data management, a specially designed mobile application was developed to support the convenient assessment of OHLs. Fig. 15 presents the user interface of the designed mobile application. It is made as intuitive and simple as possible to ensure homogeneous results from different assessors. An assessment tool is developed on the basis of the proposed methodology and all assessment criteria in eight component groups are present in the application. Additional component "markings" are also present as these do not represent the technical condition of an OHL tower, but is still important to record defects associated with markings using foot patrols. In order to perform visual inspection using the assessment tool inspectors have to select all defects that they identify on the field and the HI determination process is done in the background of the application according to (10). That approach eliminates the human factor in the assessment process as humans tend to give opinions based on the worst defects and ignore not critical defects. If no visual indicators are identified, the component will get the assessment result "no visual defects". It is also possible to add photos of each component with additional free-text comments. Photos taken with the assessment tool are used further in Section 3 for object detection model training.

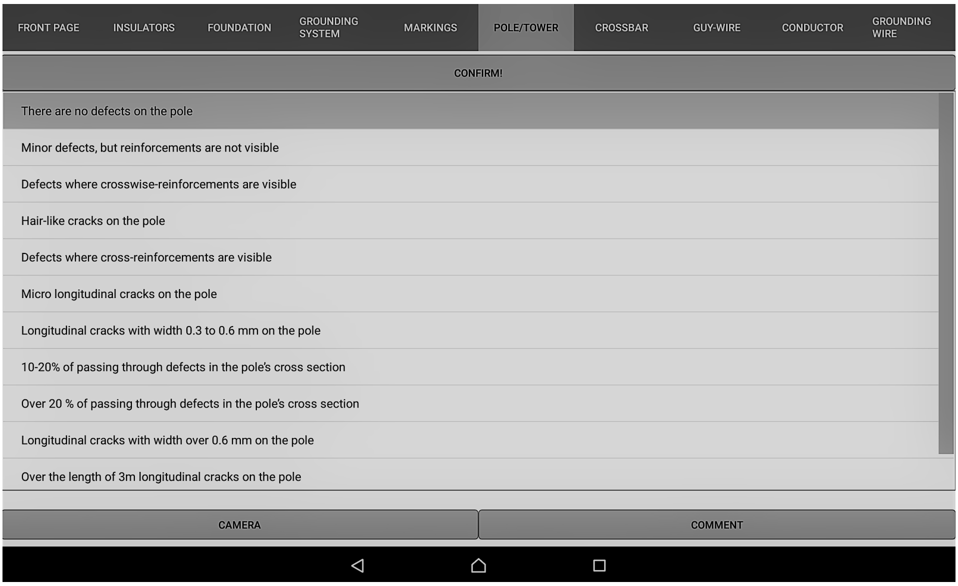


Figure 15: The user interface of the designed application. [80]

2.2.3 Health Index Aggregation

In some cases it is important to compare different OHL towers or OHLs with each other to present the worst elements in the grid. As merely using summed average values of towers will lose important information about each element then this section proposes a HI aggregation methodology that uses exponential values of HIs for clearer comparison. To do that, HIs on a linear scale will first be converted to exponential using (12). It should be noted that it is suggested to use individual component HIs in further asset management decision-making as then there is no loss in the data due to the averaging. Also, using individual component HIs enables to focus on individual asset rather than not on associations of assets.

$$HI_e = e^{\frac{\ln 101}{5} * HI_{max}} - 1, \tag{12}$$

where HI_{max} is the maximum value of the HI.

Table 4 presents an example of HI conversion from linear to exponential scale using (12) when $HI_{max} = 5$.

Table 4: A health index comparison table by using equation (12) between linear and exponential scales.

Linear	Exponential
0	0
1	1.52
2	5.33
3	14.94
4	39.13
5	100

As an OHL tower is composed of different individual components then calculationg the HI of a single tower using the combination of all components' HIs according to (13).

$$HI_{et} = \sum_{n=1}^{6} (HI_{en} * w_n) \tag{13}$$

where n is number of components on a single OHL tower, HI_{en} is HI of n-th component in exponential scale and w_n the weighting factor of n-th component.

Weighting factors used in this thesis are composed based on the best practice from Estonian TSO's according to the proportional replacement cost of that component compared to full renovation. Weighting factors are presented in Table 5.

Table 5: Weighting factors of OHL components according to the proportional replacement cost.

Component	Weighting factor (%)
Foundation	22
Support	22
Crossbar	4
Guy-wire	2
Insulation	10
Conductor	35
Grounding wire	2.5
Grounding system	2.5

The HI expressing a single OHL can be calculated using (14).

$$HI_{OHL} = \frac{\sum_{i=1}^{n_{towers}} HI_{etn}}{n_{towers}} \tag{14}$$

where n_{towers} is total number of towers on the OHL and HI_{etn} is HI of n-th tower.

For an example of why to use exponential scale instead of linear scale is presented as a single OHL with five towers. The HIs of each tower are the following: $HI_1 = 0$, $HI_2 = 0$, $HI_3 = 1$, $HI_4 = 5$, $HI_5 = 5$.

$$HI_{Linear} = \frac{0+0+1+5+5}{5} = 2.2 \tag{15}$$

$$HI_{Exponential} = \frac{0+0+1.51+100+100}{5} = 40.3 \tag{16}$$

While comparing results OHL HI calculated according to (15) and (16), using exponential scale highlights critical towers much more easily than a linear scale. When looking at Table 4, the results from (12) are worse than when the OHL had all towers with HI4 accordingly 40.3 and 39.1. In contrast, on the linear scale, the same OHL had total HI just above two and this reflects the fair condition of all towers. It must be noted that if OHLs consist of a large number of individual towers then peak values or critical towers will be highlighted less. Nevertheless, the use of an exponential scale highlights critical values much more clearly than a linear scale.

2.3 Condition Prediction and Aging Modeling

The proposed methodology for transmission OHL HI prediction is based on supervised machine learning classification algorithms combined with a HI determination framework using predefined visual indicators to predict the technical condition of OHLs on unseen data. The overview of the methodology can be found in Fig. 16. The asset HI prediction starts with the collection of OHL towers' technical features and HI data gathered as in Section 2.2. The technical features of each asset are collected and saved. After that technical data features and HI data are used as training data for the HI prediction model. The data of each OHL tower used in the training process is a combination of collected HI values and technical features as listed in Table 6. Assets that do not have complete list of technical features or HI data will not be included in the training data. This enables the trained prediction model to have already existing and relevant training data and therefore improved overall prediction accuracy by not using random variables. After model training data selection, the missing asset HI prediction and the asset aging behaviour modeling is separated. The output of both individual branches in the methodology is HI prediction results for each selected tower.

The aging behaviour modelling is also used to compare the current situation against future scenarios to detect potential situations where it might be more beneficial to hold or speed up investments. Such situations could be a complete renovation of the OHL instead of changing just a single tower or allocating resources for future investments. If in the near future a majority of assets on the OHL need to be replaced then the proposed model indicates an increase in the budget. That enables a comparison of the current situation against future scenarios to see what assets need additional resources in the near future. It also enables the replacement wave of assets to be flattened as the replacement time can be extended for longer periods to minimize peaks in annual investments. By nature, HI prediction and aging modeling are similar, however the main difference is the step where the selected asset age parameter is modified by *n*-years for the aging modeling while all other features remain unchanged. As long as modified time intervals are not unrealistic and there is a sufficient amount of training samples that support the modified age parameter in the grid, this approach enables to predict asset HI based on the best knowledge and a similar performance of assets. It is difficult to determine the realistic time interval for aging modeling, but once the limits are exceeded, the model becomes inaccurate as there is a lack of data to support accurate predictions. For example, if there are towers up to of 50 years age in the grid, then it is possible to model the selected OHL in those limits. The model does not have information about performance of towers that are older than 50 years and predictions will be done based on unrealistic data.

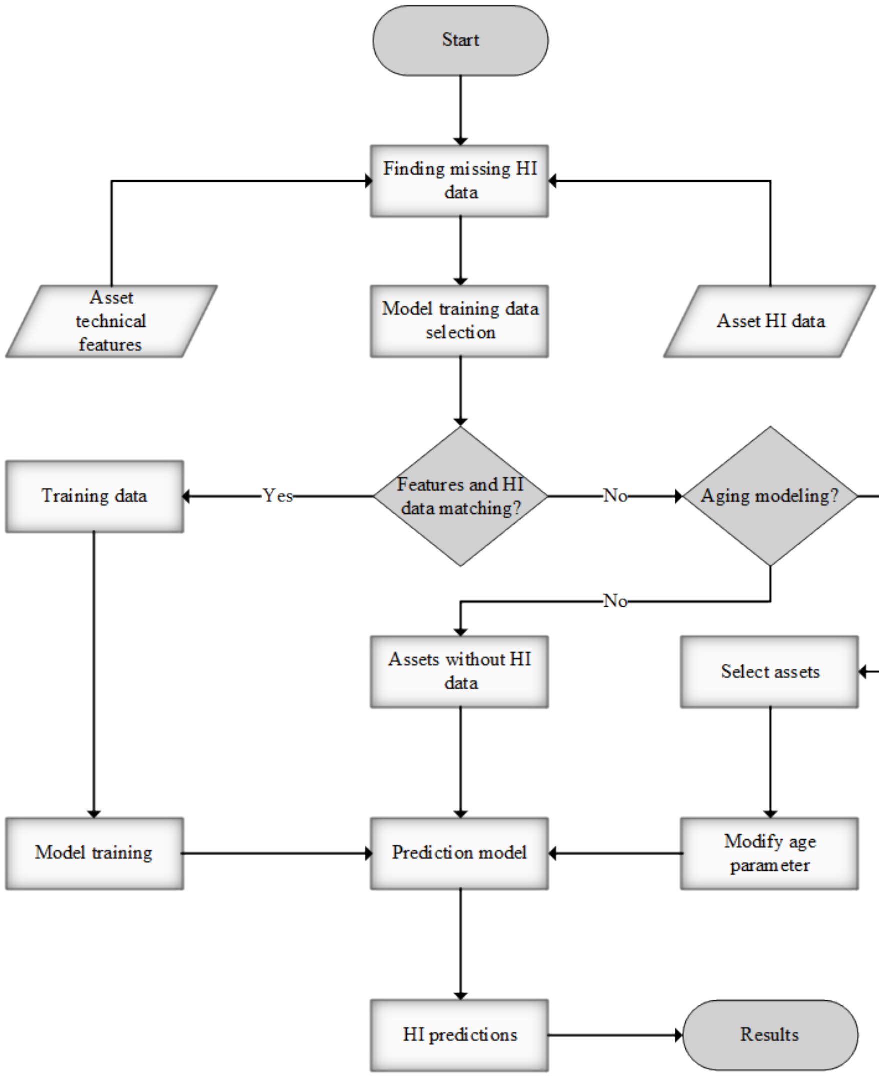


Figure 16: The flowchart of the asset HI prediction model and its input data. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process. [81]

2.3.1 Input data

The input data for the HI prediction model consists of HI data combined with the asset's technical information and specifications. Technical features used for OHLs are collected during the asset design and the construction phases are listed in Table 6. Features are selected according to the possible influence to the speed of OHLs' aging process, construction quality or mechanical stresses. All these features are fixed at the beginning of the OHL lifecycle and do not change during service. The parameters used in this work as technical features are the number of circuits on a single tower, the tower type where the support and tension tower are separated, the nominal voltage level of the OHL (110 kV or 330 kV), presence of bird protection, and tower material (reinforced concrete poles, colored steel lattice towers, zinc-coated steel or untreated steel lattice towers). The total number of data features is presented in Table 6 where these assets encompass six different manufacturers and 214 different tower configurations. Angled towers are distinguished from vertical OHL segments due to the increased mechanical stresses they experience. The final feature is the current age of the tower, calculated from the installation date.

Table 6: A list of features used by the health index prediction model.

Feature name	Number of different features
Number of circuits	4
Tower Type	2
OHL voltage level	2
Tower material	4
Tower manufacturer	6
Tower configurations	214
OHL direction changes	2
Existence of a perch guard	2
Age	1 to 67
Total	**303**

The model training and testing data is collected from periodical visual inspection as described in Section 2.2. As up to 150 individual condition criteria were used, for the simplified example, HI values were aggregated on tower level and used according to (10) for OHL-supporting components such as poles, foundations and crossbars. It is possible to increase the accuracy of this methodology further by scaling it to a more detailed approach where the HI of each component is predicted individually. The input data is manually cleaned and errors are fixed in the database. However, it is also possible to detect incorrect values in the data automatically by using outlier detection algorithms as described in [33] and done in Section 1.2.2. The data used in this paper consists of 26 273 rows described in Table 7 and presented in Fig. 17.

Table 7: Training and testing data split of the health index prediction model.

Health Index Class	Training	Testing	Total	% of Total
HI0	6 217	1 555	7 772	29.58
HI1	1 799	450	2 249	8.56
HI2	8 964	2 241	11 205	42.65
HI3	2 770	692	3 462	13.18
HI4	1 182	295	1 477	5.62
HI5	86	22	108	0.41
Total	**21018**	**5255**	**26273**	**100**

The distribution of data is presented in Fig. 17 and Fig. 18. In Fig. 17, there is no strong relationship for OHL assets between the HI and age. Table 7 also supports this, whereby HI data is concentrated around HI2 with extremely imbalanced classes with 108 samples in the HI5 class and 11 205 samples in the HI2 class. There are some HI5 assets with just ten years of service and some HI0 towers even after 60 years. The regression line based on linear regression on the graph demonstrates that there is a slight aging tendency where HI is increasing as the asset ages, but it never exceeds HI2 even after 60 years of service. Fig. 18 presents a data description chart where a majority of assets in the grid are 40 to 60 years old with HI2. There is also a larger concentration of HI0 towers with an age of around ten years.

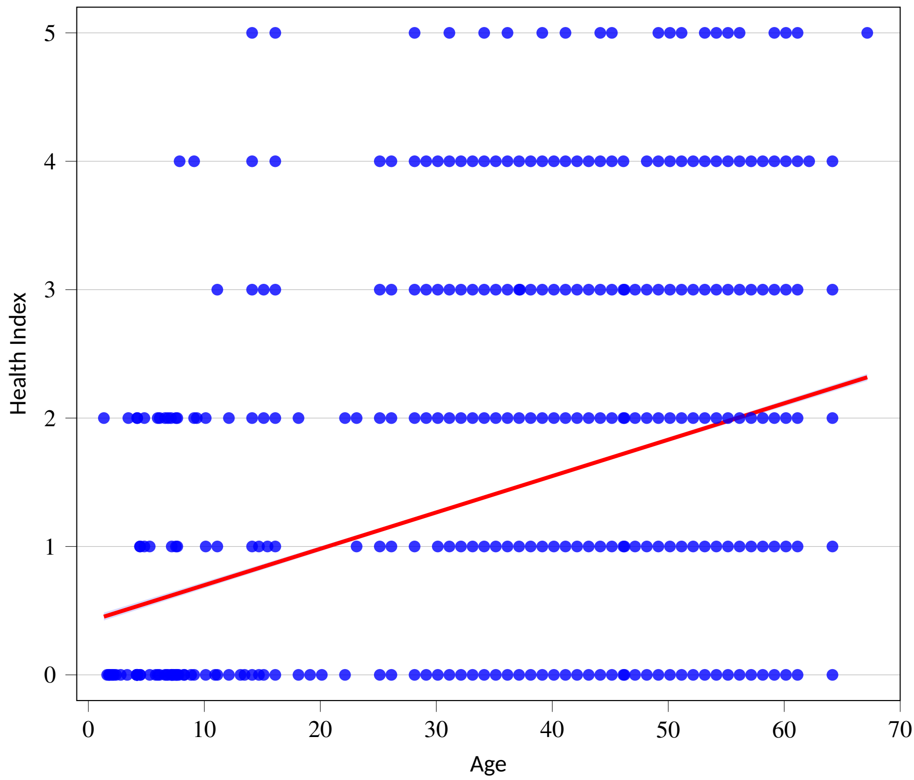


Figure 17: The scatter graph to observe relationships between tower's health index and their age. Blue is representing data point and red the linear regression line.

The data is divided into training and testing data with a ratio of 80:20 by using stratification to maintain equal class proportions for each class. That enables the percentage of samples for each class to be preserved and therefore ensures that both datasets include the necessary samples for all classes. Stratified re-sampling [82] is easy to implement and has a positive effect on both the variance and bias, especially in the case of a class imbalance. Table 7 shows that the imbalance data tendency remains the same after training and testing data split and all classes include a proportional number of training samples.

2.3.2 Class-imbalance

Fig. 18 accurately presents the impact of an interval-based preventive maintenance strategy on class imbalance that is used by TSOs to minimize risks associated with failures. As TSOs do not run their assets until failure then only a small proportion of assets reach high HI values. HI4 and especially HI5 are considered as minority classes and other HI values, such as HI0 and HI2, are considered as majority classes. This produces a significant imbalance between high and low HI classes resulting in classifiers that have poor predictive accuracy towards the minority class compared to the other classes. According to [82] and [83], class imbalance occurs when classes exhibit significant imbalances in the order of 100:1, 1 000:1 and even 10 000:1 between majority and minority classes. As looking at Table 7 there are 11 205 towers with HI2 and only 108 towers with HI5, this results in a ratio of majority vs minority class of 104:1.

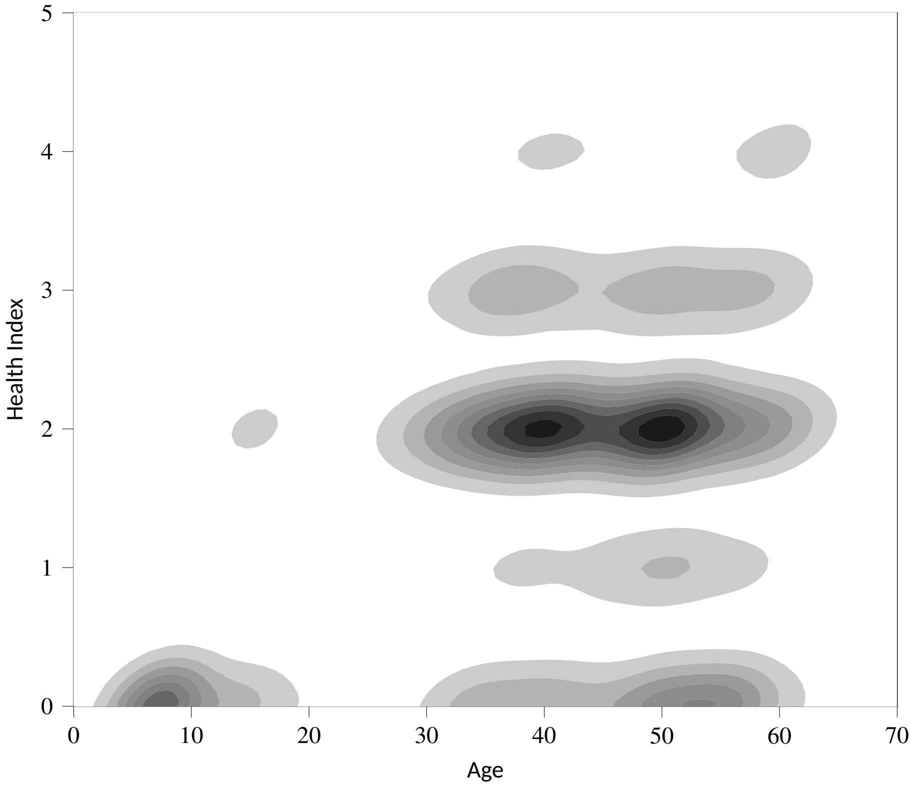


Figure 18: The concentration of overhead line towers according to health index and age. Denser distribution of towers is presented in darker colors.

In the case of class imbalance the classification prediction models are not the same as when using balanced data for model training and tend to classify most unseen samples in the majority classes. A decrease in model performance is caused by the model's loss functions, which attempt to optimize error rate or the accuracy of the model across all classes and without considering the real distribution. Thus, the model is achieving optimal results with majority classes. As this decreases the performance of the model for minority classes different methods are thus used to minimize the results of imbalanced data in prediction model learning processes. The three most widespread methods to eliminate the effects of imbalanced data are a down-sampling of majority classes [82] using random sampling, up-sampling of minority classes using random sampling and up-sampling minority classes [83] using the Synthetic Minority Oversampling TEchnique (SMOTE) [84].

Up-sampling of minority classes means that the class with the most instances is selected as a reference number and all minority classes are multiplied until all reach the same number as the majority class. In this data the majority class is HI2 with 8 964 samples in training data. There are two main ways to perform up-sampling, where the first is to multiply minority class samples by randomly duplicating existing samples in the selected class until it reaches the required population. SMOTE is a very popular oversampling method in multiple field of studies such as genetics. It was proposed to improve random oversampling by combining two similar linear samples of data from the minority class and therefore producing new data that is similar to the class average but not exactly the same data as already present in the database. Down-sampling of majority classes is similar to random up-sampling but in this case instances are deleted instead of duplicated in the dataset. It is done by selecting the class with the least instances to randomly delete all instances of larger classes until all classes have the same number of samples as the minority class. In this paper all classes are down-sampled until each class has the same amount of instances as the minority class. Training data for each HI class after up-sampling using SMOTE and down-sampling compared to unmodified data are presented in Fig. 19, where there are 86 instances in every class for down-sampling and 8 964 for SMOTE.

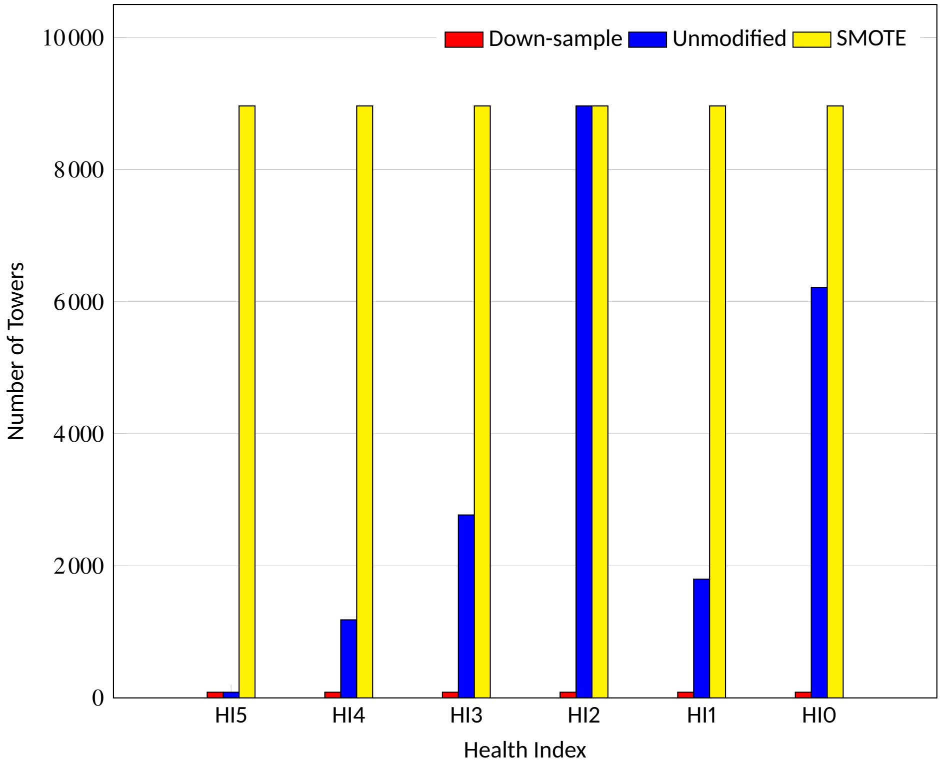


Figure 19: Training data distribution after training and testing data split for each health index class using different re-sampling techniques.

There are also other methods to reduce class imbalance that are related to balancing the training weights of models as can be done for certain algorithms such as Logistic Regression (LR) and Support Vector Machine (SVM). Some algorithms handle class imbalance better than others such as Decision Trees (DT), K-Nearest Neighbor (KNN), Random Forest (RF) and Gradient Boosting (GB) [85].

2.3.3 Steps Before Model Training

Parameter Tuning

For parameter tuning it is also important to keep track of the model learning process to prevent under- or over-fitting of models by using learning curves [86]. From Fig. 20, the GB model reaches its maximum accuracy for cross-validated score faster than RF, but the overall performance of RF is better. Also, RF improves its generalization accuracy to a larger magnitude compared to the GB model while increasing training set size. These trends show that neither model is over-fitted and the GB model is generalizing training data more compared to the RF model by reaching almost optimum training and cross-validation score ratio at the end of training process.

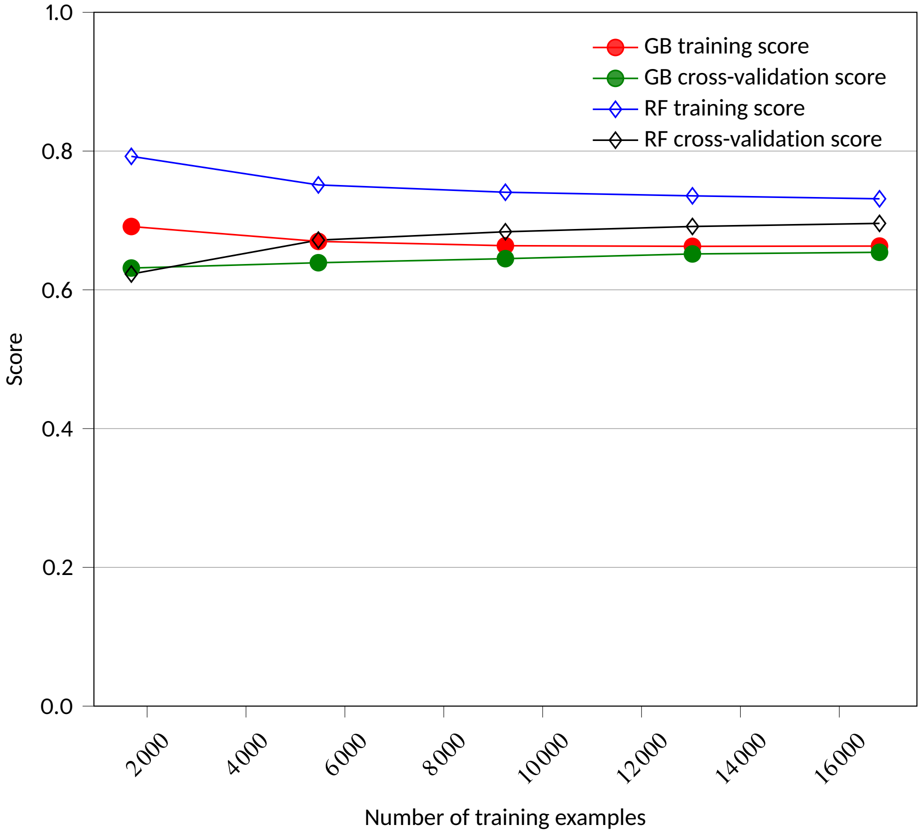


Figure 20: Learning curves for Gradient Boosting and Random Forest prediction models.

Data Cross-Validation

There are always impurities in the training dataset and not all samples are equally distributed even in the case of stratification. To overcome this issue, especially in small and imperfect datasets, nested cross-validation is used to reduce the bias for both hyper-parameter tuning and model evaluation. Nested cross-validation and its benefits in training and testing split and k-fold cross-validation are thoroughly described in [87] and [88]. In terms of computational complexity, a relatively simple 5X2 setup is used in the thesis. That means a, 5-fold cross-validation is implemented in the outer loop, and 2-fold cross-validation in the inner loop. As shown in Fig. 21, the inner loop is responsible for the model selection process, and the outer loop is responsible for estimating the generalization error.

For both inner and outer loops k-fold cross-validation described in [88] are used. The training data is divided into training and validation sub-datasets by using cross validation with five folds for the outer loop and with two folds for the inner loop. The data is divided into smaller subsets and one set is considered as validation data, and all others as training data. The model is trained using training data and performance metrics are calculated using the validation set. This is iterated as many times as are folds and mean with standard

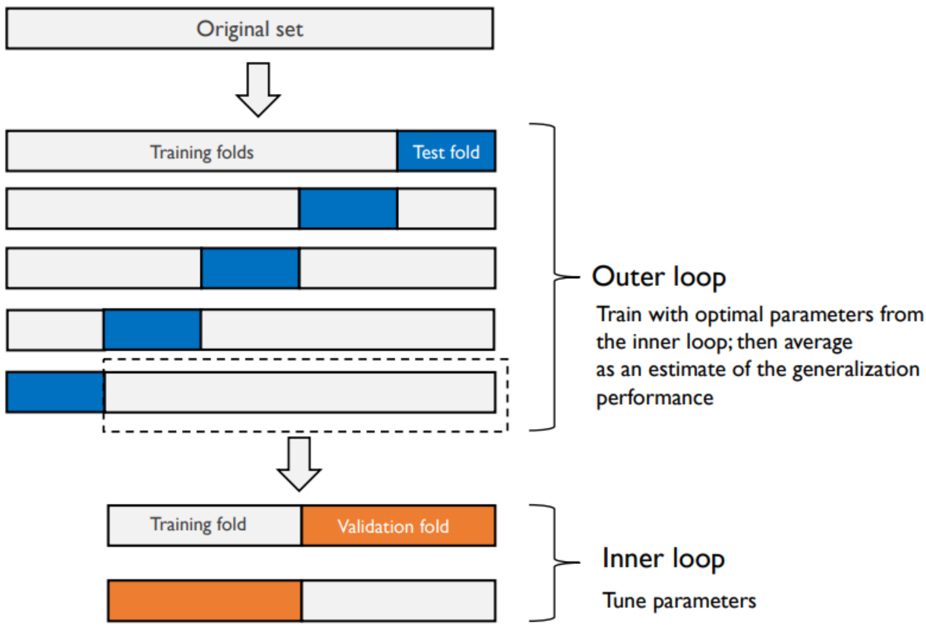


Figure 21: 5X2 setup of the nested cross-validation. [87]

deviation is calculated over all iterations. In order to preserve the ratio of minority and majority classes in subsets stratified cross-validation is used. The theory behind nested cross-validation and its benefits in over simple training and testing split or k-fold cross-validation are thoroughly described in [88] and [87].

Feature Importance

It is possible to determine the most important features of the dataset that affects model output the most. The most important features of two different and best performing prediction models, RF and GB, are presented on Fig 22. According to [89] boosting usually ignores some variables or features completely and RF should maximise the use of different features. This occurs because there is candidate split-variable selection built in RF that increases the chance of the variable being included in the model. In this case, both models are producing similar results but GB is less dependent on the features that have a low impact on the results. All features have a similar importance score for both models, where the most important feature is the age followed by tower type.

2.3.4 Supervised Machine Learning Classification Model Selection

The most common supervised machine learning classification algorithms are used to compose prediction models where the most of the major approaches are represented. The classification algorithms used are LR, SVM, Naive-Bayes Classifier (NBC), KNN, Multi-Layer Perceptron from Neural Networks (NN) and DT. There are also combined algorithms using multiple methods to improve prediction accuracy compared to a single algorithm by to converting a set of weak learners to a single strong ensemble model. Ensemble methods are used in supervised machine learning in order to obtain better predictive results from the model by using multiple learning algorithms in a single model rather than using any of the learning algorithms alone. RF [90] and GB [91] are widely used ensemble algorithms

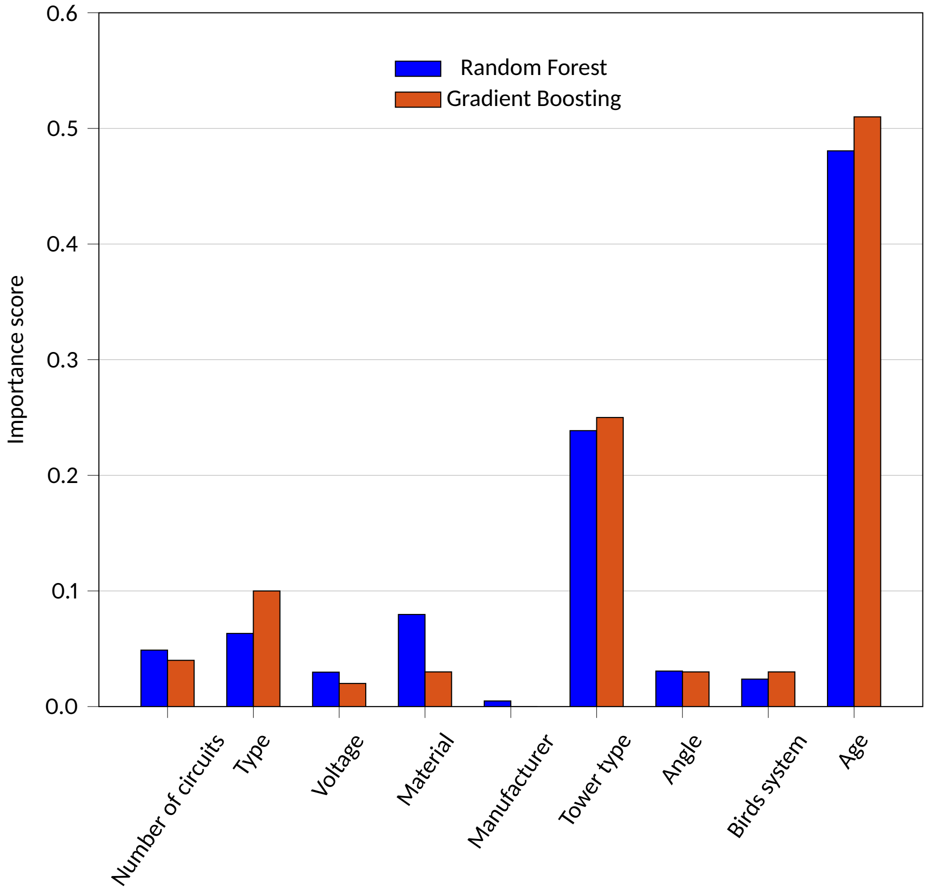


Figure 22: The feature importance of Random Forest and Gradient Boosting models.

that are also used for model selection. RF [90] is a bagging classification and regression algorithm that is based on decision trees and it has grown in popularity over the last few years because of its great performance. It is developed on an ensemble of unpruned trees, induced from bootstrap samples with random feature selection in the tree induction process of the training data. A prediction is made by using a majority vote to aggregate the predictions of the ensemble models. To overcome the class imbalance issue [85] proposes to use balanced or weighted RF models that are both evaluated during hyper-parameter tuning.

2.3.5 Performance metrics

To evaluate the performance of models, the test dataset is processed by the prediction model and compared manually to the ground truth data. The results are then classified into three categories:

- True Positive (TP) - Where the network has correctly predicted the value or identified a defect.
- False Positive (FP) - Where the network has predicted the value or detected a defect where none existed.

- False Negative (FN) - Where the network has failed to predict the value or detect an existing defect.

The aggregated values in these categories are used to calculate accuracy (17), precision (18), recall (19), F1-score (20) and specificity (21).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{17}$$

Precision or confidence, is a measure of the proportion of predicted positive cases that are correctly classified as real positives and is given by:

$$Precision = Confidence = \frac{TP}{TP + FP} \tag{18}$$

Recall or sensitivity, is a measure of the proportion of real positive test cases that are correctly predicted positive.

$$Recall = Sensitivity = \frac{TP}{TP + FN} \tag{19}$$

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{20}$$

$$Specificity = \frac{TN}{TN + FP} \tag{21}$$

Widespread performance metrics Receiver Operator Characteristic (ROC) [92] and Precision-Recall (PR) curves [89] were left out of the scope because they required to use of a One-vs-All approach that amplified the class imbalance issue and decreased the performance of models.

2.3.6 Model evaluation

Each selected classification algorithm is tested using a number of hyper-parameters specific to the algorithm. Hyper-parameters are modifiable parameters of algorithms, which affect the performance of models drastically. These could be such as the number of trees in RF or the regularization strength of an L2 penalty in the loss function of LR. As the selection of hyper-parameters relies heavily on the input data, usually this process is done manually. It requires a great number of experimentation to identify best values out of all combinations for each algorithm. This process is extremely computationally intense and time-consuming, but essential in order to increase model performance by selecting the best hyper-parameters to maximize the performance of models. Practical selection of hyper-parameters and different approaches are described in [87]. Also, three previously mentioned different datasets (unmodified, SMOTE and down-sampled) are used to test each algorithm and hyper-parameter set. Modelling and data processing is done in the Python 3.7 environment with Scikit-learn module [35]. The theory and implementation of supervised machine learning algorithms in practice is discussed thoroughly in [93], [94] and [88].

Model evaluation is done through model hyper-parameter optimization using a random search with nested 5x2 cross-validation to find the most suitable model for the used dataset and algorithm and to reduce the bias of training data. That enables the determination of the best hyper-parameters for each classification algorithm and therefore, their best performance to be compared to actual data. Model parameter optimization is performed by using random search combined with nested cross-validation. According to [95], random search over the same domain is able to find hyper-parameters for models that are as good or better within a small fraction of the computation time of a pure grid search where all possible combinations are tested. The possible number of combinations with selected limits for each models is presented in Table 8, while the random search limits tested combinations for each model by 100. A relatively simple 5x2 nested cross-validation setup is used to reduce the bias for both hyper-parameter tuning and model evaluation with a low level of computational complexity.

Table 8: Hyper-parameter combinations of the tested prediction models.

Algorithm	Tested models
Naive-Bayes Classifier	-
Neural Network	72
Random Forest	72
Logistic Regression	80
K-Nearest Neighbor	372
Support Vector Machine	400
Gradient boosting	576
Decision Trees	1 440
Total	**3 012**

Model evaluation results of 24 individual models are presented in Fig. 23, where accuracy, precision, recall and F1 score are used. Precision, recall and F1 scores are calculated by taking the average of all class values instead of the average across all results over all testing samples to indicate class imbalance more effectively. Six models RF, GB, NN, KNN with SMOTE data, RF with SMOTE data and KNN clearly perform better than the other 18 models. All models have an accuracy over 60%. The accuracy of the best performing models was almost 70% based on training data. The poorest performing models are NBC and LR algorithms, which are the most sensitive to the imbalance of the data. In terms of input data, the best results were achieved by using unmodified data. The SMOTE approach outperformed the down-sample and is comparable with unmodified data in terms of recall and F1, but worse in terms of precision. Down-sampling of training data performed worse than SMOTE and unmodified data.

Six models with the best performance are selected for more detailed analysis where the performance metrics of each class are presented individually. Fig. 24 illustrates the precision of RF and GB outperforming all other models in all classes. It is also seen that NN was not able to detect some classes at all, especially minority class HI5.

Figure 25 presents recall results for the top six models where all models that used SMOTE datasets outperformed models with unmodified data, especially the RF model with SMOTE. RF and GB with unmodified data perform poorly on minority classes compared to RF using SMOTE data and KNN algorithm. The NN model was again unable to classify minority classes correctly. Figure 26 presents the F1 scores of the top six models and the best performing models are RF, GB and KNN with unmodified training data.

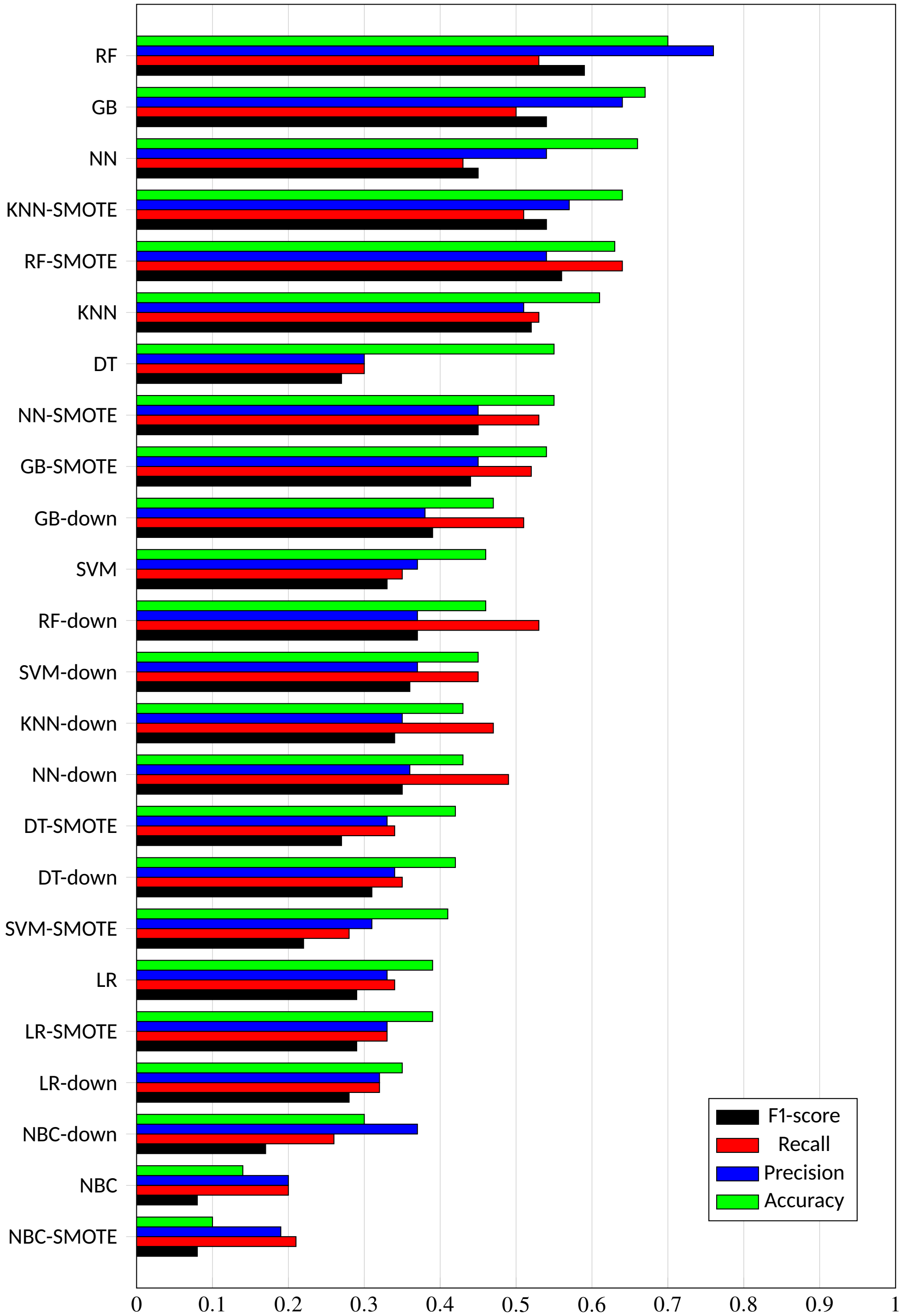


Figure 23: Performance metrics of tested models according to health index.

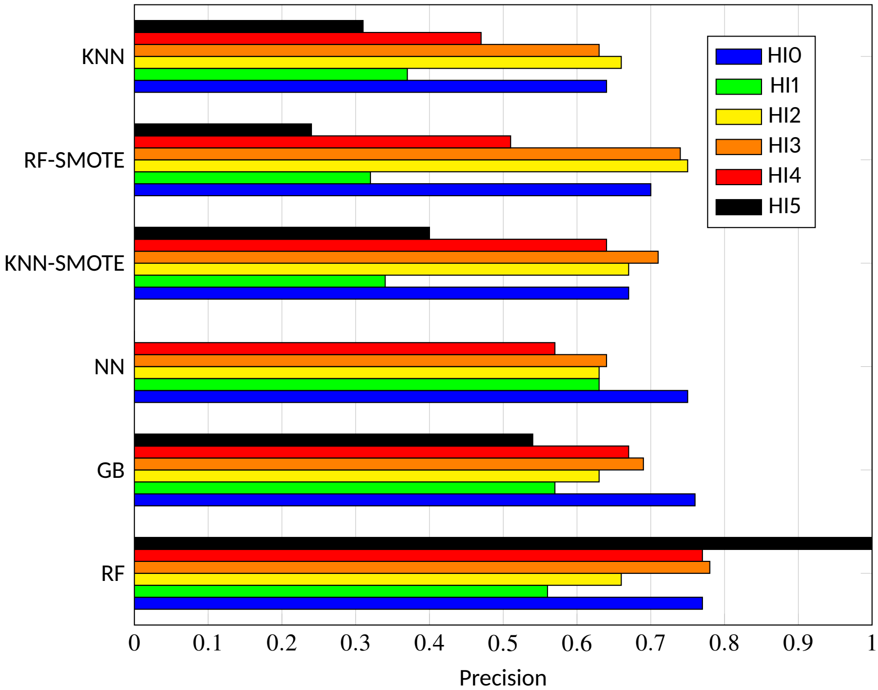


Figure 24: Precision score for top six models according to health index.

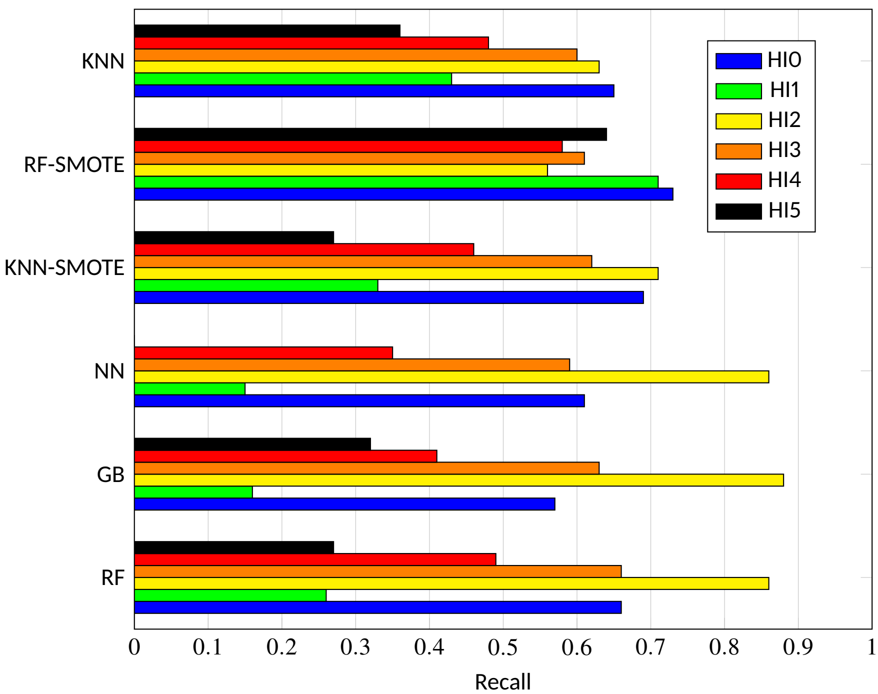


Figure 25: Recall score for top six models according to health index.

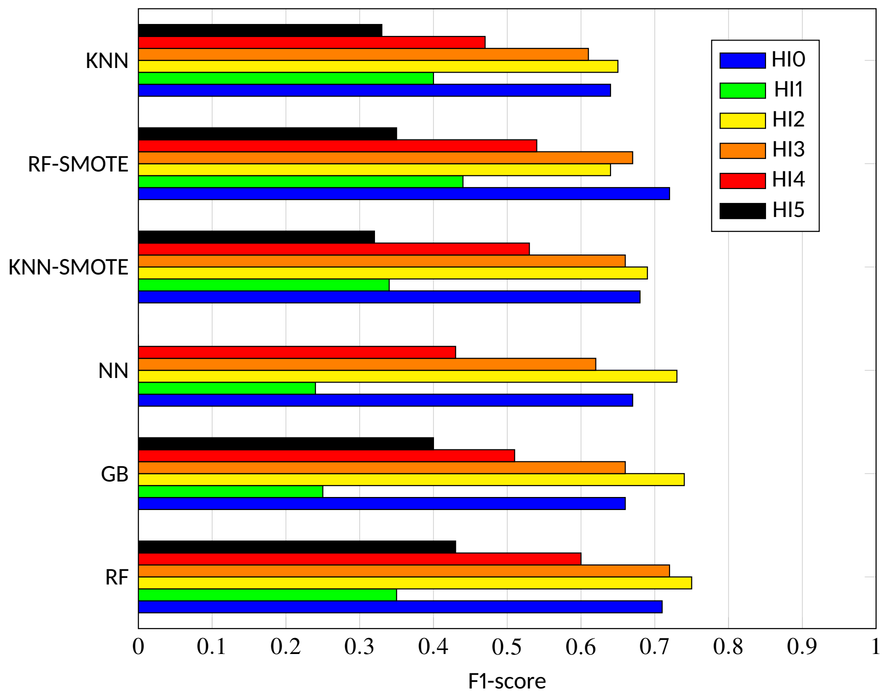


Figure 26: F1-score for top six models according to health index.

RF with unmodified training data outperformed all other models according to Fig. 24 to Fig. 26 on all performance metrics except recall, where the best performing model was the RF algorithm using SMOTE dataset. Due to the advantages in computational requirements and that RF with unmodified data outperformed other models in nearly all aspects, especially in precision, the RF model is implemented as a case study in Chapter 4. The best model was based on Random Forest algorithm and it used unmodified data. The selected hyper-parameters are:

- Number of estimators = 100.
- Minimum samples leaf = 1.
- Maximum features = sqrt.
- Class weight = None.

3 Automatic Condition Assessment using Object Detection

Implementing LiDAR technology to acquire precise data about ROW and OHL enables TSOs to take images of each OHL tower as a by-product to point-clouds at a marginal cost. Also, due to the rapid developments of drone technology, cost-effective data collection approaches emerge through the use of unmanned aerial vehicles (UAVs) and automatic flight routes, as seen in [96, 97]. As technological advancements make it possible to collect a large number of images about the grid cost effectively, there is still a large workforce required just to examine each image individually. In recent years, object detection methodologies using deep learning techniques in the field of computer vision have become a hot topic of research as they enable the possibility to detect a large variety of different objects from images. That makes it possible to process large amounts of data automatically with a fraction of the workforce compared to manual inspection. Object detection itself is the method of both recognizing an object class and predicting the location of the object via a bounding box. These techniques have been applied in the fields of medicine [98], intelligent vehicles [99], agriculture [100] and have begun to be implemented for damaged aerial power lines [101, 102] to detect the most critical defects.

This chapter presents an automated condition assessment methodology for OHLs based on deep learning object detection networks. Three object detection approaches are proposed that differ from each other by the level of detail of input images and number of defects to detect. The first object detection model is based on the publication **III** and is described in Section 3.2. It focuses on the detection of a single critical defect and it is trained on a small number of images. The second object detection model is based on the publication **V** and is described in Section 3.3. It focuses on the detection of nine different defects and is trained on larger number of images that are taken close from reinforced-concrete poles and foundations. The third model described in Section 3.4 is the most sophisticated one that uses multi-staged approach to detect defects from super-high-resolution images ($>$100 MP).

3.1 General Overview

A general overview of automatic HI determination from images is presented in Fig. 27, where the most important processes are presented. The list of poles that are inspected is based on the restrictions from the maintenance strategy. Regardless of the specific object detection methodology, the automatic condition assessment of transmission OHLs starts with the input for a OHL inspection and image collection using handheld devices, helicopters or UAVs. Each image should be saved with a geotag for further automatic image-asset correlation. If captured images do not have geotag information, they will be marked as unknown assets by the model and they are required to link with assets manually. Once all images have been collected they will be pre-processed for object detection models. Object detection models are used to detect all defects which occur in images and the output of each object detection model is a list of defects with bounding boxes. Once the defects are detected, they must be combined and mapped to a single HI of the image or component.

To support the automatic condition assessment process a simple CSV database is used. It consists of five columns and the structure is presented in Table 9. For automatic image-asset correlation at least three columns should exist: UID of the asset, latitude and longitude coordinates. HI and defects columns are used to store detections and to determine the overall HI.

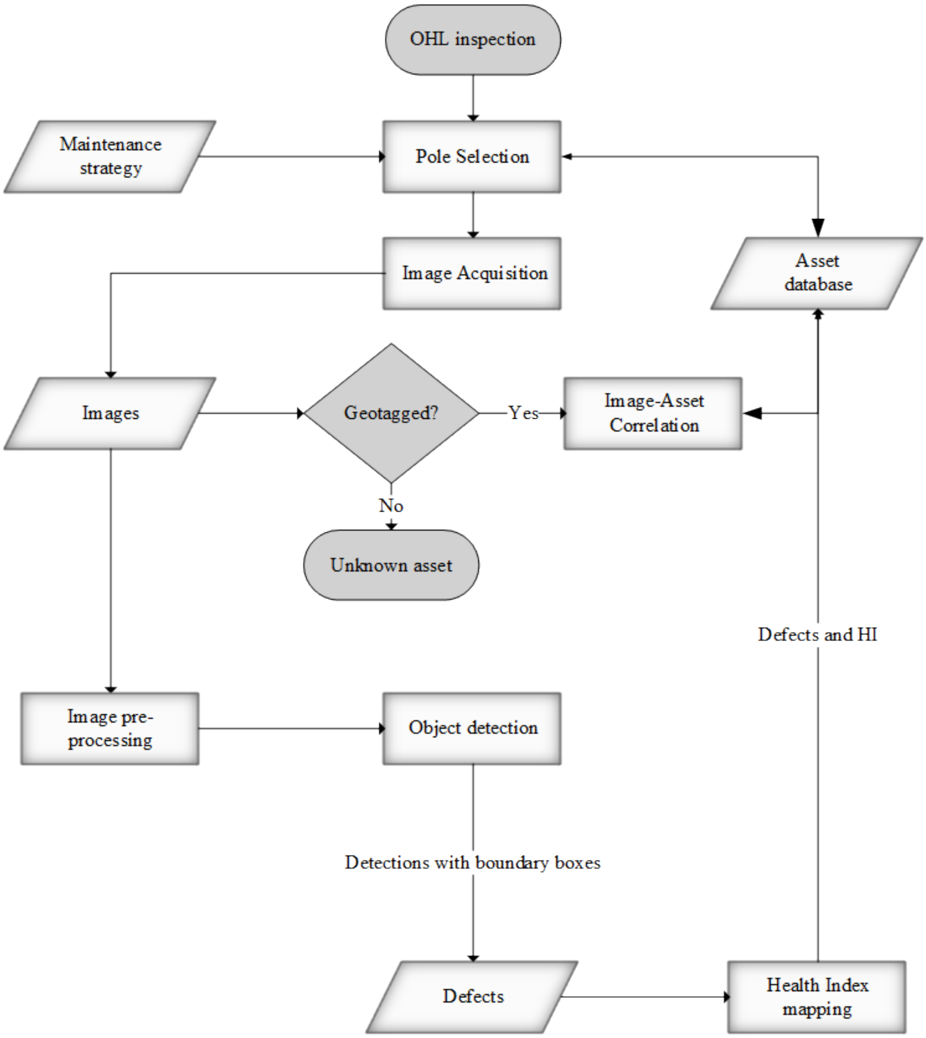


Figure 27: The flowchart of the automatic condition assessment model. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process. [79]

Table 9: A simple example of CSV database used for the automatic health index determination.

UID	Lat	Long	HI	Defects
Number	Coordinate	Coordinate	Value	List

3.1.1 Data Pre-processing

The data used to train and validate object detection models is pre-processed to decrease the computational complexity and to increase the accuracy of models. Data pre-processing is performed through the following steps:

- Resizing to required resolution.
- Manual labelling of each image.

- Splitting data into training and testing.
- Training data augmentation.

The main limitation to the usage of object detection networks is the maximum input image size as each pixel of the images is used in models. State-of-art object detection networks are optimised on input images up to 1536x1536 pixels as done for EfficientDet [103], but that increases the computational complexity of models compared to 512x512 pixels around 10 times. In order to reduce the computational complexity, each original high-resolution image was resized to 300x300 pixels for a single defect detector in Section 3.2 and 512x512 pixels for multiple defect detection models in Section 3.3 and Section 3.4. Original resolution images that were used in the model training and validation ranged from 5 MP to 100 MP before the resizing. After the resizing, bounding boxes were manually placed on each of the degradation artefacts to act as the ground truth for training and testing data. The number of individual bounding boxes for object detection models in input data is presented in Section 3.2, Section 3.3 and Section 3.4 separately. For the training and testing data split, a holdout cross validation method was used. From the dataset, 80% of the randomly chosen data was used for training and the remaining for testing.

Data augmentation is used to introduce data variability on the features and improve learning [104] without additionally collected training data. Training images of each model were augmented by varying the contrast, saturation, hue, magnification, brightness and horizontal flipping. This resulted in a higher number of training images without the requirement of additional images. To generate correct bounding boxes of augmented images, the K-medoids clustering algorithm [105] using the Intersection over Union (IoU) distance metric (22) was used to generate the anchor box sizes automatically.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{22}$$

where A denote the ground truth and B estimated bounding boxes.

In some cases several detections may be present in the near area of a single ground-truth object. In order to compensate for this, the proposal with the highest confidence score is usually selected through a process called Non-Max Suppression (NMS) as described in [106]. This approach has the potential to reduce the number of false positives and to increase the number of false negatives as the number of total detections is reduced through this process. In many cases NMS increases the precision and recall metrics and improves the overall of performance of the object detection network.

3.1.2 The Deep Learning Object Detection Networks

The Faster Region-based Convolutional Neural Network (Faster R-CNN) The Faster R-CNN deep learning algorithm [107] is a region-proposed object detection network which uses a two-stage framework that first scans the image and then focuses on regions of interest. The faster-RCNN technique advances from fast-RCNN since it does not rely on additional methods to generate a candidate pool of isolated region proposals. This results in reduced training and detection time due to the reduced computational complexity of the algorithm. As Faster R-CNN is much faster in detection than its predecessors (R-CNN and Fast R-CNN) it can even be used for near real-time object detection but it is still slower than single-stage networks due to its architecture. The faster-RCNN architecture comprises of a feature detection network (the Inception V2 model [108]), a Region Proposal Network (RPN) and a classifier. The RPN is a fully-convolutional network that generates proposals and can be trained via supervised learning techniques.

The You Only Look Once (YOLO) v2 YOLOv2 [109] real-time algorithm accomplishes object detection via fixed-grid regression. It is a one-stage framework that maps the image pixels to bounding box coordinates and class probabilities in a single step using a global regression technique, whilst region proposal frameworks such as fast/faster-RCNN, have several correlated stages that are each trained separately. The idea is to make object predictions on each feature map location without the cascaded region classification step. The feature extraction network consists of 23 convolutional, 5 max pooling, 2 routing and 1 reorganization layers. The main disadvantage of using YOLOv2 is the poor detection accuracy of small objects.

The Single Shot Detector (SSD) The SSD [110] object detection algorithm is similar to YOLOv1 as it is also a single stage network. SSD is more advanced than the YOLOv1 algorithm which results in better accuracy and faster detection times. It has also aimed to overcome some of the inabilities of the YOLOv1 algorithm, namely detecting small objects in groups. According to [111] the main issue with SSD networks their lower detection accuracy but due to the use of a single network it can be used in real-time. The feature extraction network for the SSD algorithm uses the RetinaNet backbone [112]. The RetinaNet architecture consists of a Feature Pyramid Network [113] on top of a feedforward Residual Network [114]. This topology has the benefit of using a focal loss feature which deals with the class imbalance problem.

The You Only Look Once (YOLO) v5 YOLOv5 [115] is an anchor based one-stage algorithm based on modifications of YOLOv4. The architecture consists of a backbone, neck and head. YOLOv5 uses the Cross Stage Partial Network (CSPNet) [116] as the backbone. CSPNet reduces model size and increases the speed of the model by addressing the problem of repeated gradient information in large-scale backbones. The Prototype Alignment Network (PANet) [117] method is used as a method of feature aggregation. PANet improves the location accuracy of objects by utilizing the localization signals in lower layers. Finally, the data is passed to the head layer of YOLOv5. Three feature-maps of different sizes are generated to achieve multi-scale prediction.

EfficientDet EfficientDet [118] is a single-stage object detection network that achieves efficient multi-scale feature fusion via a bi-directional feature pyramid network (BiFPN) and model scaling via compound scaling rather than using a larger input size or a bigger backbone. The EfficientDet backbone is constructed with 9 stages consisting of 18 layers. The backbone feeds into 3 BiFPN layers with 64 channels and finally the object detection network, which consists of 3 class/box prediction layers. The model consists of 3.9M parameters and has an input size of 512.

CenterNet The name CenterNet is built on the CornerNet pipeline [119] and [120]. The detection pipeline consists of the stacked hourglass-104 backbone [121] which consists of convolutional and max pooling layers to down-sample the image $4\times$. The image is then passed into two hourglass modules. Each hourglass module is a 5 layer down and 5 layer up sequence convolutional network. The information then flows into a down-sampling layer and then simultaneously to three heads: the heatmap head which estimates the keypoints for some input image, the dimension head which predicts the dimensions of the boxes and the offset head which recovers the discretization error caused by the output stride.

Models used in the thesis A list of deep learning object detection networks used in this thesis is Table 10. Faster R-CNN, YOLOv2 and SSD networks with their training parameters are further explained in [79]. YOLOv5, EfficientDet, CenterNet and the proposed ensemble model are further explained in [17].

Table 10: A list of deep learning object detection networks used in the thesis.

Used in Section	Number of models used	Name of the Network
Section 3.2	1	YOLOv2 [109]
Section 3.3	1	YOLOv2 [109]
Section 3.3	1	SSD [110]
Section 3.3	1	Faster-RCNN [107]
Section 3.4	4	EfficientDet [118]
Section 3.4	4	CenterNet [120]
Section 3.4	4	YOLOv5 [115]
Section 3.4	3	Ensemble model

3.1.3 Automatic Image-Asset Correlation

The majority of modern photo cameras, smartphones, drones and tablets save geographical information as geotags to images using Global Positioning System (GPS) receivers that can be used to link images with assets automatically. Geotaging of images ensures that the latitude and longitude are recorded while each image is taken. The automatic image-asset correlation describes a method where images taken of OHL towers during visual inspections or using UAVs are linked together with physical assets that have geographical information stored in a database by calculating the minimum distance of each image to all physical assets. Geographical information about assets the in the thesis is already gathered in an asset database by using LiDAR technology during airborne inspection. This results in precise coordinates of each individual tower and by comparing that data against an image's geotag it is possible to determine the distance between the asset and the image. The distance between two different points on the Earth's sphere can be accurately calculated using Haversine formula (23) [122].

$$D_{real} = 2r \arcsin \sqrt{\sin^2\left(\frac{lat_a - lat_i}{2}\right) + \cos(lat_i)\cos(lat_a)\sin^2\left(\frac{lon_a - lon_i}{2}\right)}, \tag{23}$$

where D_{real} is the distance between two points, lat_a, lon_a, lat_i and lon_i are the latitude and the longitude of the asset and the image respectively. And r is the radius of the Earth ($r = 6731km$). It should be noted that lon and lat of the image and the asset must be both in the same coordinate system. In this example all coordinates in the database are in $EPSG : 3301$ coordinate system and geotags are in $EPSG : 4326$ system. Therefore to use (23), all coordinates are first converted to $EPSG : 4326$.

In addition to distance calculation, it is also important to assess possible errors in GPS accuracy to determine possible confidence thresholds for the successful image-asset correlation. For the successful mapping of potential towers with images, the distance between two points must be less than the specified threshold value D_{max} that can be expressed by using (24). Once the D_{max} is chosen to be too large then there might be an issue whereby more than one asset may be linked with the image. If D_{max} is chosen to be too small then no asset may be in the radius and image-asset correlation is not giving the expected results.

$$D_{calculated} < D_{max} \tag{24}$$

Where $D_{calculated}$ is the calculated distance between the asset and taken image and D_{max} the maximum allowed distance based on the grid configuration and GPS error.

Determining distances between image and all assets and filtering all calculated distances by using limit value reduces errors that may be incurred from erroneous GPS data. The reliable value of D_{max} for OHLs must be determined empirically and it is primarily influenced by two components that are caused by the grid configuration and GPS error. Influences from the grid configuration can be caused by different OHL span distances and geographical peculiarities of terrain and the presence of parallel OHLs in the same right-of-way. The assessment of GPS errors is affected most by the used technology and disturbance of GPS signal by the environment. Transmission OHLs can be considered as open areas for GPS receivers since tall vegetation must be cut in the right-of-way to prevent flashovers from conductor to vegetation and safety distances from buildings and structures are usually measured in tens of meters. All this is beneficial in terms of GPS position accuracy as there are no disturbances caused by trees or tall buildings. Most smartphones and tablets, as used by the assessment tool in Section 2.2, use assisted GPS (A-GPS), where the mobile networks' signals are used in addition to satellites to receive the geographical position of the device by triangulating the signal. Based on [123], an open field's A-GPS mean error can be estimated at 4 meters, the consumer-grade GPS's mean error is under two meters. The standard deviation was approximately two meters for both, A-GPS and GPS. In addition to A-GPS, dual-frequency GPS receivers are used in state-of-the art smartphones that have increased accuracy compared to A-GPS. Even though A-GPS uses a network signal in addition to the GPS signal to increase the accuracy in environments where it is difficult to receive a GPS signal, it has worse accuracy in open fields compared to regular GPS that uses at least four satellites to determine the device's position. Based on that it is possible to conclude that A-GPS's parameters can be used for the D_{max} determination as the worst-case-scenario.

3.1.4 Defect Detection and Health Index mapping

Once the defects are detected using object detection models, they must be combined and mapped to a single HI for an asset to use them in PoF determination. As the HI of assets should reflect the impact to overall reliability by its actual technical condition, HI calculation is based on maximum function of all determined defects according to (10) as done in Section 2.2. There is also a possibility to count all detected defects individually to compile custom warnings if predefined criteria are exceeded but in this chapter this was not studied further as it depends on the end-user preferences. If there is more than one image detected per asset then the proposed model adds new defects to selected asset database and calculates new HI according to all determined defects using 10. That means if there are defects with HI2 in one image and defects with HI5 in another image then the asset HI=5.

3.2 Detection of a Single Defect Using Close-up Images

This section presents the simplest object detection model in the thesis, which is trained to detect a single defect from close-up images. The selected defect to detect is holes in reinforced-concrete poles. Images for the object detection model training and testing data are taken from additional periodic visual inspections by foot patrols described in Section 2.2.

3.2.1 Data Description

The dataset used in this section comprises 150 images of concrete poles with varying degrees of holes. As images were taken by foot patrols while doing visual inspections as described in Section 2.2 there was no standardization of image requirements. This resulted in the images presented in Fig. 28 that exhibited high background variability, various pole orientations, different picture angles, shadows and the presence of external objects on images such as hands, vegetation and pens.

(a) *(b)* *(c)* *(d)*

Figure 28: The input variability of images with vegetation (a), earth (b), clear sky (c) and with an hand in the foreground (d). [124]

Training and testing data is described in Table 11. A total of 150 images were used where images were split to training and testing with the ratio of 80:20. 20% of images in training were augmented. The final training dataset comprised of 144 original and augmented images with total 180 individual defects describing distinct variations of holes in reinforced-concrete poles as seen in Fig. 28. Testing dataset comprised 37 images with 43 individual identifications of holes in reinforced-concrete poles.

Table 11: *Training and testing data of a single defect detection model.*

Class	Augmented training	Training	Testing
Hole	180	132	43
Images	144	113	37

3.2.2 Performance of Networks

Performance of object detection networks is analysed by comparing its detection bounding box to all ground truth bounding boxes of the same class in the image. This determines whether a detection is true positive, false positive or false negative and precision together with recall can be calculated.

Fig. 29 presents the average precision of the object detection model as a function of various IoU thresholds. At IoU 0.4, there is also a point below which the precision does not improve while decreasing the IoU threshold. The precision and recall over the whole set of data for IoUs less than or equal to 0.4 is approximately 0.69 and 0.83 respectively. Higher IoUs indicate better detection localization, but this is not that important for this particular defect detection model as the main goal is to detects holes on the pole or not. The IoU for the object detection model should be selected according to the highest recall and precision scores and defect localization on the image is not that important. By using a lower confidence threshold it is possible to increase the recall to minimize defects not detected, but that usually results in a decrease in precision as more FP will be detected.

3.3 Detection of Multiple Defects Using Close-up Images

3.3.1 Health Index Overestimation and Underestimation Performance Metric

In addition to the performance metrics used in Section 2.3, the *IoU* and Confusion Matrix are used to evaluate the performance of object detection models. *IoU* performance measure gives a similarity between the predicted region and the ground-truth region for an object present in the image. It is defined as the size of the intersection divided by the union of the two regions and calculated using (25).

$$IoU = \frac{\text{area}(B_p \bigcap B_{gt})}{\text{area}(B_p \bigcup B_{gt})} \tag{25}$$

where B_p is the predicted bounded box and B_{gt} is the ground truth bounding box.

A *Confusion Matrix*, presented as a Table 12, is a useful metric to accurately present the outcome of classification in image recognition problems. Each row denotes the instances of an actual class and each column gives the instances of the prediction. Confusion matrices for object detection algorithms are evaluated in a similar way, however they utilise multi-class instances where the matrix compiles each object class from the same image on a single table.

For the object detection issues used for multiple defects, it is more important to analyse the network's performance in relation to estimation of the HIs and not only individual defects, which are derived from TP, FP, FN and FP performance metrics.

Two new general performance metrics for the object detection algorithm are defined as overestimation and underestimation scores of the image HI, given in (26). Usually defects escalate smoothly from one to another followed by using maximum HI value of detections as the reference point to assess networks' performance produces more realistic results than just using detections. Overestimation O_i is a measure of how much the network

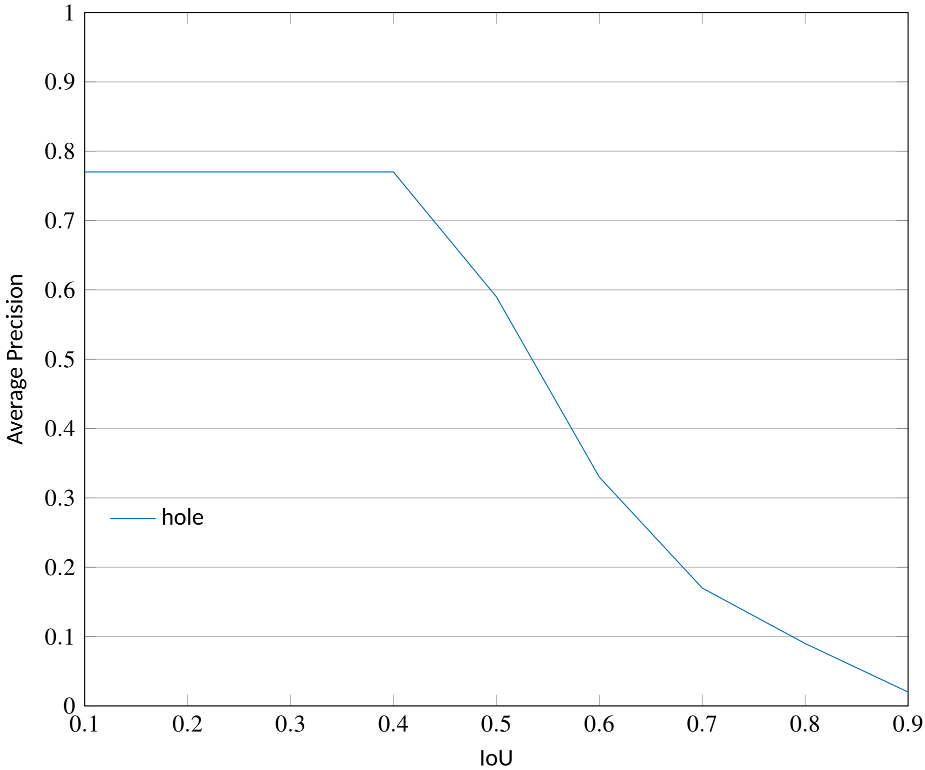


Figure 29: Average precision vs IoU of the single defect detection model.

Table 12: Confusion matrix based on estimated and actual health index. [79]

		Estimated		
		HI 1	**HI 2**	**HI 3**
Actual	**HI 1**	λ_1	$a_{1,2}$	$a_{1,3}$
	HI 2	$a_{2,1}$	λ_2	$a_{2,3}$
	HI 3	$a_{3,1}$	$a_{3,2}$	λ_3

detects and classifies defects as higher HIs than they actually are, and underestimation U_i is the converse. A Confusion Matrix is then generated with these scores as shown in Table 12. The rows of the table gives the true HI, while the column gives the estimated HI. The diagonal of the table, λ_i, therefore, represents the correctly estimated indices, the lower triangle the under-estimated HIs and the upper triangle, over-estimated HIs.

$$O_i = \frac{\sum_{j=i+1}^{5} a_{ij}}{\sum_{j=1}^{5} a_{ij}} \times 100, \quad U_i = \frac{\sum_{j=1}^{i-1} a_{ij}}{\sum_{j=1}^{5} a_{ij}} \times 100 \tag{26}$$

where a_{ij} is the element in the i-th row and j-th column of the confusion matrix.

3.3.2 Data Description

Images for training and testing data to train the object detection model were collected during the visual inspections explained in Section 2.2. A total of 1 008 images were labelled manually, which resulted in 3 544 individual labels in nine classes. Detailed input data division is listed in Table 13. Some classes such as visible reinforcements and minor defects include significantly more defects than in classes that describe end-of-life condition indicators such as significant loss of cross section and holes. This imbalance occurs because there are few critical defects in the grid and large amount of minor defects due to the PM strategy.

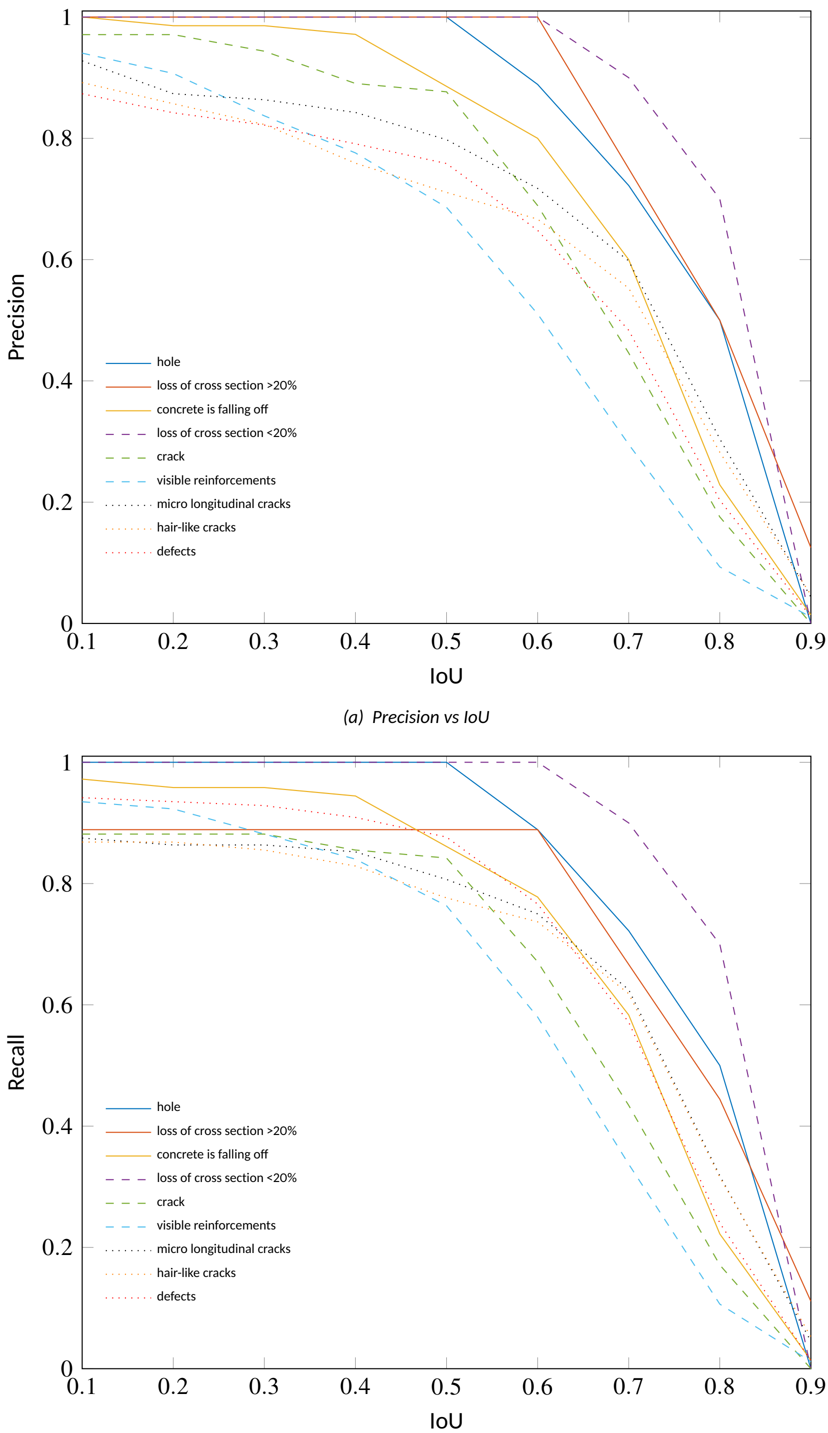
Table 13: Training and testing data description of multiple defect detection models according to defect types.

Defect	Training	Testing
Loss of cross section $> 20\%$	41	9
Hole	66	18
Loss of cross section $< 20\%$	69	20
Concrete is falling off	289	72
Hair-like cracks	351	78
Crack	356	76
Micro-longitudinal cracks	418	89
Other minor Defects	620	155
Visible Reinforcements	643	174
Total number of labels	**2 853**	**691**
Total number of images	**808**	**200**

3.3.3 Performance of Networks

Threshold values of an IoU between 0.1 and 0.9 are used for the bounding boxes to determine whether the detection corresponds to the ground truth or if the detection is not associated with the defect present in an image. Sweeps of NMS and IoU thresholds were done between 0.1 and 0.9 for each of the nine classes resulting in a 9x9x9 array for each of the detector networks (Faster R-CNN, YOLOv2, SSD). Fig. 30 to Fig. 32 present precision and recall vs IoU graphs for each object detection network separately. Based on thoughtful examination of the arrays, the NMS did not have a negligible effect on the results and its presentation on graphs was omitted out.

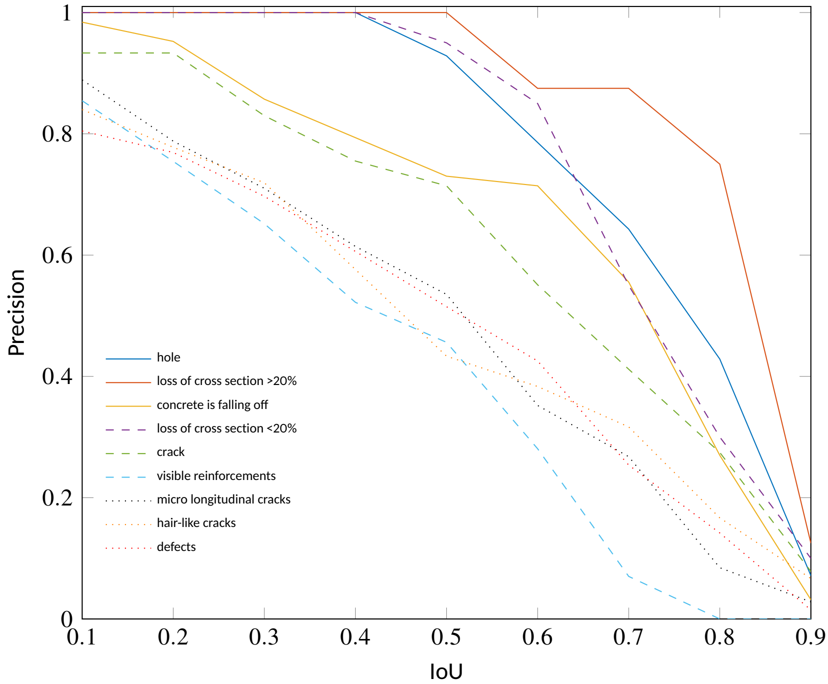
Fig. 30a to Fig. 31b indicate that the networks give the best precision-recall performance at low values for IoU thresholds. The effect of low IoU thresholds is likely to be due to the irregular defect shapes and obscure edges of artefacts such as the shape of cracks or unequal edges of various defects. That leads to increased complexity of manual labelling of bounding boxes for the ground-truth data. That kind of noisy labelling is a well known issue for multiple image classification problems and its effects have been studied in [125, 126]. While an IoU threshold of 0.5 is usually considered standard for object detection with well-defined input data boundaries, a significantly lower IoU of 0.1 was found to perform best for the defect detection problem. As the main purpose was to detect all defects on images and their exact location in image was secondary, IoU of 0.1 can be used in further detections. In further studies, methods to minimize noise in the ground-truth data should be examined to increase the performance of models. Noisy, or even wrongly labeled



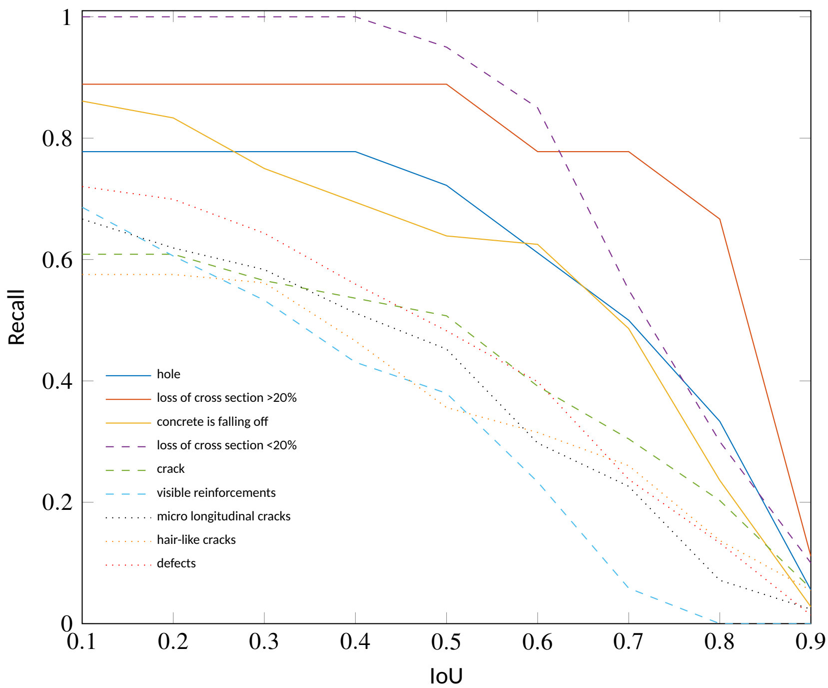
(a) Precision vs IoU

(b) Recall vs IoU

Figure 30: Precision vs IoU (a) and Recall vs IoU (b) for the YOLOv2 detector of the multiple defect detection model.

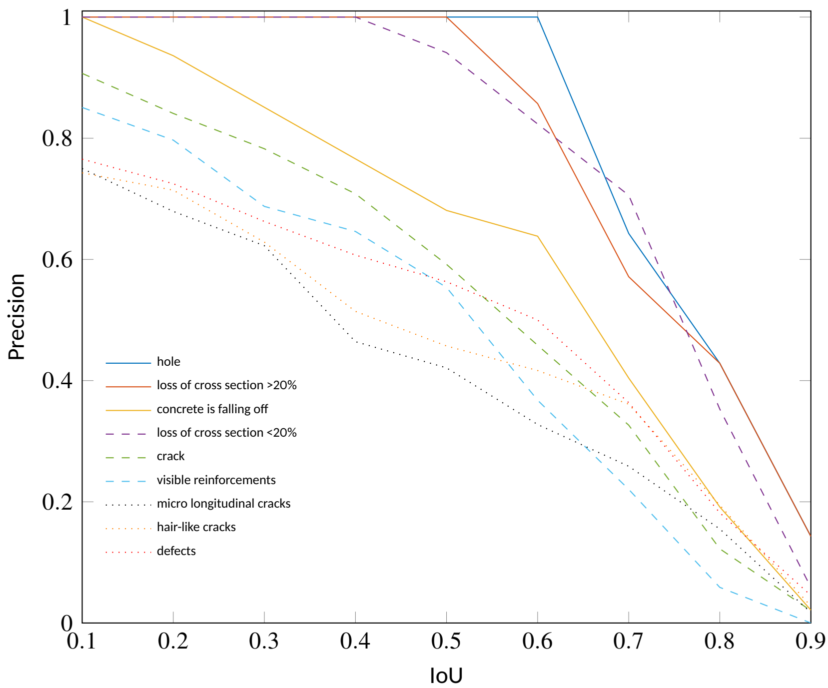


(a) Precision vs IoU

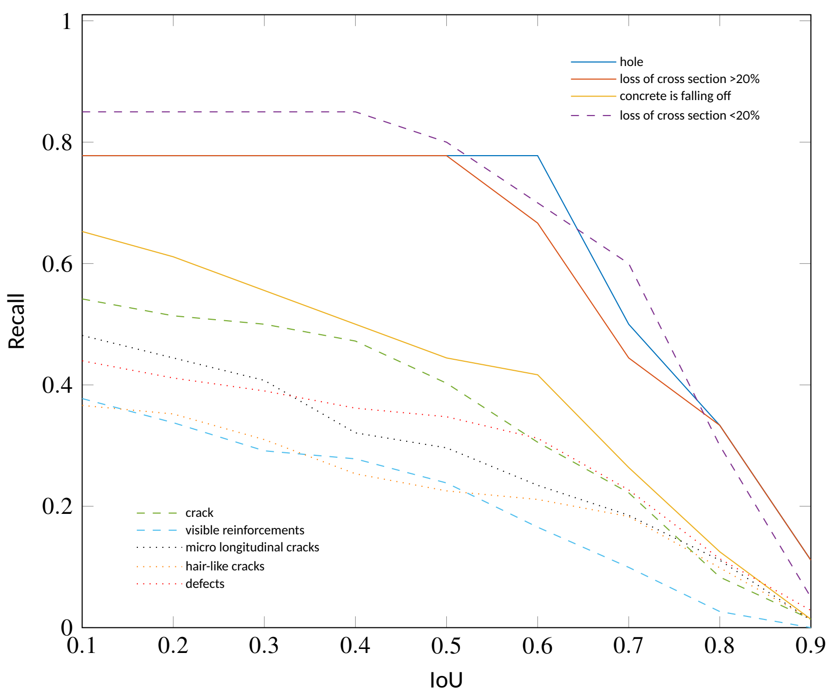


(b) Recall vs IoU

Figure 31: Precision vs IoU (a) and Recall vs IoU (b) for the Faster R-CNN detector of the multiple defect detection model.



(a) Precision vs IoU



(b) Recall vs IoU

Figure 32: Precision vs IoU (a) and Recall vs IoU (b) for the SSD detector of the multiple defect detection model.

ground-truth data is inevitable in some cases where visual boundaries between different defects are marginal. For example, hair-like cracks and micro-longitudinal cracks may be extremely similar in some cases as one defect progresses to another.

HI confusion matrices are generated by comparing the ground truths and detected defects for each asset using under- and overestimation metrics. That is done for each network and the results are presented as Table 14 for the Faster R-CNN, Table 15 for the SSD and Table 16 for the YOLOv2.

Table 14: The confusion matrix of the faster R-CNN multiple defect detection model. [79]

		Estimated HI				
		HI 1	**HI 2**	**HI 3**	**HI 4**	**HI 5**
Actual HI	**HI 1**	1	5	7	4	13
	HI 2	1	5	1	7	5
	HI 3	0	1	9	3	23
	HI 4	1	0	0	14	11
	HI 5	0	0	2	4	83

Table 15: The confusion Matrix of the SSD multiple defect detection model. [79]

		Estimated HI				
		HI 1	**HI 2**	**HI 3**	**HI 4**	**HI 5**
Actual HI	**HI 1**	11	3	2	2	12
	HI 2	0	7	2	6	4
	HI 3	4	2	7	8	15
	HI 4	0	1	2	19	4
	HI 5	7	2	2	9	69

Table 16: The confusion matrix of the YOLOv2 multiple defect detection model. [79]

		Estimated HI				
		HI 1	**HI 2**	**HI 3**	**HI 4**	**HI 5**
Actual HI	**HI 1**	19	2	3	0	6
	HI 2	1	13	4	1	0
	HI 3	2	1	23	1	9
	HI 4	0	0	1	22	3
	HI 5	0	0	1	0	88

Table 17 summarizes the overestimation and underestimation scores for each network according to (26). The networks perform better for more serious defects and the networks have a tendency to overestimate instead of underestimate. The first result is expected as the more severe defects, such as holes or concrete is falling off, have more distinguishable contours and visual changes compared to minor defects, such as hair-like cracks. Object detection networks' tendency to overestimate HI values is beneficial in terms of reliability of the grid as more critical values and defects will be highlighted and there is a smaller chance for serious damage to be left unchecked. On the bad side it leads to additional time spent manually double checking-pictures for faults to confirm the presence of a defect. As is evident, the networks have a very low underestimation score which is important to ensure that critical defects will be detected and crucial information is not ignored.

Table 17: The health index overestimation and underestimation of multiple defect detection models.

	Overestimation (%)			**Underestimation (%)**		
	R-CNN	SSD	YOLO	R-CNN	SSD	YOLO
HI 1	96.7	63.3	36.7	0	0	0
HI 2	68.4	63.2	26.3	5.3	0	5.3
HI 3	72.2	63.9	27.8	2.8	16.7	8.3
HI 4	42.3	15.4	11.5	3.8	11.5	3.8
HI 5	0	0	0	6.7	22.5	1.1
Average	55.9	41.2	20.5	3.7	10.1	3.7

To conclude the performance of three different networks, the YOLOv2 detector shows considerably better performance than both the RCNN and SSD detectors with Precision and Recall over 85% for all categories of objects. YOLOv2 has good performance when the results are mapped to the HIs as it underestimates HI4 defects in less than 4% of cases and HI5 defects in less than 2% of cases. This can be attributed to YOLOv2 scanning the entire image as opposed to regions as done by other networks, therefore extracting more contextual information for each bounding box prediction. The best performance of YOLOv2 can be also explained by the convolutional backbone architecture that was pre-trained on higher resolution images from ImageNet. This results in as the pre-trained weights being more sensitive to capturing fine-grained information such as ill-defined defect edges and incipient fault conditions such as micro-cracks.

3.4 Detection of Multiple Defects Using High-altitude Fly-by Images

Section 3.2 and Section 3.3 were trained and tested on using close-up images of OHL towers. They have a good performance on images taken from close range, but are unable to detect defects from high-altitude fly-by images as the relative size of the defect is very small compared to the size of the original image. In order to increase the applicability of the object detection methodology, a novel approach is proposed that uses images with resolutions of up to 100 MP (11 680$\times$8 708 pixels) and higher as input. At these large resolutions the computational cost of object detection is prohibitively high, but as the OHL inspections are done from high altitude due to the LiDAR then they are taken as a by-product where images are taken at multiple angles and various distances.

3.4.1 Multi-Stage Approach

In object detection, the current practice is to prepare all images (training, testing and live) for detection by resizing them to a much smaller resolution to achieve lower computational cost. Resizing the original $>$ 100 MP images to resolutions of 512$\times$512 or 1 024$\times$1 024 pixels as a whole will lose crucial details of the defect and certain defects would not be visible. To overcome the issue, a multi-stage approach is proposed that integrates multiple levels of features, object detections, cropping of relevant information from the original image and calculating the HI of assets. The multi-stage approach is explained in Fig. 33, where the proposed methodology focuses on each level of the model in the different parts of the image. Image-asset correlation module is used from Section 3.3 and HI determination from Section 3.3. For all stages images are down-sampled to size 512$\times$512 pixels and fed to the object detectors. However, the region of interest (RoI) must be cropped from the original image to ensure no loss of artefact details (for the 2nd and 3rd layers). The algorithm creates a map between the relative coordinates, which are the coordinates of the bounding boxes on the original image, and the absolute coordinates, which are the bounding box coordinates on the detected image. The proposed methodology can be separated into three cascaded levels:

1. Level 1: Tower detection layer - Detection and image cropping of transmission OHL towers using RoI.
2. Level 2: Component detection layer - Disaggregation of transmission OHL into the components using RoI.
3. Level 3: Defect detection layer - Insulator and concrete pole defect detection using image patching and ensemble modeling.

Table 18 lists all used object detection models for fly-by images' defect detection and number of detection classes for each model.

Table 18: Object detection models used in the multi-layered approach.

Object detection model	Number of classes	Number of used models
Tower detection	1	3
Component detection	3	3
Reinforced-concrete defect detection	3	3
Insulator defects detection	1	3

3.4.2 Image Disaggregation and Defect Detection

Bounding box of an object detection model consists of four coordinates: x_{min}, x_{max}, y_{min} and y_{max} where x_{min} and x_{max} are linked with the width of the image and y_{min} and y_{max} with the height of the image. In Fig. 34 $x_{min_{detection}}$, $x_{max_{detection}}$, $y_{min_{detection}}$ and $y_{max_{detection}}$ are absolute coordinates of that bounding box respectively 100, 200, 300 and 400 pixels. As the image itself is 512×512 pixels then $width_{resized} = 512$ and $height_{resized} = 512$.

To calculate relative coordinates of the bounding box (27) to (34) are used. Absolute coordinates in Fig. 34 result in relative coordinates, respectively 0.1953125, 0.5859375, 0.390625 and 0.78125. Assuming the original image is 1500×2000 pixels then the bounding box will be displayed on the original image with coordinates $x_{min_{original}} = 390.625$, $x_{max_{original}} = 878.90625$, $y_{min_{original}} = 781.25$ and $y_{max_{original}} = 1171.875$.

$$x_{min_{relative}} = \frac{x_{min_{detection}}}{width_{resized}} \tag{27}$$

$$x_{max_{relative}} = \frac{x_{min_{detection}}}{width_{resized}} \tag{28}$$

$$y_{min_{relative}} = \frac{x_{min_{detection}}}{height_{resized}} \tag{29}$$

$$y_{max_{relative}} = \frac{x_{min_{detection}}}{height_{resized}} \tag{30}$$

$$x_{min_{original}} = x_{min_{relative}} * width_{original} \tag{31}$$

$$x_{max_{original}} = x_{max_{relative}} * width_{original} \tag{32}$$

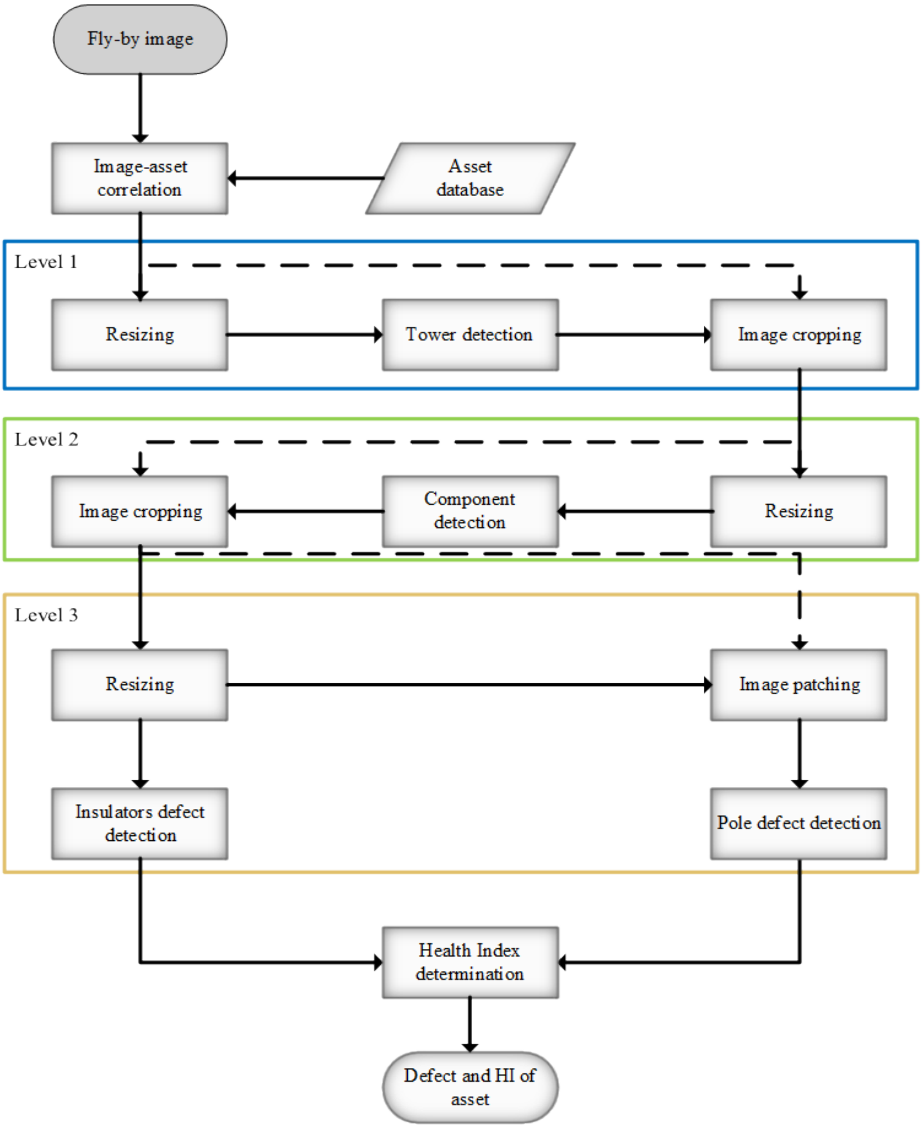


Figure 33: The flowchart of the multi-stage object detection model used for the fly-by image defect detection. Rectangles represent process steps, parallelograms data and round shapes the start and end of the process. Dashes arrows illustrate the flow of original size images.

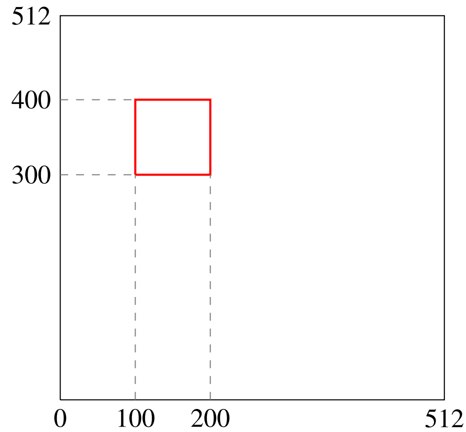


Figure 34: A simplified graph about the relative and absolute coordinates where the red square represents a bounding box.

$$y_{min_{original}} = y_{min_{relative}} * height_{original} \quad (33)$$

$$y_{max_{original}} = y_{max_{relative}} * height_{original} \quad (34)$$

To run the three-level defect detection model on high resolution images successfully, the image's width, height and detected bounding box coordinates with both absolute and relative coordinates are used. Switching from absolute to relative coordinates and back enables to detect towers, components and defects on resized images that are supported by the object detection models and to crop detections based on RoI for further levels using the maximum resolution images. It also enables to display higher level detection on the original images where it is possible to get the full overview with all detections of the tower. The method for describing this multi-stage approach is given in Algorithm 1.

Level 1: Tower detection At this level transmission towers are detected to pass onto the second stage. As there is a large variety of different angles and heights where the image was taken, then for some images the tower is only 10% of the image and in some cases the tower covers almost the whole image. The input into this level is the original images $\mathscr{O}_o$ in the range of sizes up to 11,680$\times$8,708 pixels. The system resizes the image to 512$\times$512 pixels, saves the mapping and inputs the resized image $\mathscr{O}_r$ into stage 1's detector. The output of the detector constructs bounding boxes $\mathscr{B}_r$ around the towers as the RoI. It calculates from the mapping the relative coordinates of the towers for the original image and crops the towers into a new image file $\mathscr{O}_t$ by using modified RoI where the coordinates are transformed. Note that if multiple towers (N_t) are detected on an original image, then multiple cropped images from the multiple detections are saved.

Level 2: Component detection This stage disaggregates the transmission tower $\mathscr{O}_t$ given by Level 1 into insulator strings and the concrete pole. A similar process as in Level 1 applies here, but the result is the detection of individual components. First, the tower images are resized into 512$\times$512 pixel images $\mathscr{O}_c$, and these are fed into the level 2 multi-class object detector. The output of the detector gives the bounding boxes of the insulator strings

Algorithm 1: Transmission overhead line disaggregation into components and defect detection.

Require : Original Image for Processing, $\mathscr{O}_o$
Return : Defect Bounding Boxes and Health Index of System

$[\mathscr{O}_r, \mathscr{S}_o] \leftarrow$ **Function** *Resize(image:$\mathscr{O}_o$)*
  if *($\mathscr{O}_o$ Height > 512 OR $\mathscr{O}_o$ Width > 512)* **then**
    Resize Image to 512×512 pixels $\rightarrow \mathscr{O}_r$
    Set $\mathscr{S}_o \leftarrow$ Scaling Factors
  end
  return: $\mathscr{S}_o, \mathscr{O}_r$;
end
$[\mathscr{B}_r, N_t] \leftarrow$ **Function** *TowerDet(image:$\mathscr{O}_r$)*
  return: array of tower Bounding Boxes in relative coordinates, $\mathscr{B}_r$, & no. of towers N_t; Level 1 Detection
end
for $i = 0$ *to* N_t **do**
  $\mathscr{O}_t \leftarrow$ **Function** *Scale&Crop(Bounding Boxes: $\mathscr{B}_r(i)$, factor: $\mathscr{S}_o$, image: $\mathscr{O}_o$)*
    Crop tower from $\mathscr{O}_o$ using the relative coordinates $\mathscr{B}_r(i)$ and scaling factor $\mathscr{S}$
    return: Cropped tower Image $\mathscr{O}_t$;
  end
  $[\mathscr{O}_c, \mathscr{S}_t)] \leftarrow$ *Resize(image: $\mathscr{O}_t$)*
  $[\mathscr{B}_{cr}, \mathscr{O}_{cs}, \mathscr{O}_{ct}] \leftarrow$ **Function** *ComponentDet(image:$\mathscr{O}_c$)*
    return: component Bounding Boxes in relative coordinates, $\mathscr{B}_{cr}$, & component tower cropped image, $\mathscr{O}_{ct}$, component string cropped images, $\mathscr{O}_{cs}$; Level 2 Detection
  end
  for $x = 0$ *to* $\mathscr{O}_{cs}$ **do**
    $[\mathscr{O}_c, \mathscr{S}_t)] \leftarrow$ *Resize(image: $\mathscr{O}_{cs}$)*
    $[\mathscr{C}_d, D_t] \leftarrow$ **Function** *DefDetStr(image:$\mathscr{O}_c$)*
      return: Defect Class, $\mathscr{C}_d$, & total number of defects per component D_t;
      for $y = 0$ *to* D_t **do**
        $H(y) \leftarrow$ **Function** *HI(class:$\mathscr{C}_d(y)$)*
          return: Health Index $H(y)$;
        end
      end
    end
  end
  for $x = 0$ *to* $\mathscr{O}_{ct}$ **do**
    $[\mathscr{O}_c, \mathscr{S}_t)] \leftarrow$ *Crop(image: $\mathscr{O}_{ct}$)*
    $[\mathscr{C}_d, D_t] \leftarrow$ **Function** *DefDetTwr(image:$\mathscr{O}_c$)*
      return: Defect Class, $\mathscr{C}_d$, & total number of defects per component D_t;
      for $y = 0$ *to* D_t **do**
        $H(y) \leftarrow$ **Function** *HI(class:$\mathscr{C}_d(y)$)*
          return: Health Index $H(y)$;
        end
      end
    end
  end
  return: $\max\{H(x) : x = 1 \cdots n\}$;
end

and the concrete pole using RoI, $\mathscr{B}_{cr}$. From the resized mapping these components are cropped images $\mathscr{O}_{cs}, \mathscr{O}_{ct}$ from the Level 1 output images $\mathscr{O}_t$ with the maximum resolution using the modified RoI similarly to Level 1. These components are passed to Level 3.

Level 3: Insulator and transmission tower defect detection Level 3 comprises 2 main computations: the insulator and tower defect detections and transmission tower HI computation. The defect detection is divided into separate defect detection models that are chosen according to Level 2 detection class. The insulator string defect detector $DefDetStr(\cdot)$ is trained to detect missing insulators, and the concrete pole defect detector $DefDetTwr(\cdot)$ is concerned with the most critical structural defects in the reinforced-concrete OHL tower such as cracks, holes and concrete is falling off. The insulator string detector operates similarly to Levels 1 and 2 with the actions of scaling and object detection. The classes $\mathscr{C}_d$ and number of defects D_t are stored for the HI calculation. The concrete pole defect detection is more involved as poles are usually in the ratio of 20:1 when comparing the height and width. Therefore, the tower images cannot be resized to 512×512 pixels as major details would be lost from the image due to the re-scaling and noticeable image stretching. In this case image patching is used to divide the tower into several 512×512 pixel images and passed into the tower defect detector consecutively. Here the class of defects $\mathscr{C}_d$ and number of defects D_t are more important than the bounding boxes. The detected classes of faults for both the insulator string and concrete pole are sent to HI determination where the HI of the image and its components is determined.

3.4.3 Data Description

Table 19 shows Level 1 tower detection training and testing data with the used detection classes. A total of 331 images were used to train and test the model. 80% of images were randomly chosen as training and 20% as testing, where all training images were augmented, resulting in total of 526 training images with 554 labels.

Table 19: The level 1 tower detection training and testing data.

Class	Augmented training	Training	Testing
Tower	554	277	67
Images	526	264	67

Table 20 shows Level 2 component detection training and testing data with the used detection classes. Three different detection classes are used as they represent different components of OHL. A total of 689 images were used and split into training and testing as 80% and 20%. Training images were augmented, resulting in total of 1 098 training images with 8 378 labels.

Table 21 shows Level 3 concrete defect detection and Table 22 Level 3 broken insulator detectors training and testing data. A total of 249 images were used in the concrete defect detector training process and 509 images for broken insulator detector. All the data is split into training and testing as 80% and 20%, where training data is augmented. That results in 399 images and 613 labels for concrete defect detection, and 814 images and 882 labels for broken insulator detector.

Table 20: The level 2 component detection training and testing data.

Class	Augmented training	Training	Testing
Insulation	5 566	2 783	727
Concrete pole	904	452	112
Steel structure	1 908	954	231
Total	8 378	4 189	1 070
Images	1 098	550	139

Table 21: The level 3 concrete defect detection training and testing data.

Class	Augmented training	Training	Testing
Concrete is falling off	102	52	14
Crack	222	111	23
Hole	289	146	36
Total	613	309	73
Images	399	199	50

Table 22: Level 3 broken insulator detection training and testing data.

Class	Augmented training	Training	Testing
Broken insulator	882	441	110
Images	814	407	102

3.4.4 Ensemble Model

To improve the results of the individual networks, an output-based ensemble technique is used to compose a novel ensemble model to exploit the diversity in feature extraction techniques employed by the different networks. The idea behind the ensemble model is depiced in Fig. 35, where precious detectors' results can be combined as they are failing on different images. The increase in confidence levels produces more missing detections for each network, but reduces the amount of false positives. The proposed ensemble model enables to increase precision and recall, as different networks do not always miss the same images, and when combining them there will be more positive detections. To increase the precision and to minimize the amount of false positives, NMS was used to retain only the highest confidence detections among the networks.

3.4.5 The Performance of Object Detection Models

The performance of tested Level 1 object detection models is presented in Fig. 36, where all networks have nearly perfect performance, except Centernet with 0.8 recall, even at a high IoU threshold. Fig. 37 illustrates the performance of Level 2 models, where recall was almost perfect for all models, but precision was around 0.8 for insulator detection. YOLOv5 performed nearly perfectly, while Centernet had the worst performance. Examination of erroneous predictions shows that they primarily occur in steel structure towers where the

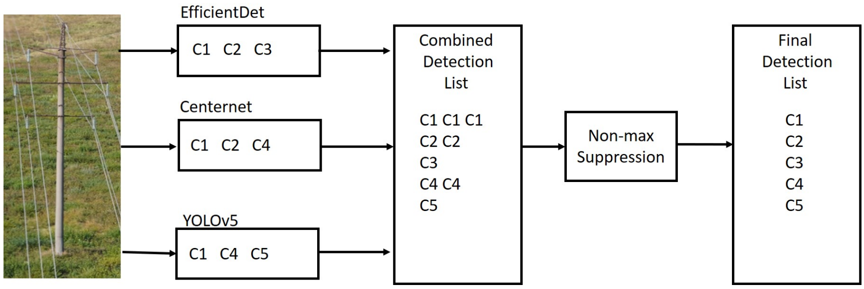
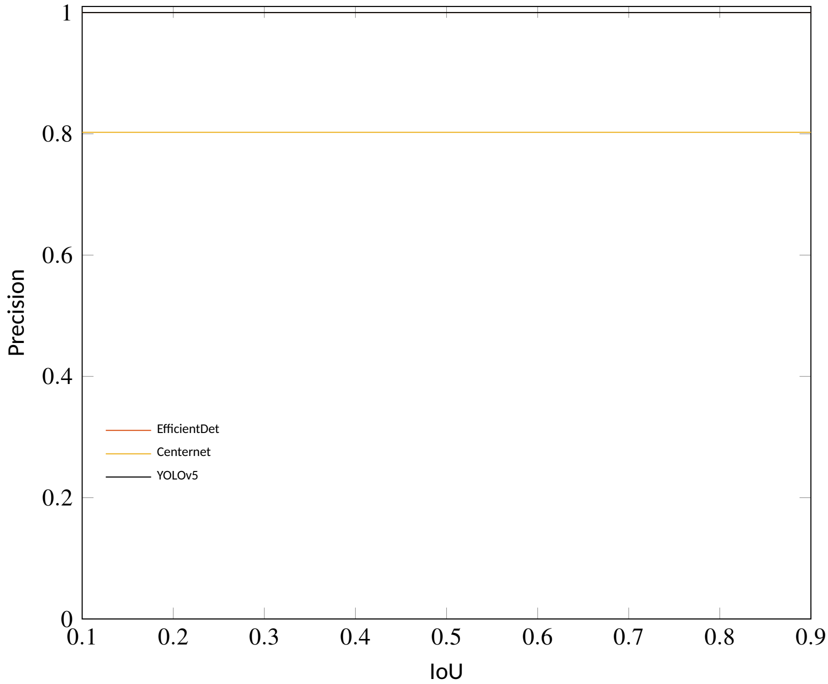


Figure 35: The principle scheme of an ensemble detection using outputs of individual detectors. [17]

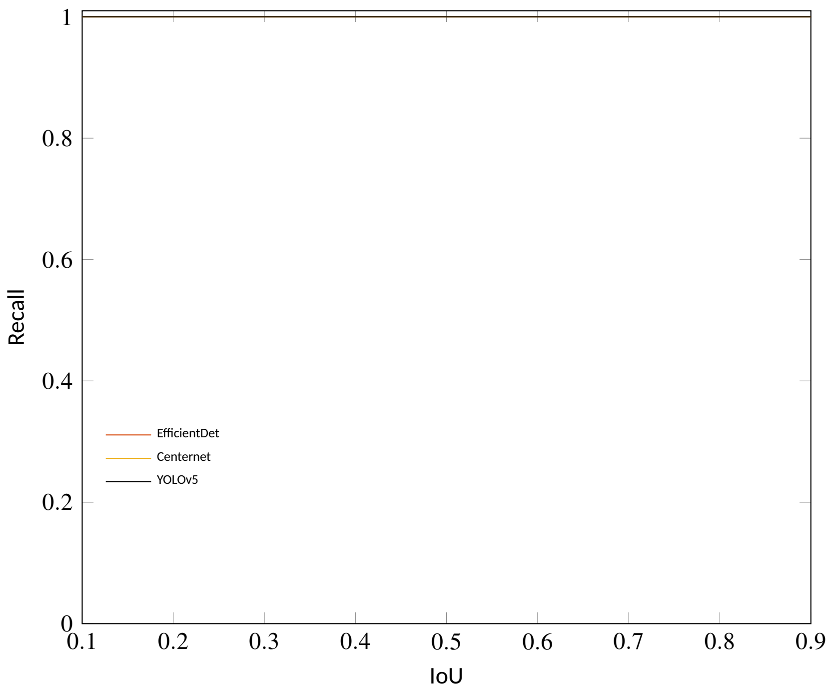
dense lattice image makes detection difficult. As Level 1 and Level 2 are image cropping layers, it is important to use a high IoU threshold to increase the accuracy of lower level models.

The performance of Level 3 models is presented in Fig. 38 and 39, where the performance of concrete defect detection is significantly worse at high IoU threshold than for Level 1, Level 2 and broken insulator detection. Level 3 differs from previous layers as the objects being detected are much less defined, subject to more labelling noise and those detections will not be used for further image cropping. Therefore, IoU and confidence thresholds should be set low for Level 3 models as the precise localization of detections is not required. Since it is important to detect all defects as they may affect OHL condition significantly, recall is more important than precision to minimize the amount of missed detections. A higher amount of false positives will increase the workload of asset managers, but increases the reliability of OHLs.

Fig. 36 to Fig. 39 present that Level 1 detector has nearly perfect performance, while Level 3 concrete pole defect detectors have poor precision and recall on some defects, such as cracks. Fig. 40 illustrates the results of ensemble models for Level 2 and Level 3 detection. An IoU of 0.9 is tested for Level 2 since it is a cropping layer and 0.1 for Level 3 since it is the defect identification layer. Using the ensemble model enabled nearly perfect performance of Level 2 and Level 3 broken insulator detectors and significantly increased the performance of Level 3 concrete defect detector. At low confidence thresholds the results of all detection models improved significantly.

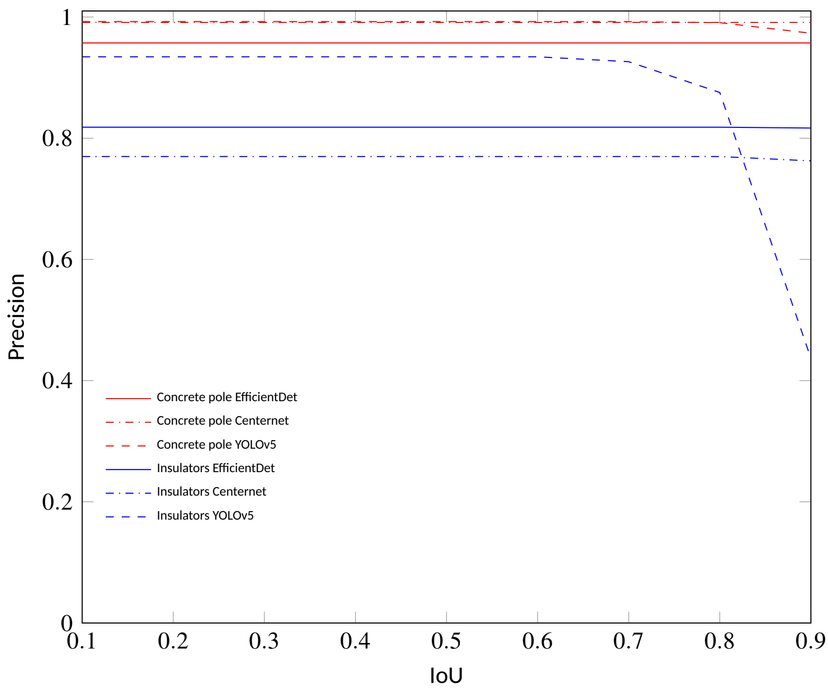


(a) Precision vs IoU

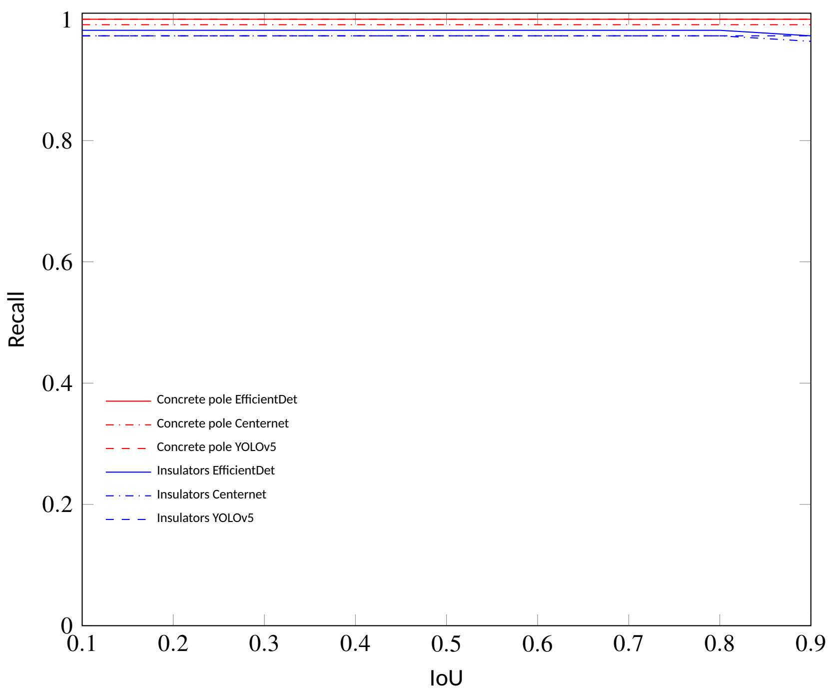


(b) Recall vs IoU

Figure 36: Precision vs IoU (a) and Recall vs IoU (b) for fly-by Level 1 tower detectors.

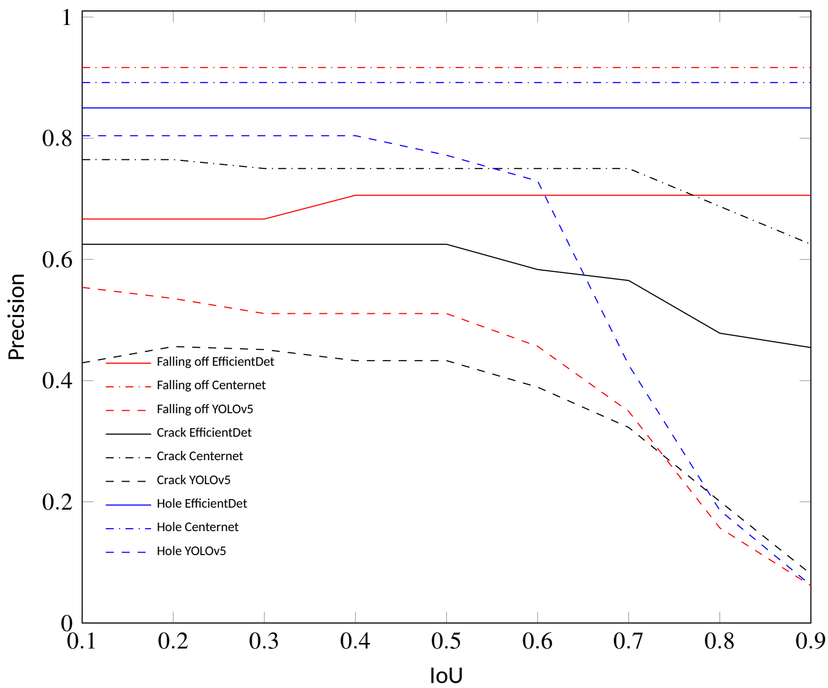


(a) Precision vs IoU

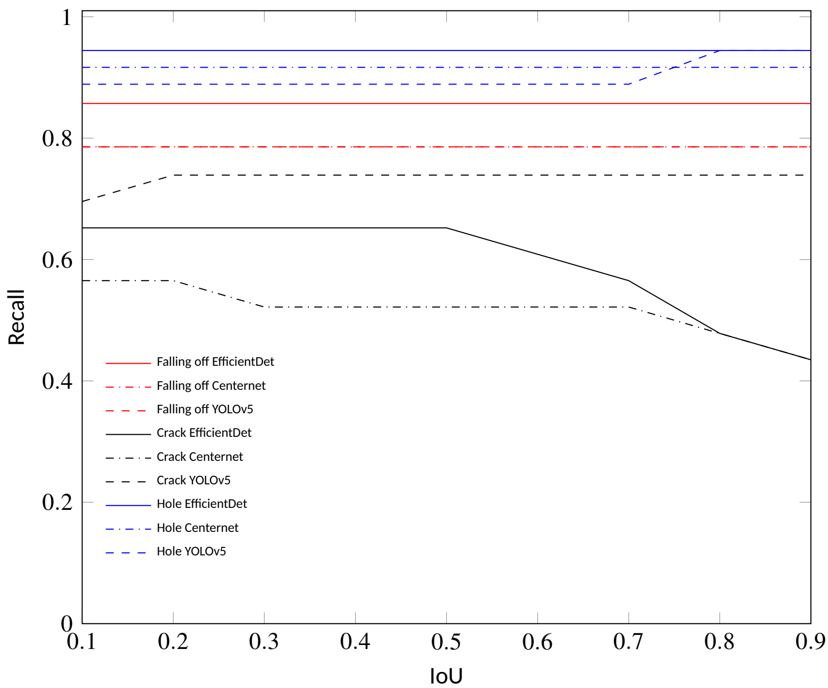


(b) Recall vs IoU

Figure 37: Precision vs IoU (a) and Recall vs IoU (b) for fly-by Level 2 component detectors.

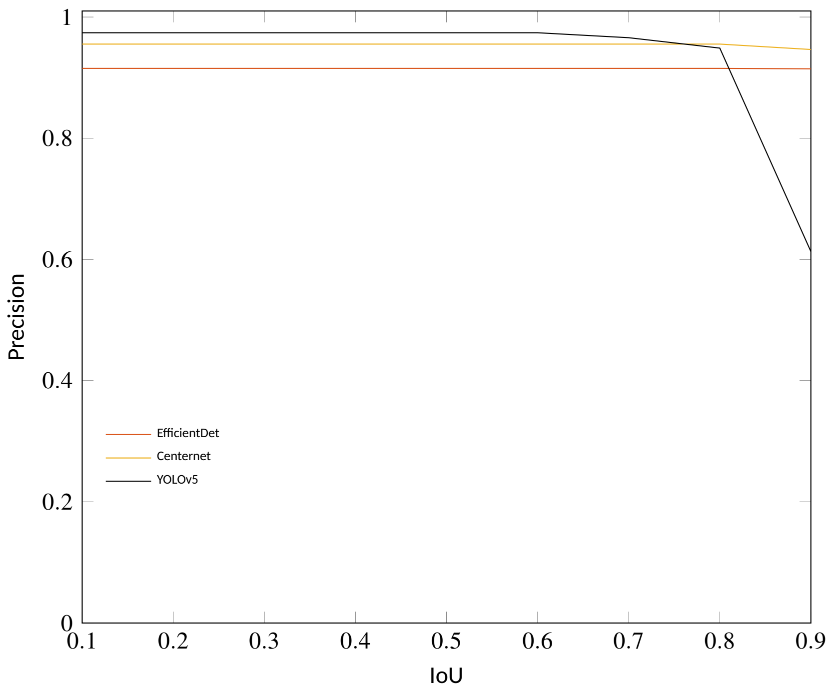


(a) Precision vs IoU

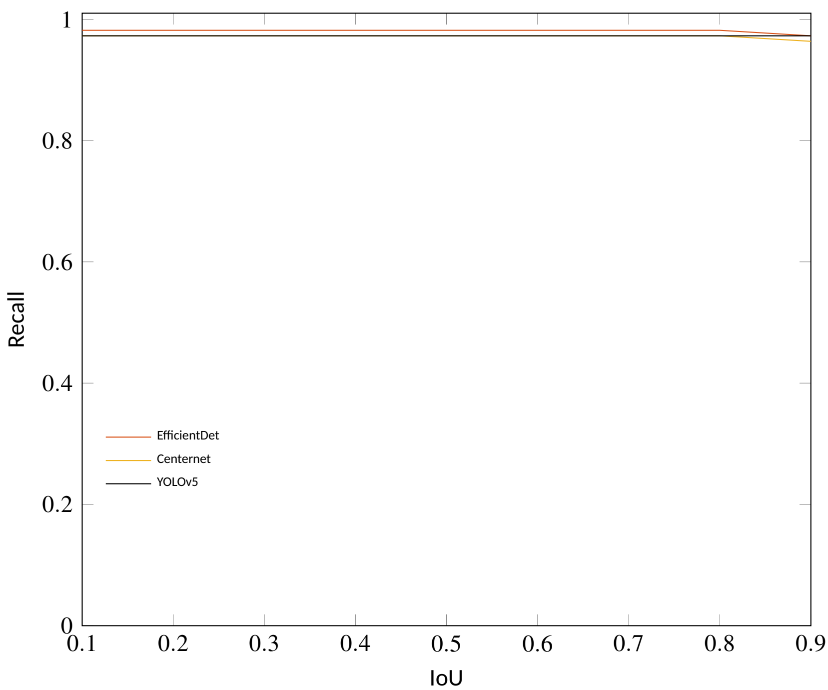


(b) Recall vs IoU

Figure 38: Precision vs IoU (a) and Recall vs IoU (b) for fly-by Level 3 concrete defect detectors.

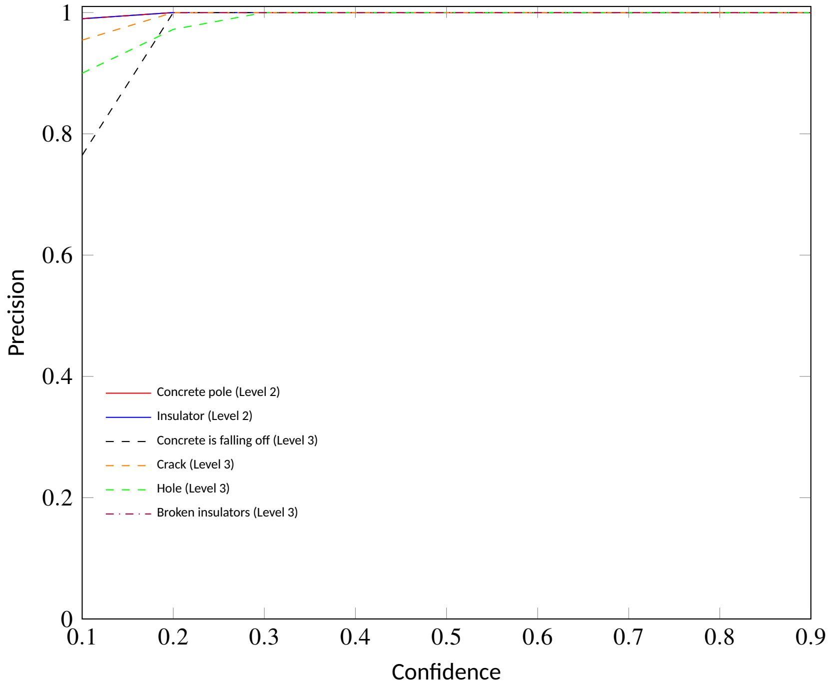


(a) Precision vs IoU

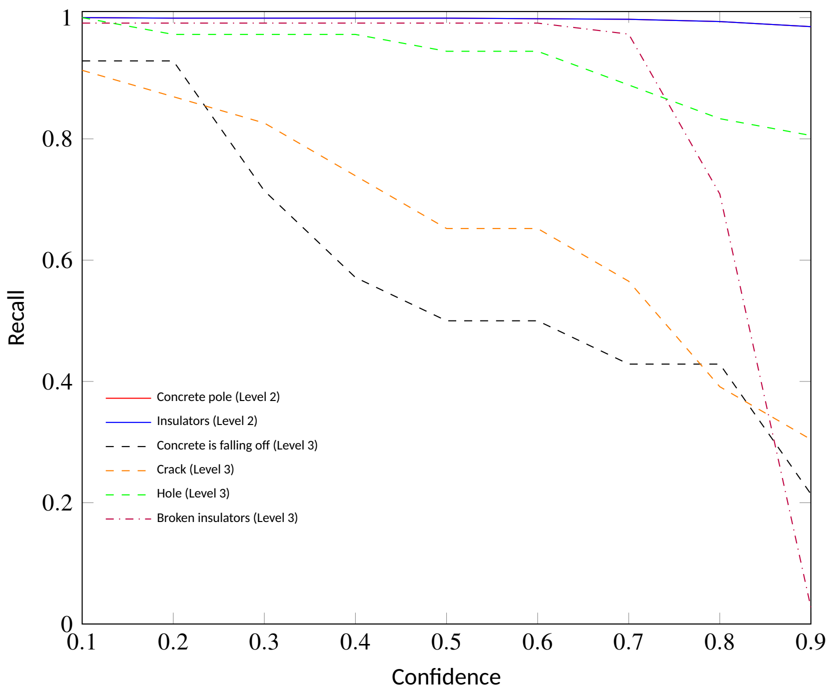


(b) Recall vs IoU

Figure 39: Precision vs IoU (a) and Recall vs IoU (b) for fly-by Level 3 broken insulator detectors.



(a) Precision vs Confidence



(b) Recall vs Confidence

Figure 40: Precision vs Confidence (a) and Recall vs Confidence (b) for ensemble models.

4 The Practical Implementation of the Methodology in a Case Study

In order to illustrate the theoretical methodology proposed in the thesis, a case study using Estonian TSO data is created. The case study starts with the condition assessment of transmission OHLs and ends with the analysis of the proposed asset management decision-making methodology. The case study proposed three individual condition assessment methods of OHLs to determine the HI based on the input data type as described in Section 2 and illustrated in Fig. 13. After the OHL condition assessment, PoF and CoF are determined as done in Section 1.3 and Section 1.4. PoF is calculated based on the asset HI values using survival analysis and historical failure data. CoF is determined by VOLL and direct costs related to the failure elimination.

A brief overview of the data used is presented in Table 23. HI prediction model's training and testing data is collected from visual inspection conducted with the proposed assessment tool in Chapter 2.1. Close-up images from OHL towers and foundations are also collected from visual inspections by using the assessment tool and object detection models are trained as described in Chapter 3.

Table 23: The input data used in the case study.

Explanation	Count
Number of OHLs	200
Number of substation	131
Total number of individual towers	16 823
Number of towers with HI data used to train prediction model	26 206
Number of towers with deleted HI data	67
Number of recorded historical failures	277

The proposed risk-based asset management methodology is compared against the most widespread approaches among TSOs that are TBM and CBM. As there are variations of exact replacement strategies in TBM and CBM then the following approaches are compared in the case study:

- Time-based maintenance (TBM) where assets will be replaced after 50 years of service
- Time-based maintenance (TBM) where assets will be replaced after 60 years of service
- Condition-based maintenance (CBM) where assets will be replaced when reaching HI=5
- Condition-based maintenance (CBM) where assets will be replaced when reaching HI=4 or HI=5
- The proposed risk-based maintenance (RBM) where assets will be replaced according to risk

4.1 Condition Assessment of Transmission Overhead Lines

Condition assessment of transmission OHLs is divided into three different approaches that all have the same aim - to determine the health index of OHL tower and its components. First a HI determination methodology is tested by using specially designed tablet applications combined with visual inspections done by foot patrols. As there is always missing data from inspections then a HI prediction methodology is tested in this case study to predict HIs of missing assets using supervised machine learning algorithms. Automatic HI determination using deep learning techniques on images of OHL towers is performed in Section 4.2.

4.1.1 Health Index Determination Using Visual Inspection

Health index determination using visual inspections methodology was implemented by an Estonian TSO and a pilot project to assess the full grid was successfully done in 2018. Additional full grid inspection using the same methodology was also performed in 2019. Inspections were carried out by foot patrols using specially designed tablet applications where inspectors had only to select noticed defects and HI determination took place in the background of the system as described in Section 2.2 in nearly 100% of the grid. Predefined defect lists consisting of 150 assessment criteria divided between nine component classes enabled to minimize the subjectivity in the assessment process. With additional images and comments it produced 186 rows of data for nearly 17 000 towers resulting in almost 3.1 million columns of new data relating to OHLs. Combined, this allowed to compare different towers and OHLs on the same basis and highlighted OHLs and individual towers that are in a bad technical condition.

Fig. 41 depicts the condition assessment results of a single OHL in the grid. Every tower on that OHL is presented separately with three different component classes: support, foundation, insulation. Fig. 41 shows that this OHL is in relatively bad technical condition as there are eight towers with critical supports and six with critical foundations. That enables to easily see the technical condition of each tower and its components and gives the first hints for the decision-making process. If assuming OHLs' projected lifetime of all components is 50 years then according to 11 it is possible to plan new investments in a certain time-frame. That means replacing or renovating eight towers and six foundations with HI5 will increase the lifetime of OHL by ten years as the remaining lifetime of other towers is at least ten years due to the HI4 and lower. That enables to acquire invaluable information for potential further investments by knowing the exact technical condition on a common basis of each component individually. Therefore it is possible to minimize total expenditures by only renovating those components that are in the worst technical condition.

Fig. 42 presents a comparison of different OHLs according to their technical condition aggregated on an exponential scale. Such an approach does not give an intuitive technical condition of given OHLs but it enables comparison on the same basis where OHLs in the worst condition will be highlighted. In order to look into each OHL in more detail and plan for further maintenance works it is recommended to investigate each OHL on linear scale with all towers as done in Fig. 41.

In addition to showing all OHLs on an exponential scale aggregated to OHL level, it is also possible to present each OHL tower individually using geographical data as done in Fig. 43. That approach requires additional tower coordinates and a geographic information system (GIS), but it enables to plot each tower individually while selecting a single component's HI or aggregated HI as done using (13). A brief glimpse of Fig. 43 will provide a total overview of the grid and provide the opportunity to intuitively highlight all critical towers in the grid.

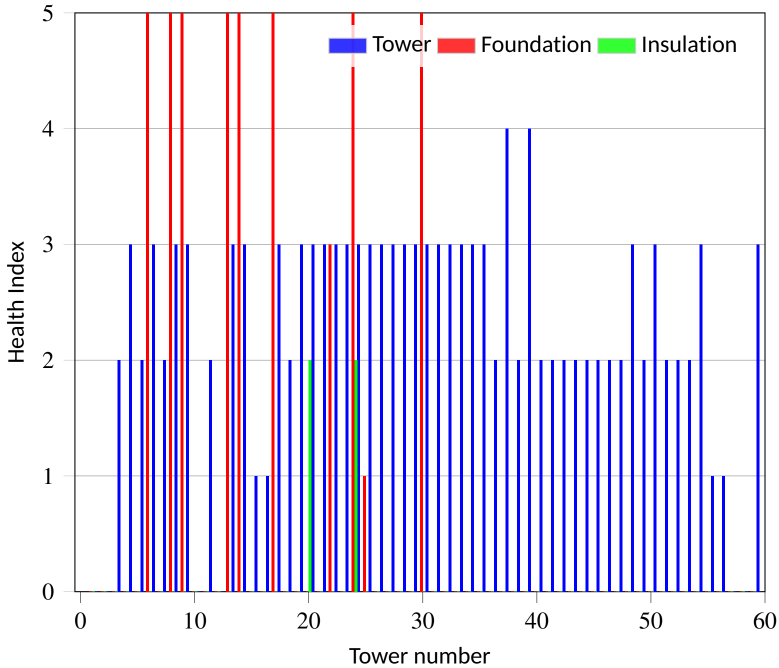


Figure 41: Health index of each tower for a single OHL where foundation (red), towers (blue) and insulators (green) are assessed separately.

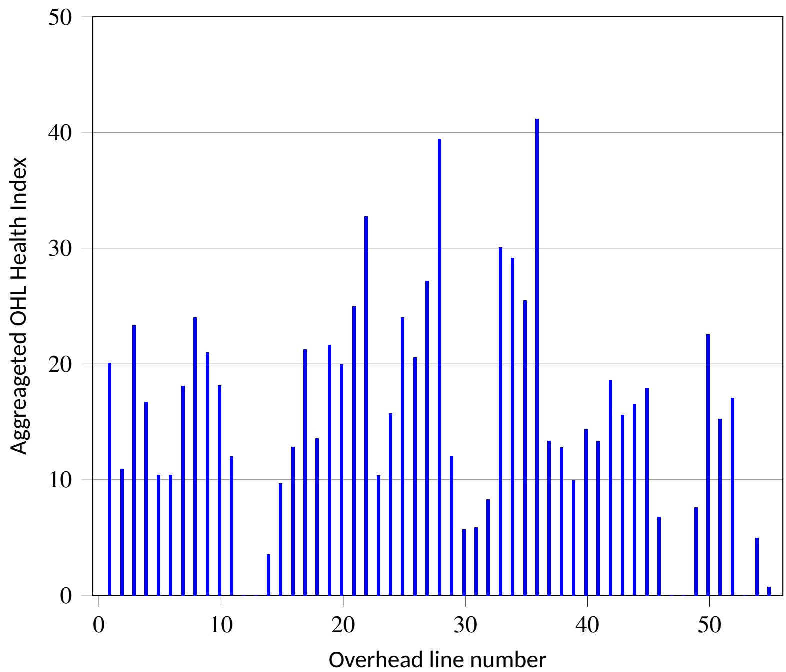


Figure 42: Aggregated health index values of overhead lines.

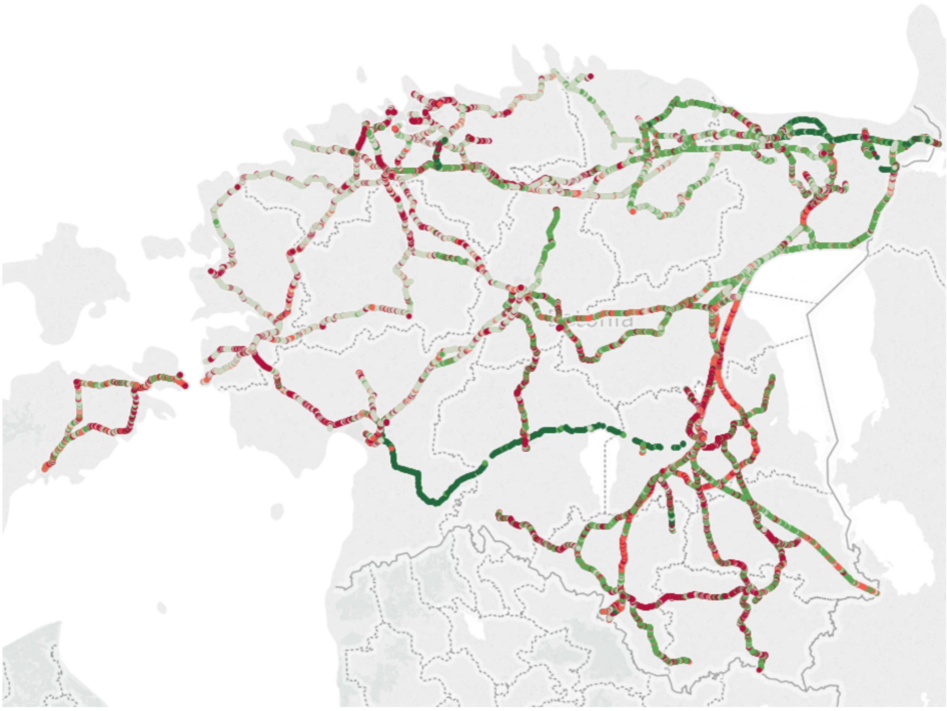


Figure 43: Health index values of each tower in the grid where red represents HI5 and green HI0. HI1 to HI4 are distributed smoothly from green to red.

4.1.2 Health Index Prediction Using Machine Learning

The health index prediction model uses the results of the visual inspection in the previous section as input data to train and validate the prediction model. As these visual inspections produced nearly two million individual defects in OHLs then they were aggregated on the tower level. Nearly all towers in the grid were assessed and a single HI value of each tower was generated on the basis of detected defects according to 11. This resulted in a dataset described in Table 23 with 16 071 individual towers with corresponding HI values. All towers were linked in the asset database to associate the technical features of each tower with HI.

The training data for the HI prediction model is described in Table 24. For the training data 26 206 samples of data of OHL towers were selected along with a single OHL to test the prediction model. All remaining towers in the grid were used to train the prediction model using the proposed methodology. To test a proposed methodology a single OHL that had HI values from visual inspections was selected and all data about HIs was deleted. This resulted in a OHL without HI data consisting of 67 individual towers while all technical features remained unchanged. The selected OHL for testing had a variety of different tower types and installation dates as it had been constructed and renovated in sections and in different timeframes. The age of these OHL towers is in the range of 13 to 61 years, and ten different tower configurations are used. The asset HI prediction model is tested using a Random Forest algorithm with optimal hyper-parameters using unmodified data as it provided the best results for the model validation in Section 2.3.

The results from health index prediction model compared with the results from actual visual inspections are presented in Fig. 44, where it can clearly be seen that 80% of OHL towers have identical HI with actual values. Based on those results it is possible to conclude

Table 24: *Training and testing data of HI prediction model for a case study. [81]*

Health Index Class	Selected OHL	Training
HI0	15	7 757
HI1	0	2 249
HI2	2	11 203
HI3	32	3 430
HI4	16	1 461
HI5	2	106
Total	**67**	**26206**

the HI prediction model is performing significantly better than classifying towers into six HI classes randomly. The model had an accuracy of around 80% in predicting the HI of OHL towers that the model has not previously seen while randomly classifying towers into six different categories only gives an accuracy of around 16.6%. Out of 67 towers, 54 were predicted correctly by the model and 13 incorrectly.

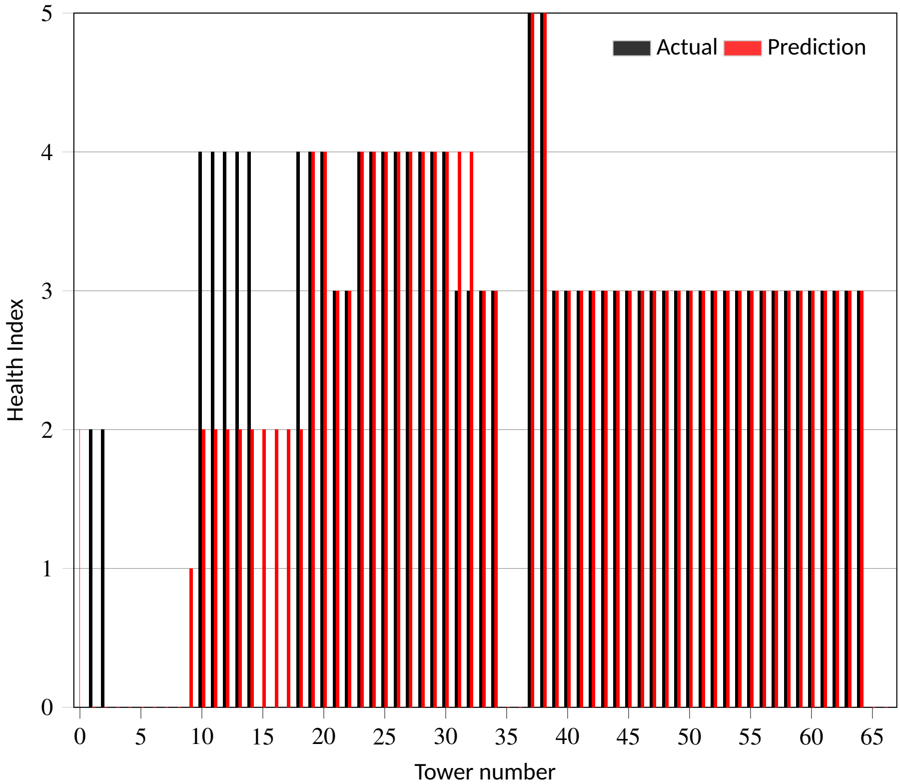


Figure 44: An actual and the predicted health index of the selected overhead line using the random forest prediction model.

4.1.3 Modeling of Aging Behaviour

The aging behaviour model works best on the basis of the HI prediction model where the age parameter will be modified and HIs of each tower will be predicted based on the new age of the tower while all other technical features of assets remain unmodified. To present a clear example of aging behaviour modeling in the grid, then the age parameter of the same OHL with 67 towers in Section 4.1.2 was used, but the age of each tower was increased by ten years. The results are presented in Fig. 45, where it is clearly seen that for nearly all towers HI values increased compared to the results presented in Fig. 44. It is also seen that not a single HI of towers decreased, even the towers with maximum HI value five remained the same. In some cases where towers were in the age of 50 to 67 years, HI was increased more than just a single step as expected according to a linear increase of the HI in the methodology. That can be explained by the prediction process where the model does not just linearly increase the HI of the towers but rather predicts the most probable output based on the data.

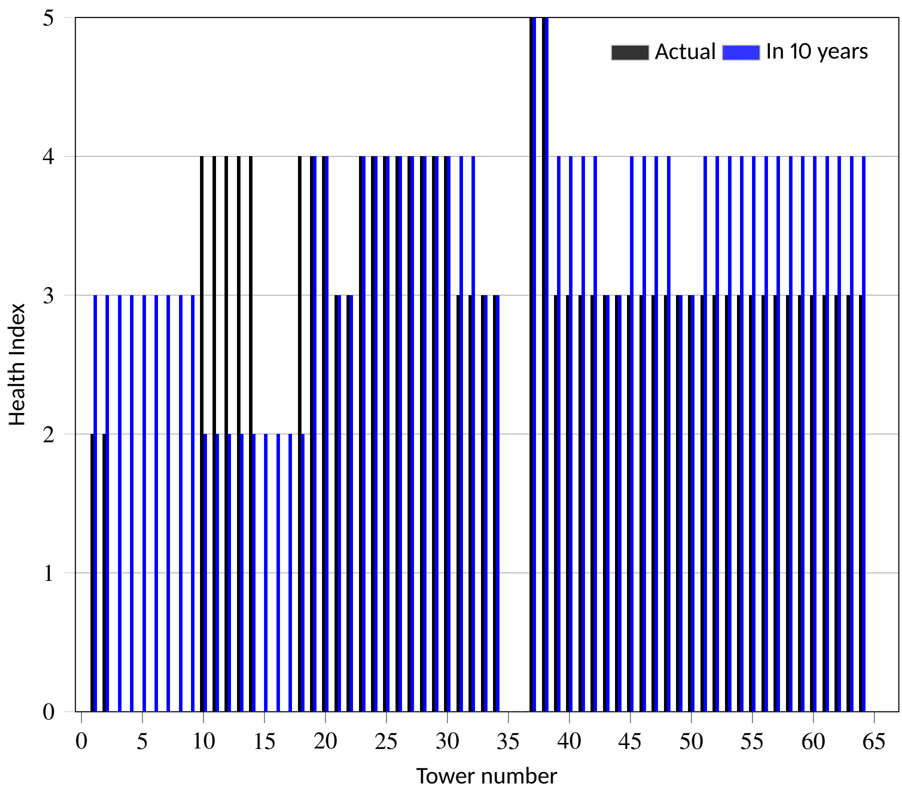


Figure 45: The comparison of an actual and the modelled health index in ten years using the prediction model with Random Forest algorithm.

One part of the TSO's grid is modeled to predict the distribution of HI values in ten years. That situation is presented as a 3D chart in Fig. 46. The number of towers with HI0 and HI2 has decreased while the number of towers with higher HI values such as HI4 and HI5 has increased. In the medium HI range such as HI1 and HI3 the number of towers in those categories has remained almost unchanged. It is clearly seen that the age of the grid has increased and concentrations of assets have moved from lover HI values to higher

HI values such as from HI0 and HI2 to HI4 and HI5. This reflects the situation where the overall condition of the grid has decreased and the reliability of the grid has decreased as there are more towers with critical or end-of-life than before.

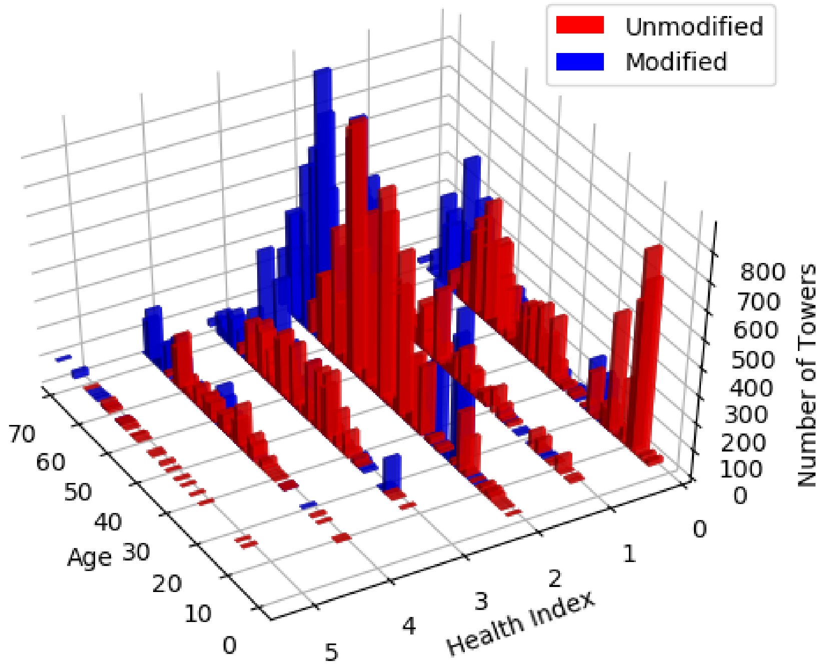


Figure 46: The 3D chart of health index and age distribution according to the aging modeling. [81]

The aging behaviour modeling reflects a realistic situation where no investments are done in the grid. Due to the insufficient data about old assets, there are limitations for the implementation of that methodology where towers with an old age are difficult to model accurately as there is no data about assets of that age. As there are always limits to obtaining a sufficient amount of data and there are impurities in the input data for the grid, this case works well on large datasets where there are many samples from different HI and age ranges but it should be noted that predicting far into the future will cause inaccuracies due to the lack of reliable data.

4.2 Automatic Condition Assessment using Object Detection

4.2.1 Automatic Image-asset Correlation

Image-asset correlation was performed by using data about all 16 823 individual towers with their geographical coordinates. The minimal distance from each tower to the nearest tower was calculated using (23) where the minimal distance for each tower was saved. The results are visualised in Fig. 47 and in Fig. 48 where the cumulative distribution of the tower-to-tower distances is presented. It is clearly seen that 71% of OHL towers in the grid are more than 100 meters from their nearest neighbour. That value also corresponds to the average span length in 110 kV OHLs. It is also seen that around half of the towers are further than 200 m from each other, which is also in the correlation with 330 kV OHL span lengths. 4% of towers are less than 30 m apart and around 1% of towers are within 20 m

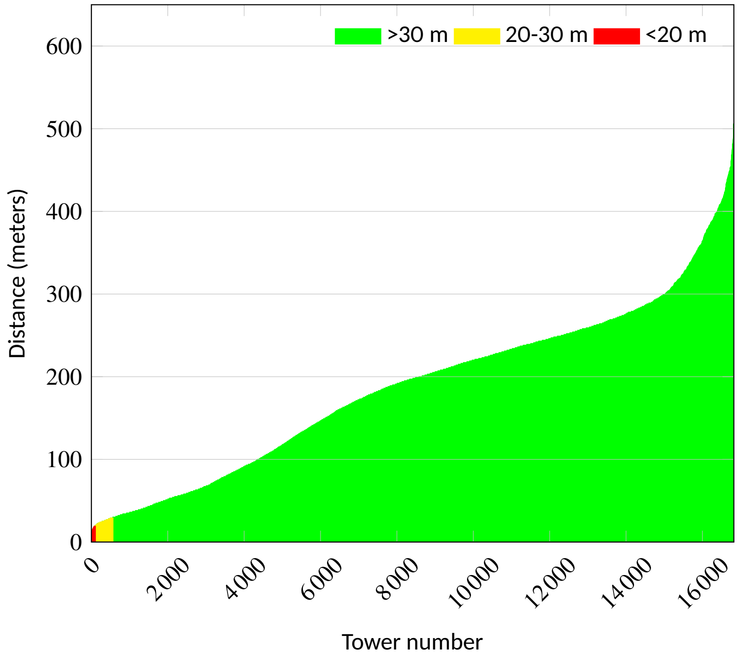


Figure 47: The cumulative distribution of minimal distances between the two nearest towers in the grid.

of the nearest tower. That is well explained by the usage of parallel OHLs in Estonia that are sharing ROW. In some cases the nearest tower is not exactly the tower of the same OHL but rather the tower from parallel OHL.

Next, 1 871 images taken from periodic visual inspections of the transmission grid were used to determine an appropriate value for D_{max}. All images were taken by using the assessment tool described in Section 2.2. Visual inspections were carried out by using tablets that enabled geotaging of images using the A-GPS technology and the geographical location of each tower was saved into the asset database from LiDAR inspections. By knowing the coordinates of each tower and image with the association to asset it was then possible to calculate geographical distance from each image to corresponding asset by using (23). The results of that calculation for each image/asset pair are presented in Fig. 49. 89% of the images were identified to be in the range of 20 meters of the corresponding assets. 82% of images were taken less than 15 meters from the assets and 60% of the images were less than ten meters. As those images were taken during visual inspections then those results are expected because all images were taken next to the tower itself. Around 9% of images were further than 100 meters from the OHL tower and in some cases almost up to five kilometers from the asset. The reason for these large offsets is likely to be due to a GPS drift on tablets or the GPS signal being lost while taking the picture. It must also be noted that image-asset correlation using images taken from visual inspections with tablets using A-GPS can be considered as baseline accuracy when compared to more precise technologies that are widespread with modern drones.

Even using input data from tablets, a threshold value of D_{max} for image-asset correlation in the range of 10 to 15 meters could be used for asset identification with an accuracy of about 90%.

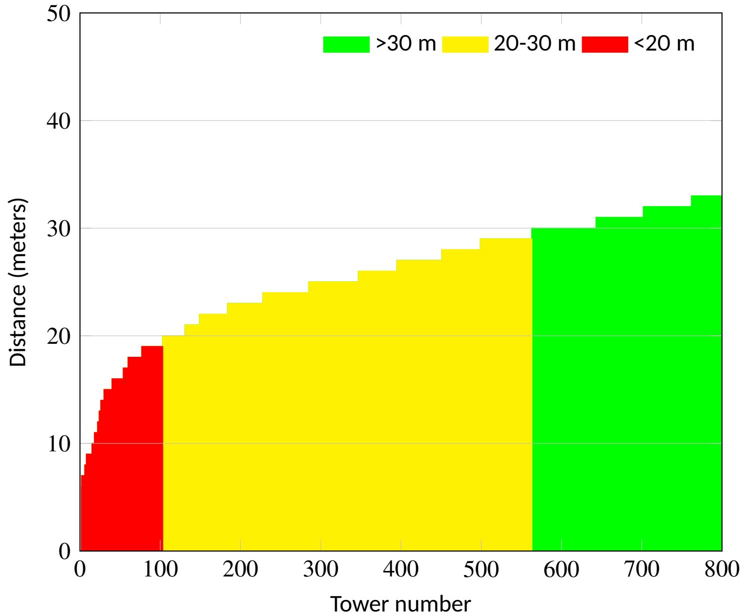


Figure 48: The cumulative distribution of minimal distances between the two nearest towers limited with the closest 800 towers.

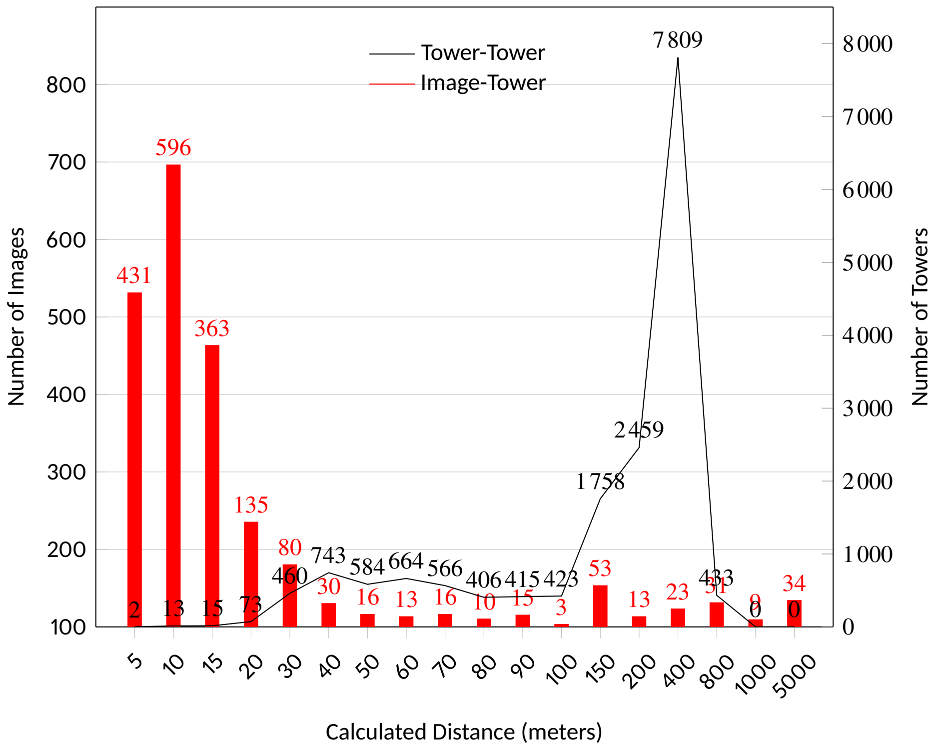
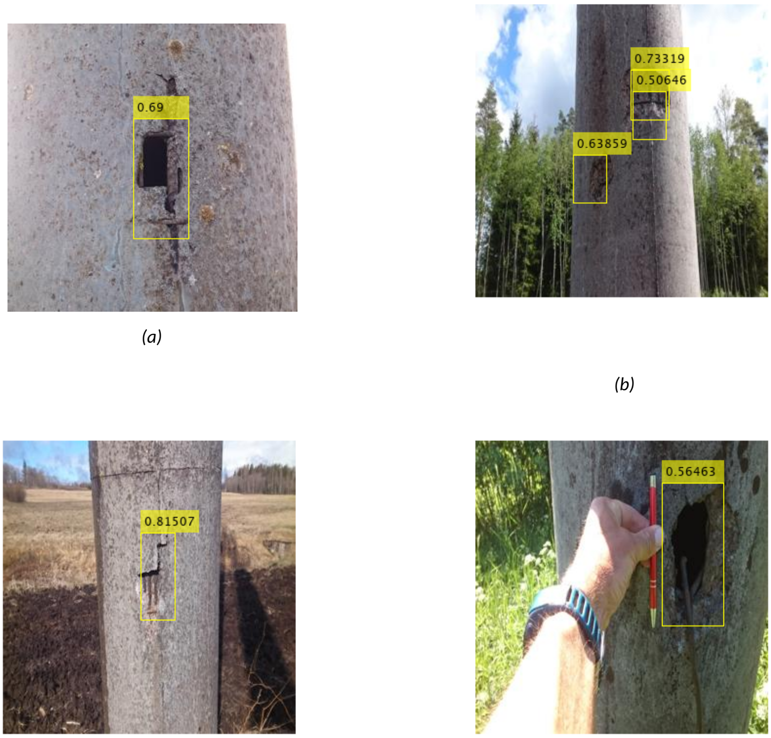


Figure 49: The calculated minimal distance between taken images and towers and between two individual towers.

4.2.2 Detection of a Single Defect Using Close-up Images

Fig. 50 shows example outputs for different backgrounds with the bounding boxes overlaid on the images and their corresponding confidence scores. The threshold confidence score assigned to a positive detection was set as 0.5. The object detection model successfully detected all holes in reinforced concrete poles and similarly across the entire dataset was able to the detect all instances of damage as long as the hole was up close.



(a) (b)

(c) (d)

Figure 50: Detections with different backgrounds of pole (a), forest and sky (b), field (c) and hand in the foreground (d). [124]

When images were taken at a distance, so that the damaged areas were difficult to distinguish due to the image size, several instances of missed detections were noted as in Fig 51a and Fig. 51b. For some cases the confidence threshold for a successful detection was reduced to a very low value such as 0.2, but even then the hole could not be detected although false positive detections began showing up in other images. The occurrence of poor performance for detections of small objects is a recognized issue with the used YOLOv2 algorithm, and would have to be taken into consideration when specifying guidelines for taking images.

Manual inspection of the boxes indicates that higher ground truth labelling noise leads to variations in bounding box sizes or multiple detections of a single damaged area, as presented in Fig. 52. Irregularity in hole shape, backgrounds, shadows and lack of defined boundaries in damaged areas also increase the ground truth labelling noise. Since the

(a) Single hole

(b) Multiple holes

Figure 51: Missed detections of far taken images with single hole (a) and multiple holes (b) where green box represents ground truth. [124]

intersection between the bounding boxes is limited to the area of the smaller box, a small predicted bounding box which perfectly overlaps with a larger ground truth box would have a small IoU even thought it is detected correctly.

Figure 52: Multiple detection of the single damaged area. [124]

If the aim of defect detection model is to highlight all possible critical defects and detect all possible holes then a lower threshold confidence score can be used. It may be desirable but could also increase the number of false positives or multiple detections and flags more defects for maintenance personnel to double-check. Using higher threshold confidence values will decrease false positives but then some holes will be not detected.

4.2.3 Detection of Multiple Defects Using Close-up Images

An example image from the Faster R-CNN object detection model output is presented in Fig. 53 showing the overlays of bounding boxes of the ground truth data (green), and the bounding boxes of detector (red). Three different defects (holes, micro-longitudinal cracks and minor defects) were successfully detected by the Faster R-CNN algorithm for an IoU of 0.1 that were also marked manually as ground truth. The example gives a good indication of a detector's performance to an image with multiple classes of features.

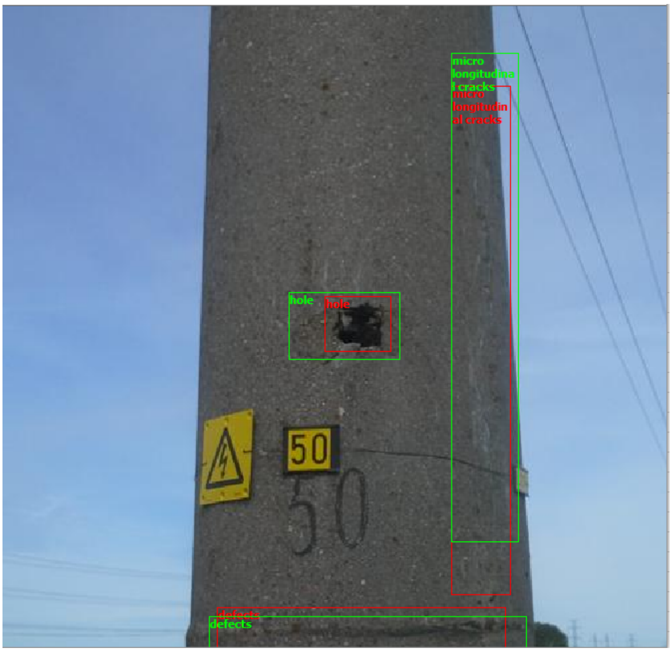


Figure 53: Bounding boxes overlayed on a concrete pole image where green is the ground truth and red is the detection. [79]

Table 25 lists all detected defects from Fig. 53 and links corresponding HI values using Table 3's defects list. Maximum HI of the selected image is calculated according to all detections and by using (10).

Table 25: Detected defects in Fig. 53 using object detection model.

Detected defect	Health Index Class
Holes	5
Micro-longitudinal cracks	2
Minor defects	1
Maximum HI value	**5**

4.2.4 Detection of Multiple Defects Using High-altitude Fly-by Images

Results from the multiple defect detection model using high-resolution fly-by images are presented in multiple steps where each level of detection is explained. Fig. 54 presents a tower detection from high-resolution image where a single tower from the high-resolution image is detected and cropped for further steps. The tower is successfully detected with a confidence level of 100%. After this, the detection bounding box's coordinates are used to crop tower section out of the full-resolution image. There is also another 330 kV OHL tower seen in the image, but as it is not entirely fit, it is correctly not considered as a tower detection. If the proposed model detects more than one OHL tower on the image then all detections will be extracted to Level 2 models.

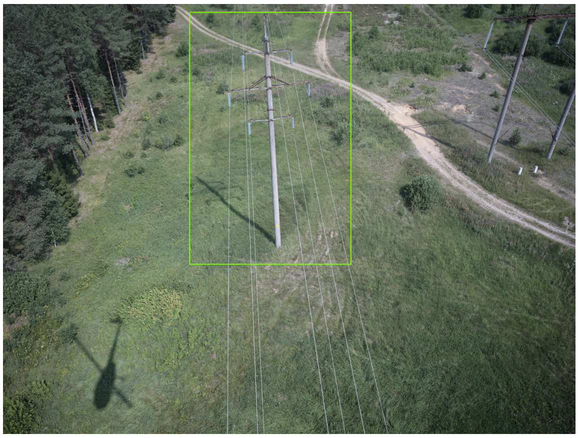


Figure 54: Level 1 tower detection from the input image. Green box is representing detection of the tower.

After successful Level 1 tower detection, Level 2 component detection is used. Fig. 55 presents a component detection where six different insulator strings, concrete pole and three crossbars that are considered as steel structures are detected. Confidence levels of each detection are over 90%, except for the top crossbar that had a confidence level of 68%. By decreasing the confidence level, more detections will appear that might also produce false-positive detections that should be eliminated in Level 1 and Level 2 detections. Each individual detection's labels with bounding box coordinates are used to crop only detected components of full-resolution images to perform Level 3 detections without sacrificing quality. All Level 2 detections are used by component specific Level 3 defect detectors.

Detected components are saved according to bounding box coordinates as individual images for the Level 3 defect detection model. Individual insulator strings are presented as in Fig. 56 where all images are detailed and give a good overview of the condition of glass insulators. There are no broken insulators in the insulator string but two different insulator types are seen.

All Level 3 detections for the reinforced-concrete pole are presented in Fig. 57. Fig. 57d and Fig. 57e are false-positive detections of cracks on the background noise with confidence around 50%. Concrete is falling off and holes are all correctly detected with high confidence levels. In Fig. 57h cracks are detected with 27% confidence and there is doubtful existence of cracks in Fig. 57f. The accuracy of severe and clearly bordered defects, such as holes and concrete falling off, is promising as all detections on the test images are correct. Cracks are a difficult defect to detect because of the variety of the visual appearance of cracks. In this section all different crack types such as micro-longitudinal,

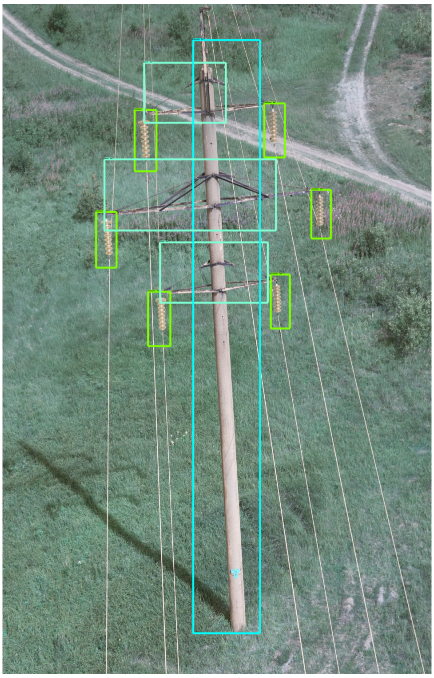


Figure 55: Level 2 component detection from the cropped image according to the Level 1 detection's bounding box coordinates. Green boxes are representing detected insulators, blue concrete pole and cyan steel structures.

hair-like cracks or large cracks are also combined into a single class. A larger amount of training data is required to detect cracks of reinforced concrete poles accurately.

Health index determination of the fly-by image is done according to all detected defects using (10). The results of defect detection model for fly-by image are simplified in Table 26 where all detected defects are counted and HI of the image is determined. False-positive detections are also counted in the HI determination process as the model is producing results automatically. False positives could be minimized by increasing the confidence thresholds of defect detection models but that may result in missed detections.

Table 26: Detected defects from the fly-by image and overall health index of the image.

Detected defect	Number of detections	Health Index Class
Hole	4	5
Concrete is falling off	1	5
Cracks	3	4
Maximum health index		**5**

4.3 Probability of Failure

Probability of failure of each tower is calculated based on the methodology described in Section 1.3. Survival analysis is used to find each tower's PoF according to the asset's HI and historical failures and critical replacements. PoF for OHLs is divided into two categories based on voltage levels where for both voltage levels a separate hazard curve is composed

(a)

(b)

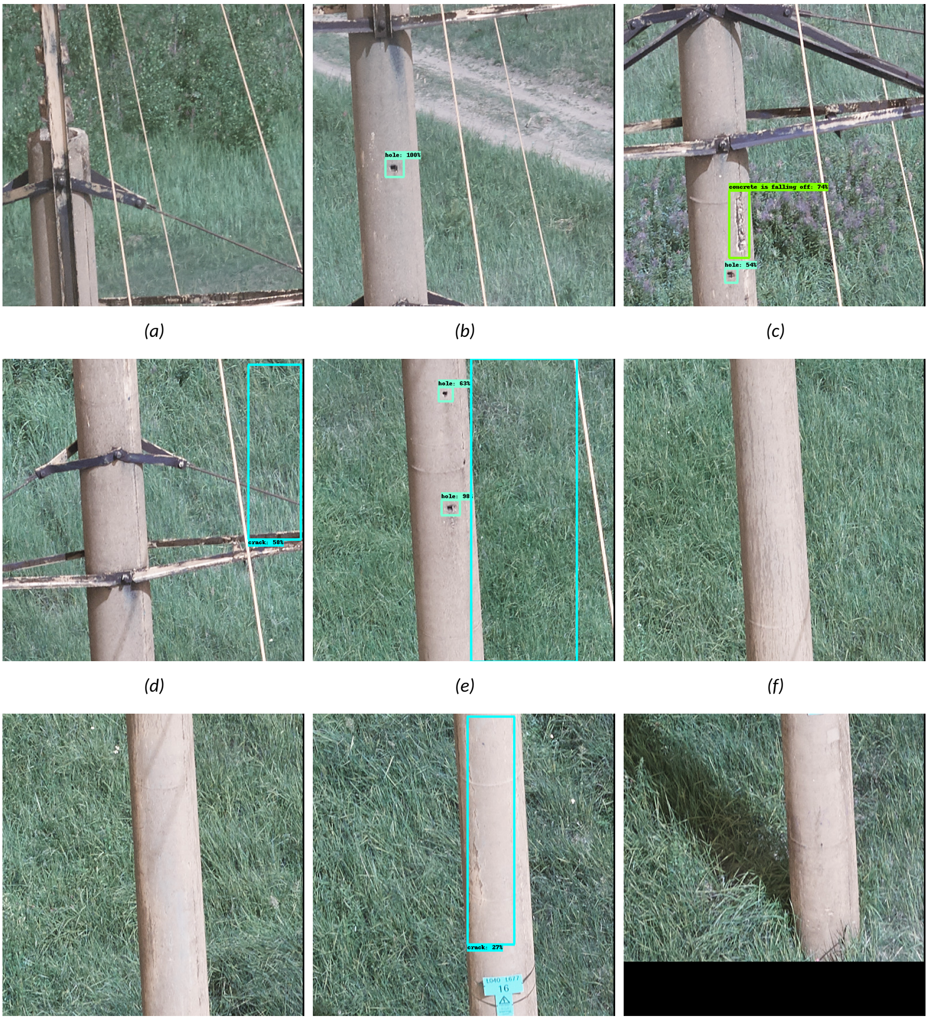
(c)

(d)

(e)

(f)

Figure 56: Cropped insulator images using the Level 2 component detection model with the original resolution image.



(a) *(b)* *(c)*

(d) *(e)* *(f)*

(g) *(h)* *(i)*

Figure 57: Defect detection using the Level 3 reinforced-concrete pole model.

as presented in Fig. 58. It is clearly seen that PoF values for 110 kV towers are higher at all HI levels compared to 330 kV towers. At higher HI values the PoF icreases rapidly where it reaches a maximum value of 0.89 at HI5 for 110 kV and 0.36 for 330 kV towers. PoF for other HI classes is significantly lower, where at HI4 it is respectively 0.08 and 0.02, at HI3 0.02 and 0.03, at HI2 0.007 and 0.001. At low HI classes such as HI1 and HI0 values are almost zero for 330 kV towers and for 110 kV towers they are at HI1 0.002 and at HI0 0.0008. It is also seen that PoF for 330 kV network is lower compared to 110 kV towers at the same HI classes. It is also an expected result as 330 kV grid has higher reliability requirements. These results reflect a realistic situation of extremely reliable assets such as transmission OHLs that are in service for 50 or more years without a single failure. Once these assets reach critical technical condition, they are prone to failures while assets in a good condition will have PoF nearly zero due to the good technical condition. That is also explained from the results by the PoF of HI0 to HI3 where it increases only marginally but once reaching HI4 and especially HI5 then PoF increases rapidly. That can be explained with the exponential nature of degradation processes where mechanical strength decreases rapidly at the end of their lifetime. It is also seen at HI5 that confidence intervals of 95% are increasing significantly by reaching almost 20% of the initial value due to the uncertainties in the data and lack of assets with bad technical condition. This mainly explained with the usage of predictive maintenance in the TSOs grid where the majority of assets will be replaced before reaching their end-of-life condition.

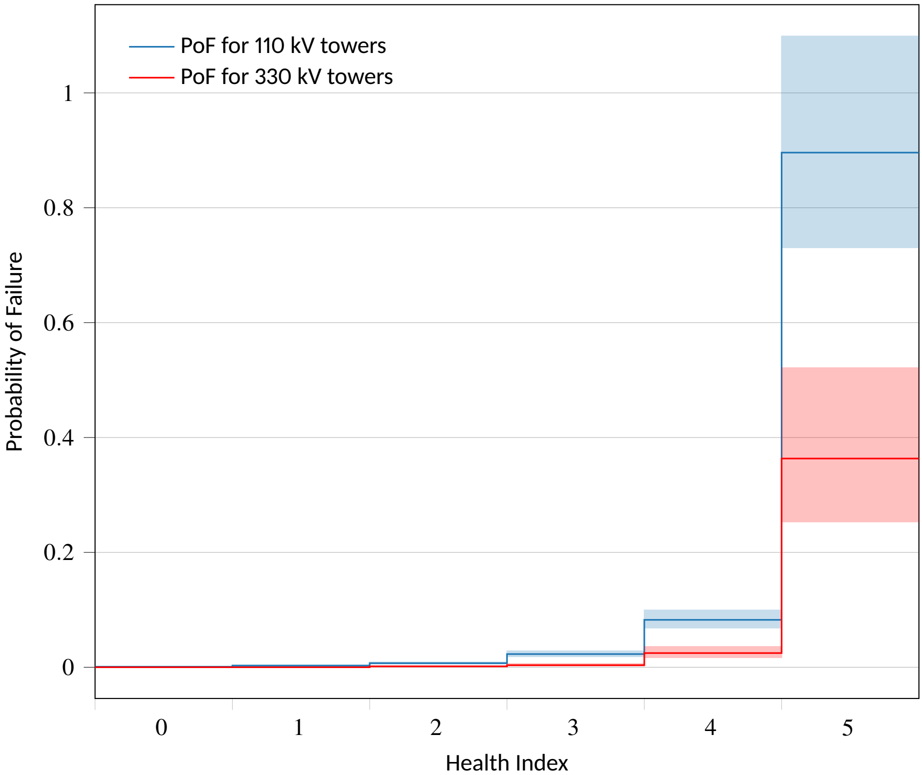


Figure 58: The probability of failure of 110 kV and 330 kV towers based on the health index. 95% confidence intervals are represented with light-colors. [32]

4.4 Consequences of Failure

Consequences of failure are calculated here according to (7) where the expected outage duration is determined individually for each tower. VOLL calculation example of a single substation is done based on a 110 kV substation's real consumption. CoF for each tower is calculated on the basis of replacement of the asset and VOLL. Replacement costs are determined based on the best knowledge from the TSO and values used in this case study for different tower configurations are presented in Table 27.

4.4.1 Estimated Outage Duration

To acquire the estimated outage duration for each tower it is essential to find the nearest distance from roads. To achieve this, each OHL tower's geographical coordinates were compared against the OpenStreetMaps road network and minimal distances from the

Table 27: Expected replacement costs for OHL towers according to tower types and voltage levels used in the case study.

Tower type	Cost of replacement (€)			
	110 kV		330 kV	
	Steel	Concrete	Steel	Concrete
Suspension tower	16 000	8 000	60 000	20 000
Tension tower	16 000	8 000	60 000	20 000

nearest road were saved according to predefined proximity zones. The results of this are presented in Fig. 59.

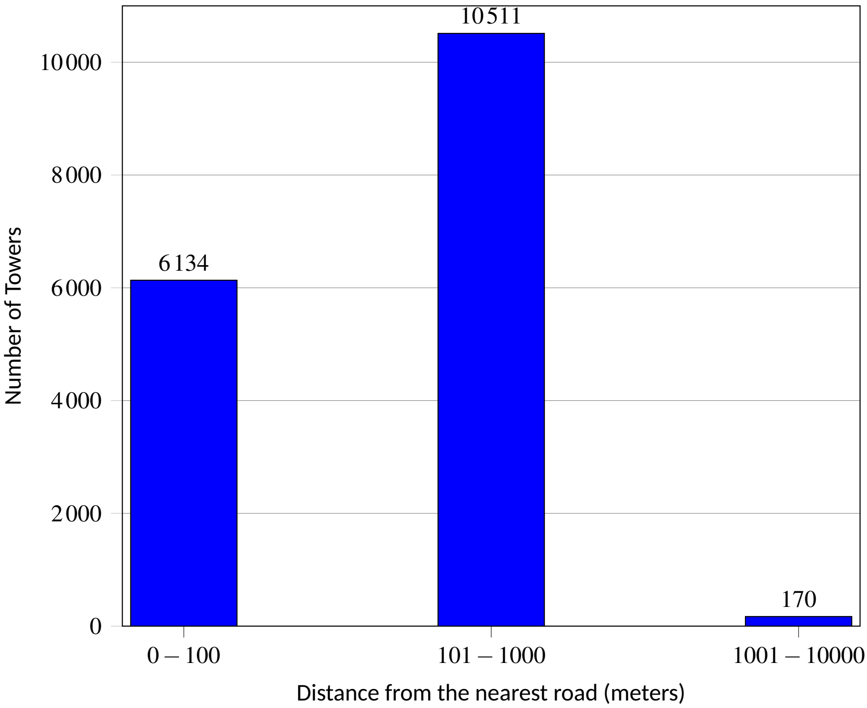


Figure 59: Summarized minimal distances from the nearest roads of each tower according to predefined proximity zones.

Expected outage duration values based on the distance from the nearest road, tower type and voltage level for each tower are presented in Table 2. Realistic results of possible outage durations according to predefined rules are presented in Fig. 60. The estimated outage duration based on the repair times is usually 24 hours for 8 947 towers or 12 hours for 4 829 towers. For 187 towers the estimated outage duration time is longer than 24 hours, 36 hours for 80 towers, 48 hours for 79 towers and 72 hours for eight towers. It seems to be a logical distribution of estimated outage durations as Estonia is well covered with road networks. This approach can be further improved by defining more and smaller proximity zones with additional rules.

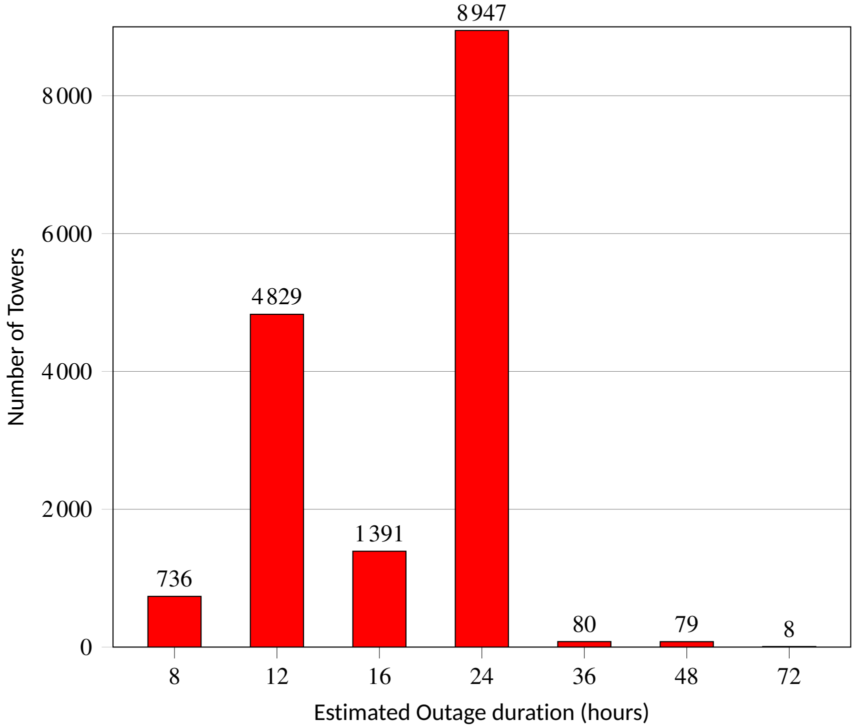


Figure 60: The estimated outage duration of each tower based on the rules from the Table 2.

4.4.2 Value of Lost Load

Distribution of each customer sector is calculated according to annual load distribution between different sectors of final consumers. Customer sectors and each sector's CENS values are acquired from Section 1.4.1. CENS for substations in general is calculated by using (8), where weighted average CENS of the substation is calculated for each substation. An example of a single substation's CENS determination with the distribution of different sectors according to consumption of five years is presented in Table 28. In that substation the majority of consumers are from commercial services, which has the highest cost of CENS per sector. It is also affecting the substation CENS the most, as weighting factor of that sector is almost 83%. There are no agricultural consumers and only a marginal proportion, 0.36%, of industrial consumers connected to the substation.

Table 28: Customer's structure, consumption and CENS values of the selected substation according to Fig. 10 for the estimated 8 hour outage.

Customer sector	Consumption of 5 years (MWh)	Distribution (%)	CENS (€/MWh)
Industry	1 539	0.36	10 890
Commercial services	349 151	82.77	25 210
Agriculture	0	0	13 570
Households	71 156	16.87	17 520
Substation	**421 847**	**100**	**23 861**

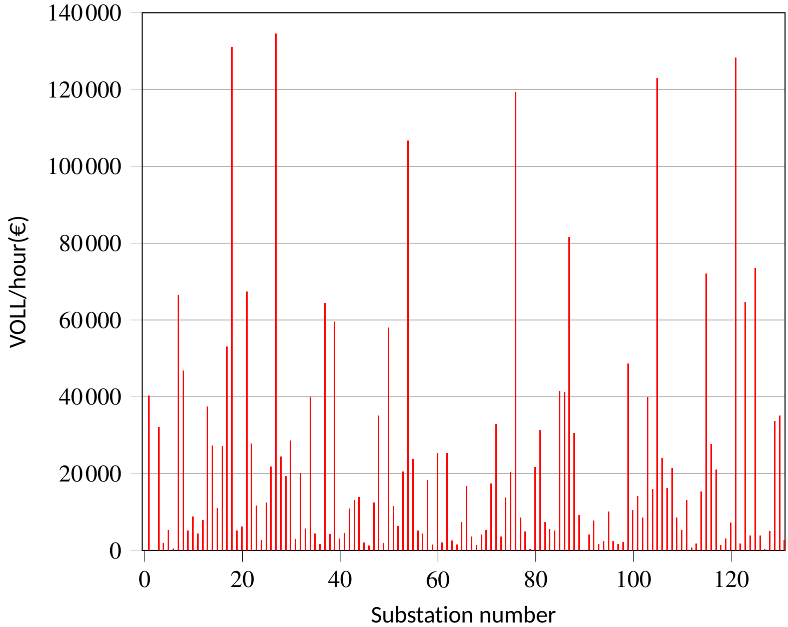


Figure 61: Calculated value of lost load in an hour values for all substations.

Fig. 61 presents all substations with calculated VOLL values in an hour. VOLL in an hour values are used here because the expected outage duration is determined with the tower failure and its repair time. VOLL/hour is calculated according to (8) for all substations with the outage duration of one hour as shown in Table 28 based on five-year consumption data for each sector. The results show large variations between different substations where the maximum VOLL/hour is nearly 140 000 € and minimum ones just around 300 €. This is because there are some substations with large consumption and some substations with nearly no consumption at all. Those VOLL values for substations are further used in combination with estimated outage duration to calculate CoF.

4.5 Risk-based Decision-making

Risk determination results of all individual towers are presented in Fig. 62. It clearly illustrates that the proposed methodology enables to determine the most critical towers in the grid. That enables to move to the efficient asset management decision-making where towers 1 032, 13 854, 15 207 and 15 319 should be focused upon first as they have significantly higher risk values. As the majority of towers have the risk just around a few euros due to the good technical condition and therefore low PoF then some towers have bad technical condition in combination with great CoF leading to risk values exceeding 0.25 M€.

The final results of the proposed risk-based approach are presented in Fig. 63. Different AM strategies are compared in parallel and the proposed RBM achieved the best results in terms of remaining risk and in the lowest total cost of replacements in the grid. As expected, the TBM scenario where all assets were replaced after 50 years of service produced the highest cost of replacement resulting in 161 M€. The TBM approach, where all assets that are older than 60 years were replaced, produced a replacement cost of around 41 M€ while

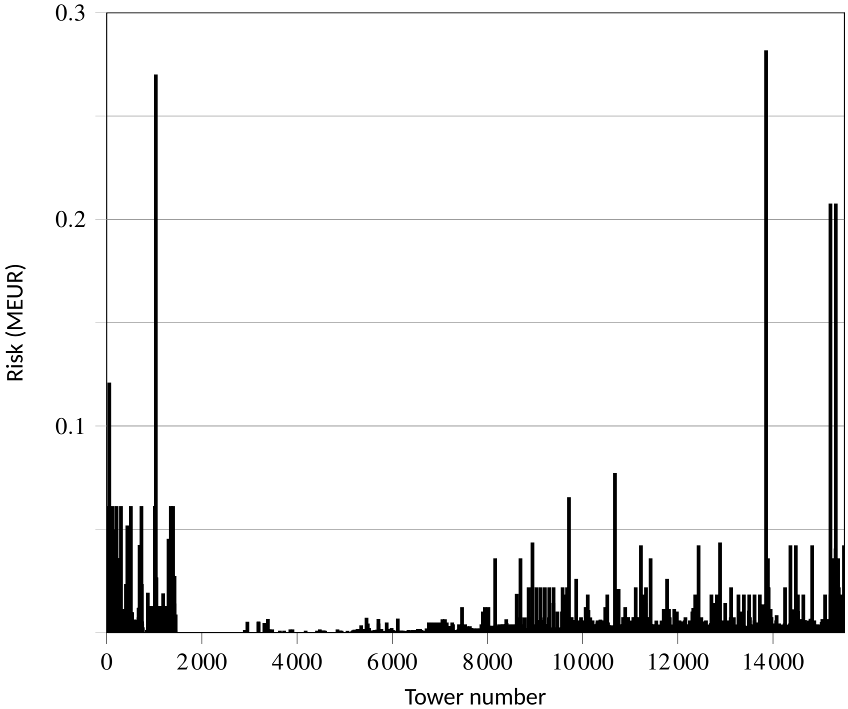


Figure 62: Calculated risk values of each tower in the grid.

the cost of remaining risk remained higher than at the TBM50. The remaining risk of the grid for the TB60 is 8.84M€ and for the TBM50 6.43 M€ when doing no replacements in the grid produced the remaining risk of 8.94 M€. For the CBM scenarios replacement costs are lower compared to the TBM approaches while the cost of remaining risk is also less than in TBM approaches. Replacing all assets that have a higher HI than four requires around 25 M€ compared to the CBM5 with 1.9 M€ where only assets with HI5 will be replaced. That results in remaining risk, respectively 4.9 M€ and 7.75 M€ for the CBM4 and the CBM5. The proposed RBM methodology produces replacement costs comparable with the CBM5 strategy while providing risk reduction in the grid similarly to the CBM4. This clearly illustrates that the proposed RBM approach produces the lowest cost of replacement with the lowest cost of remaining risks as it takes into account both, technical condition of assets and their importance in the transmission system.

4.6 Optimisation

The results of using dynamic knapsack optimisation are presented in Fig. 64. The optimisation gives only a slight benefit compared to selecting the towers with the highest risk. But still, all optimised scenarios from 250 000 € to 1 500 000 € are one percentage point better than not optimised. It must be noted that this kind of knapsack optimisation is really computationally resource expensive and the increase in efficiency is small. However, using more advanced optimisation algorithms will most probably produce similar results with less computational resources.

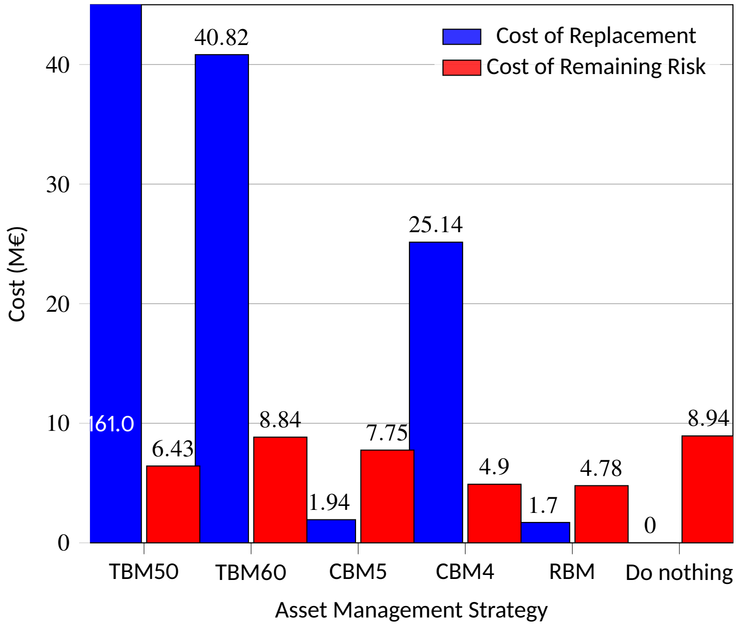


Figure 63: The comparison of different asset management strategies using the cost of total replacement and remaining risk.

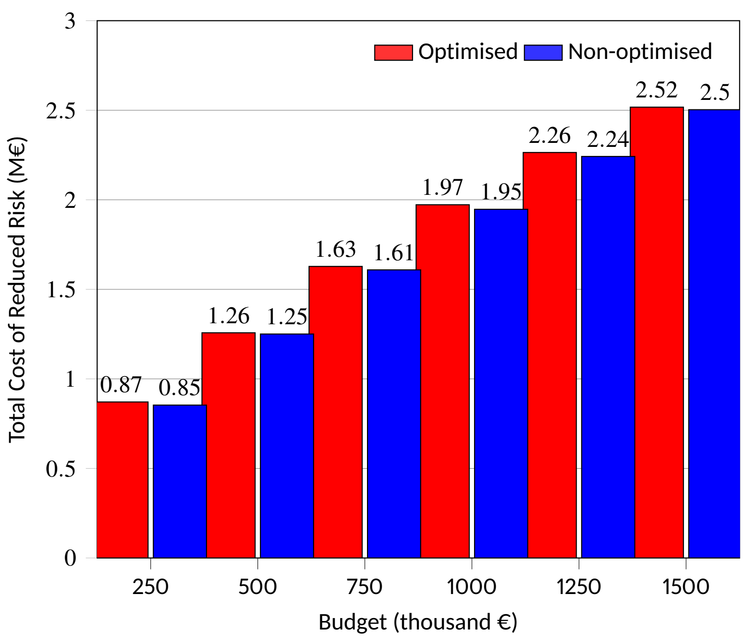


Figure 64: Optimised decision-making under budget restrictions vs non-optimised.

5 Conclusions and Further Work

The thesis presents a complete framework of asset management decision-making methodology for transmission OHLs to increase the efficiency of condition assessment and LCM of OHLs. All the hypotheses of this thesis were thoroughly confirmed and proved to be possible. The proposed methodology substantially increases decision-making efficiency compared to TBM and CBM approaches and clearly determines the most critical towers in the grid. The proposed asset management decision-making methodology is a RBM approach that enables to overcome the main issues related to classical RBM implementations such as the transparency related to input parameters and decision-making. The main drawbacks of classical RBM methodologies are eliminated by using tower-specific PoF determination based on the actual technical condition of assets and CoF based on precise VOLL determination. For the more cost-effective decision making, each tower in the grid is treated separately with tower-specific PoF, CoF and risk. Focusing on a tower level enables a well-argued and transparent asset-management decision-making process, as each element in the grid can be observed separately.

Sophisticated condition assessment methodologies of OHLs were developed, where it was proven that the age alone is an inaccurate parameter for technical condition assessment, given that, when focusing on the towers of the same age, some were in good condition and others in poor condition. The unambiguity of traditional visual inspection was reduced by using predefined visual indicators and specially designed mobile applications to ease the assessment process. The health index prediction methodology of overhead transmission lines using supervised machine learning models demonstrated that it is possible to predict missing HI values of high voltage OHL towers based on the asset's technical features and HI results of already existing data with the accuracy of 80%.

Object detection networks demonstrate a significant potential in the near future to replace traditional visual inspections with aerial inspections as they provided the opportunity to detect the most critical defects from concrete poles while mapping them to a HI. The multi-stage approach shows even greater potential in the future to be implemented as it enabled the possibility to use super-high-resolution images taken during LiDAR data collection as an input to determine the technical condition of OHL towers. The results imply that automated state-of-the-art methodologies such as HI prediction and object detection demonstrated cost-effective and reliable results, that could potentially enable to decrease the workload of maintenance engineers.

Practical Implementations

For the practical implementation of the proposed methodology, it is essential to focus on the collection of input data. The condition information of assets is required for accurate PoF determination together with grid topology such as possible outage combinations of OHLs and division of consumers into groups. The cost of asset replacements and VOLL determination using estimated outage duration of OHL towers are TSO specific, where input parameters are specified using the best practice.

The most practical and also cost-effective approaches to determine the technical condition of OHL elements is by using multi-stage object detection model described in Section 3.4. To acquire the technical condition of assets that were not assessed it is recommended to use HI prediction model as described in Section 2.3. Training and validation process of all models should be done based on the real TSOs' data as images of OHL towers, manual labelling of defects and a large-scale use of technical data about the grid and technical condition of assets.

Further Work

The proposed methodology is developed for transmission OHLs, but it is possible to use the framework for other infrastructure systems as well. To implement the methodology for other infrastructure elements it is essential to develop reliable condition assessment methodologies for the selected assets and analyse the results of possible failures.

The most critical part of the methodology is the condition assessment of asset, which affects the PoF and Risk assessment results significantly. It is possible to develop precise and well analysed condition assessment criteria for all OHL components and to increase the accuracy of the proposed decision-making methodology by focusing on the component level instead of tower level. More parameters could be involved in the CoF determination to also assess the CoF arising from the hazards to safety. However, it must be taken with extreme care, as it may distort results, when using unrealistic parameters.

To increase the accuracy of the HI prediction model, it is possible to involve additional training data of multiple-year inspections. Using the model to predict the HI of each OHL component or even defects, similarly as done for OHL towers, might have a positive effect on the final results. That also enables to move to a more detailed asset management approach where all decisions are done on the component level.

The proposed multi-stage object detection model should be used to increase the efficiency of airborne OHL inspections. Additional Level 3 object detection models should be developed for other OHL components such as steel structures to detect all critical defects of OHLs. It is also important to develop guidelines to take fly-by images of OHLs to cover all sides of assets. Standardized guidelines also enable to increase the accuracy of object detection models by using images taken from similar angles.

List of Figures

List of Tables

References

[1] G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surfaces*. USA: Wiley, 1987.

[2] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, J. Sanchez-Gasca, R. Schulz, A. Stankovic, C. Taylor, and V. Vittal, "Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1922 – 1928, 12 2005.

[3] Y. Makarov, I. Reshetov, A. Stroev, and I. Voropai, "Blackout prevention in the united states, europe, and russia," *Proceedings of the IEEE*, vol. 93, pp. 1942 – 1955, 12 2005.

[4] E. Andreini *et al.*, "Facing the replacement wave," tech. rep., ENTSO-E, 2015.

[5] CIGRE WG B3.03, "TB300 Guidelines to an optimized approach to the renewal of existing air insulated substations," tech. rep., CIGRE, 2006.

[6] ISO 55000:2014(en), "Asset management — Overview, principles and terminology," standard, International Organization for Standardization, Geneva, CH, Jan. 2014.

[7] ISO 55001:2014(en), "Asset management. Management systems. Requirements," standard, International Organization for Standardization, Geneva, CH, Jan. 2014.

[8] I. 55002:2014(en), "Asset management. Management systems. Guidelines for the application of ISO 55001," standard, International Organization for Standardization, Geneva, CH, Jan. 2014.

[9] PAS 55:2008-1, "Specification for the optimized management of physical assets," standard, British Standards Institution, Sept. 2008.

[10] PAS 55:2008-2, "Guidelines for the application of PAS 55-1," standard, British Standards Institution, Sept. 2008.

[11] A. Jahromi, R. Piercy, S. Cress, J. Service, and W. Fan, "An approach to power transformer asset management using health index," *IEEE Electrical Insulation Magazine*, vol. 25, no. 2, pp. 20–34, 2009.

[12] T. Matsui, Y. Nakahara, K. Nishiyama, N. Urabe, and M. Itoh, "Development of remaining life assessment for oil-immersed transformer using structured neural networks," in *2009 ICCAS-SICE*, pp. 1855–1858, 2009.

[13] CIGRE WG A2.49, "TB761 Condition assessment of power transformers," tech. rep., CIGRE, 2019.

[14] F. Yang, M. S. Habibullah, T. Zhang, Z. Xu, P. Lim, and S. Nadarajan, "Health index-based prognostics for remaining useful life predictions in electrical machines," *IEEE Transactions on Industrial Electronics*, vol. 63, pp. 2633–2644, April 2016.

[15] T. Kato and Y. Suzuoki, "Monte-carlo-based modeling of asset management of electric power apparatuses for assessing influence of wide variety of degradation extent and failure probability," in *2009 IEEE 9th International Conference on the Properties and Applications of Dielectric Materials*, pp. 232–235, July 2009.

[16] A. Janssen, D. Makareinis, and C. Sölver, "International surveys on circuit-breaker reliability data for substation and system studies," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 808–814, 2014.

[17] H. Manninen, C. J. Ramlal, A. Singh, J. Kilter, and M. Landsberg, "Multi-stage deep learning networks for automated assessment of electricity transmission infrastructure using fly-by images," *Electric Power Systems Research*, no. Under second review, 2022.

[18] J. D. Spinhirne, "Micro pulse lidar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 1, pp. 48–55, 1993.

[19] ASPRS, "Las specification 1.4 - r14," tech. rep., The American Society for Photogrammetry & Remote Sensing, 2019.

[20] CIGRE WG B2.42, "TB601 Guide for thermal rating calculations of overhead lines," tech. rep., CIGRE, 2014.

[21] R. Ross, *Reliability Analysis for Asset Management of Electric Power Grids*. John Wiley & Sons, 01 2019.

[22] W. Shakoor, "Qhse risk assessment matix," in *Integrated Management System (IMS) Project*, 03 2018.

[23] P. Thomas, R. Bratvold, and J. Bickel, "The risk of using risk matrices," *SPE Economics and Management*, vol. 6, 09 2013.

[24] CIGRE WG B2.13, "TB175 Management of existing overhead transmission lines," tech. rep., CIGRE, 1999.

[25] CIGRE WG B3.32, "TB660 Saving through optimised maintenance in air insulated substations," tech. rep., CIGRE, 2016.

[26] G. Balzer and C. Schorn, *Asset Management for Infrastructure Systems - Energy and Water*. Springer, June 2015.

[27] CIGRE WG B3.06, "TB576 it strategies for asset management of substations - general principles," tech. rep., CIGRE, 2014.

[28] D. Hughes, "Condition based risk management CBRM — enabling asset condition information to be central to corporate decision making," in *CIRED 2005 - 18th International Conference and Exhibition on Electricity Distribution*, pp. 1–5, 2005.

[29] S. Raghavan and B. Chowdhury, "State diagram-based life cycle management plans for power plant components," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 965–972, 2015.

[30] B. O. Mkandawire, N. M. Ijumba, and H. Whitehead, "Asset management optimization through integrated systems thinking and n-1 contingency capability for refurbishment," *IEEE Systems Journal*, vol. 5, no. 3, pp. 321–331, 2011.

[31] L. Pokorádi, "Fuzzy logic-based risk assessment," *Academic and Applied Research in Military Science*, vol. 1, 01 2002.

[32] H. Manninen, J. Kilter, and M. Landsberg, "A holistic risk-based maintenance methodology for transmission overhead lines using tower specific health indices and value of loss load," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107767, 2022.

[33] V. Hodge, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85–126, 10 2004.

[34] F. T. Liu, K. M. Ting, and Z. Zhou, "Isolation forest," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, 2008.

[35] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[36] A. Shaheen and A. Sleit, "Comparing between different approaches to solve the 0/1 knapsack problem," *International Journal of Network Security*, vol. 16, pp. 1–10, 07 2016.

[37] M. Trick, "A dynamic programming approach for consistency and propagation for knapsack constraints," *Ann Oper Res*, vol. 118, 12 2001.

[38] F. Emmert-Streib and M. Dehmer, "Introduction to survival analysis in practice," *Machine Learning and Knowledge Extraction*, vol. 1, pp. 1013–1038, 09 2019.

[39] "EVS-EN 50341-2-20:2018 Overhead electrical lines exceeding AC 1 kV - Part 2-20: National Normative Aspects (NNA) for Estonia (based on EN 50341-1:2012)," Dec. 2018.

[40] "IEC 62271-1:2017 High-voltage switchgear and controlgear - Part 1: Common specifications for alternating current switchgear and controlgear," 2017.

[41] E. Colosimo, F. Ferreira, M. Oliveira, and C. Sousa, "Empirical comparisons between kaplan-meier and nelson-aalen survival function estimators," *Journal of Statistical Computation and Simulation - J STAT COMPUT SIM*, vol. 72, pp. 299–308, 04 2002.

[42] W. Kuckshinrichs and T. Schröder, "Value of lost load: An efficient economic indicator for power supply security? a literature review," *Frontiers in Energy Research*, vol. 3, 12 2015.

[43] P. Raesaar, E. Tiigimägi, and J. Valtin, "Assessment of electricity supply interruption costs in estonian power system," *Oil Shale*, vol. 22, pp. 217–231, 01 2006.

[44] W. Kuckshinrichs and T. Schröder, "Value of lost load: An efficient economic indicator for power supply security? a literature review," *Frontiers in Energy Research*, vol. 3, 12 2015.

[45] The Royal Academy of Engineering, "Counting the cost: the economic and social costs of electricity shortfalls in the UK," 2014.

[46] L. Ilaria and B. Osvaldo, "Regulation of continuity of supply in the electricity sector and cost of energy not supplied," 2009.

[47] Eurostat, "NACE rev. 2: Statistical classification of economic activities in the european community," 2008.

[48] K. Sokk, “Methods of power outage damage assessment and their applicability in the estonian electricity system outage cost assessment methods and application in estonian electricity system,” Master’s thesis, Tallinn University of Technology, 2019.

[49] CEER, “Guidelines of good practice on estimation of costs due to electricity interruptions and voltage disturbances,” 2010.

[50] OFGEM and DECC, “The value of lost load VoLL for electricity in great britain,” 2013.

[51] OpenStreetMap contributors, “Planet dump retrieved from https://planet.osm.org .” `https://www.openstreetmap.org` , 2017.

[52] N. Kilroe, “Line inspections - eyes and ears of rcm,” in *Transmission and Distribution Construction, Operation and Live-Line Maintenance*, pp. 18 – 24, 05 2003.

[53] CIGRE WG B2.08, “TB230 Assessment of existing overhead line supports,” tech. rep., CIGRE, 2003.

[54] CIGRE WG 37.27, “TB176 Ageing of the system—impact on the planning,” tech. rep., CIGRE, 2000.

[55] CIGRE WG 37.27, “193 ageing of the system—impact on the planning,” tech. rep., CIGRE, 2000.

[56] R. Eyre-Walker, G. Howarth, R. Ahmed, and J. Lewin, “Application of advanced condition assessment and asset management techniques on steel tower overhead line electricity networks,” *Asset Management Conference 2014*, 2014.

[57] D. Havard, M. Bissada, C. Fajardo, D. Horrocks, J. Meale, J. Motlis, and M. Tabatabai, “Aged ACSR conductors. II. prediction of remaining life,” in *Proceedings of the 1991 IEEE Power Engineering Society Transmission and Distribution Conference*, IEEE, 1991.

[58] Y. Tsimberg, K. Lotho, C. Dimnik, N. Wrathall, and A. Mogilevsky, “Determining transmission line conductor condition and remaining life,” in *2014 IEEE PES T&D Conference and Exposition*, IEEE, 2014.

[59] S. Gusavac, M. Nimrihter, and L. Geric, “Estimation of overhead line condition,” *Electric Power Systems Research*, vol. 78, no. 4, pp. 566–583, 2008.

[60] P. Thongchai, P. Pao-La-Or, and T. Kulworawanichpong, “Condition-based health index for overhead transmission line maintenance,” in *10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pp. 1–4, May 2013.

[61] G. Ricci, R. Bertella, S. Prato, and M. Zippo, “A.RI.EL: A structural integrity condition based method for renewal assessment of pylons in overhead transmission lines,” in *2018 AEIT International Annual Conference*, pp. 1–6, October 2018.

[62] L. A. Cardenas, D. L. Alvarez, S. Rodolfo Garcia, and F. A. Herrera, “Overhead lines assessment using health index,” in *2019 FISE-IEEE/CIGRE Conference - Living the energy Transition (FISE/CIGRE)*, pp. 1–6, 2019.

[63] Y. Liu *et al.*, “Health assessment and prediction of overhead line based on health index,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5546 – 5557, 2019.

[64] K. Tsang and W. Chan, “Non-destructive stiffness detection of utility wooden poles using wireless MEMS sensor,” *Measurement*, vol. 44, no. 6, pp. 1201–1207, 2011.

[65] S. Mudiyanselage, P. Rajeev, E. Gad, B. Sriskantharajah, and I. Flatley, “Application of stress wave propagation technique for condition assessment of timber poles,” *Structure and Infrastructure Engineering*, vol. 15, no. 9, pp. 1234–1246, 2019.

[66] Y. Yu, M. Subhani, U. Dackermann, and J. Li, “Novel hybrid method based on advanced signal processing and soft computing techniques for condition assessment of timber utility poles,” *Journal of Aerospace Engineering*, vol. 32, no. 4, p. 04019032, 2019.

[67] U. Dackermann, Y. Yu, E. Niederleithinger, J. Li, and H. Wiggenhauser, “Condition assessment of foundation piles and utility poles based on guided wave propagation using a network of tactile transducers and support vector machines,” *Sensors*, vol. 17, no. 12, p. 2938, 2017.

[68] U. Dackermann, B. Skinner, and J. Li, “Guided wave-based condition assessment of in situ timber utility poles using machine learning algorithms,” *Structural Health Monitoring*, vol. 13, no. 4, pp. 374–388, 2014.

[69] CIGRE WG 22.07, “TB141 Refurbishment and upgrading of foundations,” tech. rep., CIGRE, 1999.

[70] N. F. Stanley and H. F. Heap, “Reliability-centered maintenance. Report Number AD-A066579,” tech. rep., United States Department of Defense, 1978.

[71] Y. Liu, J. Xv, H. Yuan, J. Lv, and Z. Ma, “Health assessment and prediction of overhead line based on health index,” *IEEE Transactions on Industrial Electronics*, vol. 66, pp. 5546–5557, July 2019.

[72] Y. Tsimberg, K. Lotho, C. Dimnik, N. Wrathall, and A. Mogilevsky, “Determining transmission line conductor condition and remaining life,” in *2014 IEEE PES T&D Conference and Exposition*, pp. 1–5, 04 2014.

[73] K. Ayu and A. yunusa kaltungo, “A holistic framework for supporting maintenance and asset management life cycle decisions for power systems,” *Energies*, vol. 13, p. 1937, 04 2020.

[74] M. Dong and A. B. Nassif, “Combining modified weibull distribution models for power system reliability forecast,” *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1610–1619, 2019.

[75] P. Blackmore and N. Leeprechanon, “Improving network reliability through effective asset management,” in *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, pp. 1–9, 11 2015.

[76] R. Hashim, F. Usman, and I. Baharuddin, “Determining health index of transmission line asset using condition-based method,” *Resources*, vol. 8, p. 80, 04 2019.

[77] Construction Mechanics Catheter, *Manual of reinforced concrete manufacturing and inspections (in Estonian)*, 1977.

[78] S. Millard and L. Sadowski, “Novel method for linear polarisation resistance corrosion measurement,” 01 2009.

[79] H. Manninen, C. J. Ramlal, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques," *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106726, 2021.

[80] H. Manninen, J. Kilter, and M. Landsberg, "Advanced condition monitoring method for high voltage overhead lines based on visual inspection," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, Aug 2018.

[81] H. Manninen, J. Kilter, and M. Landsberg, "Health index prediction of overhead transmission lines: A machine learning approach," *IEEE Transactions on Power Delivery*, vol. 37, no. 1, pp. 50–58, 2022.

[82] T. Hoens and N. Chawla, *Imbalanced Datasets: From Sampling to Classifiers*, pp. 43–59. John Wiley and Sons, Inc, June 2013.

[83] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[84] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, March 2013.

[85] C. Chen and L. Breiman, "Using random forest to learn imbalanced data," *University of California, Berkeley*, January 2004.

[86] M. J. Anzanello and F. S. Fogliatto, "Learning curve models and applications: Literature review and research directions," *International Journal of Industrial Ergonomics*, vol. 41, no. 5, pp. 573 – 583, 2011.

[87] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," *ArXiv*, vol. abs/1811.12808, 2018.

[88] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 ed., 2009.

[89] S. Marsland, *Machine Learning: An Algorithmic Perspective, Second Edition*. Chapman and Hall/CRC, 2nd ed., 2014.

[90] L. Brieman, "Random forest," *Machine Learning*, vol. 45, p. 5–32, July 2001.

[91] J. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis*, vol. 38, pp. 367–378, February 2002.

[92] T. Fawcett, "ROC graphs: Notes and practical considerations for researchers," *Machine Learning*, vol. 31, pp. 1–38, 01 2004.

[93] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1 ed., 2007.

[94] E. Alpaydin, *Introduction to Machine Learning*. The MIT Press, 2014.

[95] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *The Journal of Machine Learning Research*, vol. 13, pp. 281–305, March 2012.

[96] B. Van den Bergh and S. Pollin, "Keeping UAVs under control during GPS jamming," *IEEE Systems Journal*, vol. 13, no. 2, pp. 2010–2021, 2019.

[97] H. Kinjo, M. Morita, S. Sato, T. Suriyon, and T. Anezaki, "Infrastructure (transmission line) check autonomous flight drone (1)," in *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, pp. 206–209, 2017.

[98] C. Han, Y. Kitamura, A. Kudo, A. Ichinose, L. Rundo, Y. Furukawa, K. Umemoto, Y. Li, and H. Nakayama, "Synthesizing diverse lung nodules wherever massively: 3d multi-conditional gan-based ct image augmentation for object detection," in *2019 International Conference on 3D Vision (3DV)*, pp. 729–737, IEEE, 2019.

[99] C. Guindel, D. Martin, and J. M. Armingol, "Traffic scene awareness for intelligent vehicles using convnets and stereo vision," *Robotics and Autonomous Systems*, vol. 112, pp. 109–122, 2019.

[100] G.-J. Horng, M.-X. Liu, and C.-C. Chen, "The smart image recognition mechanism for crop harvesting system in intelligent agriculture," *IEEE Sensors Journal*, 2019.

[101] B. Jalil, G. R. Leone, M. Martinelli, D. Moroni, M. A. Pascali, and A. Berton, "Fault detection in power equipment via an unmanned aerial system using multi modal data," *Sensors*, vol. 19, no. 13, p. 3014, 2019.

[102] X. Lei and Z. Sui, "Intelligent fault detection of high voltage line based on the Faster R-CNN," *Measurement*, vol. 138, pp. 379–385, 2019.

[103] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10778–10787, 2020.

[104] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.

[105] H.-S. Park and C.-H. Jun, "A simple and fast algorithm for K-medoids clustering," *Expert systems with applications*, vol. 36, no. 2, pp. 3336–3341, 2009.

[106] A. Neubeck and L. Van Gool, "Efficient non-maximum suppression," in *18th International Conference on Pattern Recognition (ICPR'06)*, vol. 3, pp. 850–855, IEEE, 2006.

[107] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, pp. 91–99, 2015.

[108] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[109] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7263–7271, 2017.

[110] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *European conference on computer vision*, pp. 21–37, Springer, 2016.

[111] S. Sanchez, H. Romero, and A. Morales, "A review: Comparison of performance metrics of pretrained models for object detection using the tensorflow framework," *IOP Conference Series: Materials Science and Engineering*, vol. 844, p. 012024, 06 2020.

[112] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.

[113] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117–2125, 2017.

[114] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[115] G. Jocher, A. Stoken, J. Borovec, L. Changyu, and A. Hogan, "ultralytics/yolov5: v3. 0," *Zenodo*, 2020.

[116] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "Cspnet: A new backbone that can enhance learning capability of cnn," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pp. 390–391, 2020.

[117] K. Wang, J. H. Liew, Y. Zou, D. Zhou, and J. Feng, "Panet: Few-shot image semantic segmentation with prototype alignment," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9197–9206, 2019.

[118] M. Tan, R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10781–10790, 2020.

[119] H. Law and J. Deng, "Cornernet: Detecting objects as paired keypoints," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 734–750, 2018.

[120] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "Centernet: Keypoint triplets for object detection," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6569–6578, 2019.

[121] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *European conference on computer vision*, pp. 483–499, Springer, 2016.

[122] R. Nitin and K. N. Mangesh, "Landmark based shortest path detection by using a* and haversine formula," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 1, no. 2, pp. 298–302, 2013.

[123] R. Garnett and R. Stewart, "Comparison of GPS units and mobile Apple GPS capabilities in an urban landscape," *Cartography and Geographic Information Science*, vol. 42, no. 1, pp. 1––8, 2014.

[124] C. J. Ramlal, H. Manninen, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automated utility pole condition monitoring: A deep learning approach," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 255–259, 2020.

[125] D. F. Nettleton, A. Orriols-Puig, and A. Fornells, "A study of the effect of different types of noise on the precision of supervised learning techniques," *Artificial intelligence review*, vol. 33, pp. 275–306, 2010.

[126] X. Zhu and X. Wu, "Class noise vs. attribute noise: A quantitative study," *Artificial intelligence review*, vol. 22, no. 3, pp. 177–210, 2004.

[127] H. Manninen, J. Kilter, and M. Landsberg, "Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices," in *2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, pp. 1–6, June 2019.

[128] CIGRE WG B3.48, "TB858 Asset health indices for equipment in existing substations," tech. rep., CIGRE, 2021.

Acknowledgements

This thesis is dedicated to the memory of my brother, Ergo, who always believed in my ability to successfully finalize my PhD studies. You are gone but your belief in me has made this journey possible.

I would like to acknowledge my supervisors, Jako Kilter and Mart Landsberg, for their guidance and support during my studies. I would also like to thank the Power Line Maintenance Unit at Elering AS for the early implementation of novel methodologies in Estonian transmission grid.

Special thanks to Arvind Singh and Craig Ramlal whom I had an opportunity to work together during my stay in The University of West Indies.

My greatest gratitude goes to my parents and to Liina for the consistent support and encouragement throughout the most difficult times.

Abstract

Data-driven Asset Management and Condition Assessment Methodology for Transmission Overhead Lines

Electrical systems all around Europe are aging, resulting in a significant replacement wave of assets in the next decade. The majority of those costs will be for the replacement of overhead lines (OHLs) that consist of up to millions of individual elements and cross harsh terrain. Due to the changes in the historical consumption and generation pattern since the construction of OHLs, the replacement of worst-condition assets in modern electricity grids may lead to ineffective decision-making. Traditionally, transmission system operators (TSOs) have used time-based maintenance that is easy to implement, but may lead to overinvestments. This thesis presents a holistic framework of a cost-effective asset-management decision-making methodology including state-of-art condition assessment approaches.

Cost-effective risk-based maintenance (RBM) approaches are becoming widespread among TSOs but usually lack transparency in terms of decision-making. The thesis proposes an improved RBM approach, where the probability of failure and consequences of failure are based on the actual technical condition of assets and precise value of lost load. The proposed framework is a data-driven approach, where it is essential to ensure the quality of the input data. The proposed methodology outperformed the most widespread asset-management decision-making methodologies. It also enabled to detect critical assets in the grid and optimise lifecycle costs of transmission OHLs.

Condition assessment of OHLs predominantly relies on manual visual inspections that are costly and may produce vague results on such a geographically expansive system. Unambiguous results of visual inspections are achieved by dividing OHLs into components. Specific visual indicators are developed according to the material's physical fatigue that represent different life-stages of each component using asset health index (HI). To acquire the HI of not evaluated assets, an asset HI prediction methodology based on supervised machine learning is developed and used. The proposed methodology was applied in the case study based on Estonian TSO data, where it enabled to determine HI of each tower in the grid. The HI prediction methodology achieved nearly 80% accuracy without any additional measurement.

TSOs are moving to aerial inspections using high-resolution imaging, but there is still an enormous data processing burden that falls to the asset managers. The thesis proposes an automatic condition assessment methodology based on image recognition using deep learning techniques. The proposed methodology automatically detects transmission poles, disaggregates their components and detects defects on concrete structures and insulators. Detections are mapped onto established HIs of each component. Various state-of-the-art deep learning networks are tested and new performance metrics, specific to this problem, are defined to evaluate their performance based on HI. Automatic condition assessment approaches allowed to determine HI of individual OHL components from high resolution images.

The results illustrate that the proposed methodology enables TSOs to significantly reduce costs related to OHLs' lifecycle management. The novel methodologies hold promise for significantly reducing cost and manual labour associated with condition assessment of transmission OHLs, especially using HI prediction and automatic defect detection from images. The proposed methodology enables to minimize risks cost effectively in the grid compared to traditional approaches and highlights the most critical elements.

Keywords: Aging, Asset Management, Condition Assessment, Deep Learning, Health Index, Machine Learning, Object Detection, Probability of Failure, Risk Assessment, Survival analysis, Transmission Overhead Lines, Value Of Lost Load

Kokkuvõte
Ülekandevõrgu õhuliinide andmepõhine varahalduse ja seisundi hindamise metoodika

Euroopa elektrisüsteemide vananemine toob järgmise kümnendi jooksul kaasa märgatava seadmete väljavahetamise laine. Valdav enamus teostatavatest investeeringutest tuleb suunata kõrgepinge õhuliinide tehnilise seisukorra parandamiseks. Elektri ülekandevõrkude hooldamisel on traditsiooniliselt kasutatud ajapõhist lähenemisviisi, mida on lihtne rakendada, kuid mis viib olemuslikult üleinvesteerimiseni. Samuti ei pruugi tänapäeva elektrisüsteemides investeeringud kõige halvemas seisus olevatesse seadmetesse tähendada kõige kuluefektiivseimaid otsuseid, sest nii tarbimise kui ka tootmise mustrid on võrkude ehitamise ajast tugevalt muutunud. See lõputöö hõlmab endas terviklikku ja kuluefektiivset õhuliinide varahalduse raamistikku, mis sisaldab ka tänapäevastel tehnoloogiatel põhinevaid seisundihindamise metoodikaid.

Riskipõhised varahalduse lähenemisviisid muutuvad ülekandevõrkudes järjest enam laialdasemalt kasutatavaks ja on kuluefektiivsed, kuid nende suurimaks puuduseks on otsuste läbipaistmatus. Selles lõputöös esitletakse andmetel tuginevat riskipõhise varahalduse lähenemisviisi, kus sisendandmete kvaliteet mõjutab tugevalt otsuste efektiivsust. Rikke tekkimise tõenäosuste arvutamisel kasutatakse õhuliinide terviseindekseid, mis on määratud igale üksikelemendile, tuginedes nende tegelikul seisukorral. Rikete tagajärgede määramiseks on välja pakutud andmata jäänud energial põhinev metoodika, mis hõlmab endas ka eeldatava rikke kestvuse prognoosimist ja kaudsete kulude arvestamist.

Traditsiooniliselt tugineb ülekandevõrkude õhuliinide seisundi hindamine manuaalsetel visuaalsetel ülevaatustel, mida on kulukas läbi viia ning mis võivad anda subjektiivseid tulemusi. Paremate tulemuste saavutamiseks jagatakse käesolevas lõputöös õhuliinid väiksemateks vaadeldavateks osadeks, kus igale komponendile töötati välja selle eluetappe kirjeldavad hindamiskriteeriumid. Sellise lähenimisviisi rakendamine võimaldas tõsta visuaalsete ülevaatuste kvaliteeti ja vähendada subjektiivsust. Lisaks töötati välja masinõppel põhinev terviseindeksi prognoosimise mudel, mis võimaldas prognoosida Eesti ülekandevõrgu andmetel koostatud juhtumiuuringus mastide tehnilist seisukorda ligikaudu 80% täpsusega ilma täiendavate tegevusteta.

Õhuliinide ülevaatustel kasutatakse aina rohkem õhusõidukite abi, mis võimaldavad hõlpsasti koguda suurel hulgal kõrge kvaliteediga pilte. Piltide automaatseks ja efektiivseks kasutamiseks pakub see lõputöö välja sügavõppe närvivõrkudel põhinevad pildituvastuse meetodid. Väljapakutud lähenemised suudavad automaatselt tuvastada piltidelt õhuliini defekte ilma käidupersonali abita ja nende alusel hinnata komponentide tegelikku seisukorda. Terviseindeksi määramise täpsuse hindamiseks arendati töö raames välja uued seadmete vananemise eripärasid arvestavad tulemuslikkuse mõõdikud.

Lõputöö tulemused näitavad, et väljapakutud metoodika võimaldab õhuliinide elutsükli halduse kulusid ülekandevõrkudes märgatavalt vähendada. Uute seisundihindamise metoodikate rakendamise tulemused annavad lootust, et nende abil on võimalik oluliselt vähendada õhuliinide ülevaatuste ja hindamise aega ning maksumust, seejuures samal ajal suurendades tulemuste usaldusväärust. Praktilise rakendatavuse osas on suure potentsiaaliga terviseindeksi prognoosimise mudel ning pildituvastusel põhinevate metoodikate kasutuselevõtt õhuliinide ja nende komponentide seisundi hindamiseks. Väljapakutud metoodika võimaldas juhtumiuuringus minimeerida elektrivõrgu riske kuluefektiivsemalt võrreldes üldlevinud metoodikatega ja samuti võimaldas see tuvastada kõige kriitilisemad elemendid võrgus.

Märksõnad: Andmata jäänud energia hind, Elulemusanalüüs, Kõrgepinge õhuliinid, Masinõpe, Pildi-tuvastus, Rikke tõenäosus, Riskide määramine, Sisundi hindamine, Sügavõpe, Terviseindeks, Vanane-mine, Varahaldus

Appendix 1

I

H. Manninen, J. Kilter, and M. Landsberg, “Advanced condition monitoring method for high voltage overhead lines based on visual inspection,” in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, Aug 2018

Advanced condition monitoring method for high voltage overhead lines based on visual inspection

Henri Manninen
and Jako Kilter
Department of Electrical Power Engineering and Mechatronics
Tallinn University of Technology
Tallinn, Estonia
Henri.Manninen@ttu.ee, Jako.Kilter@ttu.ee

Mart Landsberg
Grid Maintenance Department
Elering AS
Tallinn, Estonia
Mart.Landsberg@elering.ee

***Abstract* - This paper presents the advanced condition monitoring method for high voltage overhead lines (OHLs) using visual inspection without the need of highly qualified overhead line experts. The method is for technical condition assessment by minimizing subjectivity in the assessments process and thus maximizing the efficiency of visual inspections. Unambiguous results between inspection patrols are achieved by dividing OHL into assessable components with specific predetermined criteria. This enables to extend OHL's expected lifetime with minimal resources by replacing only components that are in critical technical condition and therefore minimizing life-cycle costs. Methodology is designed for reinforced concrete and steel lattice towers using inspection patrols with specially designed tablet computer (tablet) applications. The aim of the proposed methodology is to minimize risks associated with poor technical condition and optimize replacement strategies throughout life-cycle of an OHL. This paper includes the results of implementing the methodology in Estonian transmission system.**

***Index Terms* - Asset management, health index, life-cycle management, overhead line, visual inspection**

I. INTRODUCTION

The availability and reliability of electrical energy has become the backbone of our society and requirements for electrical energy have increased rapidly in last decades with possible severe economic consequences from short interruptions and technical malfunctions. The most common way to transfer electrical energy from producers to consumers is by using high voltage overhead lines (OHLs) that are well described in [1]. Projected lifetime of OHLs is commonly between 50-80 years [2], but it depends on various conditions such as material, construction quality, climate, maintenance quality etc. At the same time majority of the European transmission grid (110 kV and higher voltage) has been built in 1960s to 1980s [2] and therefore the expected lifetime of OHLs is reaching the final phase of its lifecycle. In Europe alone transmission system operators (TSOs) have to spend over 52 billion euros before 2030 just to maintain the current reliability level and up to 77% (over 40 billion euros) of that budget will be spent on overhead lines replacements or refurbishments [2]. Aging of OHLs and how to overcome the problem of infrastructure deterioration is discussed in papers [3-7]. One of the most complicated challenges is determining the actual technical condition of OHLs, because reaching the final phase of OHLs projected lifetime is not always inferring to its deteriorated technical condition. Therefore more advanced methods to define operational lifetime and real technical condition of OHLs needs to be used in order to obtain more precise assessment than just using age as a deterioration characteristic. One possibility is to use the condition index of OHL by applying fuzzy systems theory as discussed in paper [7]. However, as technical condition of OHL is the basis of all possible further calculations and decisions, then it is important to determine it as accurately as possible.

The most common method to estimate technical condition of OHLs by network operators is using visual inspection performed by inspection patrols. They record detected defects and majority of the maintenance works will be planned based on their information. However, visual inspection is considered as an ineffective way to assess technical condition of OHLs because line inspectors usually note only major defects such as broken insulators or mechanical defects and therefore will not give full information to TSOs about the aging of the OHLs. Improving the visual inspection of overhead lines using reliability-centered maintenance (RCM) is discussed in paper [8], where the key element for success is outlined as education of the line inspectors. There are mainly two options to enhance the quality of visual inspection, where the first is using experts as line inspectors and the second one is to improve assessment technique and use self-explanatory assessment criteria as it is discussed in the aforementioned paper. Visual inspection of OHLs is time consuming and costly and without specific assessment criteria, it can be subjective, depending on assessor's previous experience on the field. Therefore evaluation results from different inspection patrols can vary between regions and assessors on a large scale and that makes comparison of different OHLs inadequate. In order to decrease subjectivity in the assessing process the most inaccurate link, human's capability to see same thing in different angles, needs to be minimized. Therefore, it is essential to introduce highly specific and unambiguous evaluation criteria for all OHL components that not only describe defects but also give indication of the technical condition of the selected component. In doing so, assessors do not have to know background or physical processes behind the assessment but only need to find predetermined visual indicators and record them in the specially designed tablet application. This enables to get good quality results from inspection patrols after just short training and consequently total expenses for inspection patrols decrease, which is further enhanced with possibility to use unmanned aerial vehicles (UAVs) coupled with machine learning techniques to replace foot patrols in the near future.

This paper presents the advanced condition assessment method for reinforced concrete and steel lattice towers based on visual inspection with detailed example of reinforced concrete poles. The paper is divided into five main topics that cover the following: introduction, overview of the new methodology, assessment tool for OHL's condition monitoring, case study in Estonian transmission system and conclusion.

II. Overview of the Methodology

Advanced condition monitoring method for high voltage OHLs based on visual inspection is a new and unique methodology that combines the best of asset management, engineering and material science to achieve reliable results from visual inspections and optimize the lifecycle cost of OHLs. This is achieved by dividing OHLs into components, developing health index (HI) that describe different lifecycle stages for each component and then assessing the components separately to obtain technical condition of all elements in the grid. As the aim of the methodology is to reduce the subjectivity of visual inspection and enhance the quality of results then it is important to ease the assessment process for the line inspectors. That has been achieved by developing a tablet application based on the same methodology that has ancillary online capability of sending all data to the asset database.

It should be noted that visual inspection is only the first indication of OHL's technical condition. Possible decision process is described in paper [9] where inspections are divided into 4 levels e.g. visual inspection, mobile measurements, laboratory measurements, full-scale tests. All levels differ from each other by the level of detail and cost. As visual inspection is the first level of inspections then after receiving critical values from visual inspection it is important to investigate the critical component on next levels to confirm the actual technical condition before committing to expensive investments.

A. Components of high voltage overhead line

In order to get results that are accurate and less biased while simplifying assessors' job on the field, the OHL is divided into smaller observable parts – components. Each component differs from others by construction or by the task it performs and thus it is possible to examine and assess them separately. In this method the OHL is divided into eight different groups of components according to Fig. 1 [10].

Each group of components will be assessed individually using different assessment criteria for every component group. Variety of different materials and designs of the same component turns the assessments of different OHLs into complex task where all alternative materials of same components (for example glass insulators versus composite insulators or reinforced pole versus steel towers) need different assessment criteria but the output must be HI in the range of 0 to 5. In order to assess different alternatives of the component or different components it is essential to focus on the physical processes that occur in the materials as they are aging such as carbonization of reinforced concrete or loss of zinc coating in steel structures and thus describe component's visual image and appearance in selected life stage.

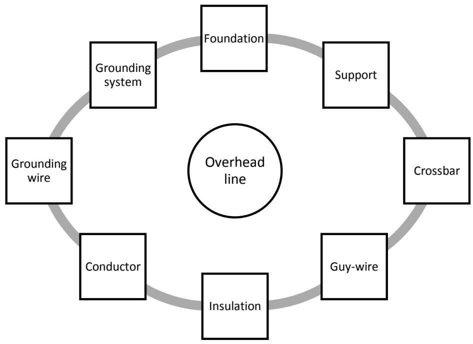


Figure 1. Overhead line structure

As OHL is divided into components then it is important to determine weighting factors for each component to calculate the overall condition index. The current paper proposes values for weighting factor calculated on the cost proportion of selected component to OHL in general based on Estonian practice. This gives opportunity to estimate replacement costs of OHL components as they will be in reality and therefore there is no need to focus on each component separately. Weighting factors for components are the following: foundation – 22%, support – 22%, crossbar – 4%, guy-wire – 2%, insulation – 10%, conductor – 35%, grounding wire – 2.5% and grounding system 2.5%.

B. General description of the methodology

The basic principle of the advanced visual inspection method for OHLs is development of such assessment criteria where every detectable visual sign of selected component matches certain period of the same component's life stage. Therefore, there is HI behind every developed assessment criteria that describes selected component's technical condition in the range of 0 to 5. HI 5 means that selected component is in poor technical condition and its operational lifetime has reached its final stage, while HI 0 means that selected component is in excellent technical condition and its operational lifetime is at its maximum value. HI values are in the range of 0 to 5 are divided linearly and describe technical condition of selected component in that range. Since the basis of the methodology is knowing of what visual signs will emerge when asset ages, then it is possible to deduce current technical condition and residual lifetime of the asset. Calculating the predicted lifetime of the OHL's component is done by using (1).

$$A_{predicted} = \frac{5-HI_i}{5} * A_{projected}, \tag{1}$$

Where $A_{projected}$ is the projected lifetime of the component in the beginning of its lifecycle and HI_i is health index of the same component.

C. Development of health index for reinforced concrete foundations and poles

Methodical example of the derivation of the HI by visual inspection results is done with reinforced concrete structures such as foundations and tower poles. In order to understand the aging of reinforced concrete structures it is important to explain the basic changes in the material's physical properties. The aging of reinforced concrete is described thoroughly in [11], but in general reinforced concrete structures are very durable by construction and their potential lifetime is influenced mostly by external environment. The aging of reinforced concrete structures is determined by the decrease of structural strength in time and it is caused by the change of the concrete's properties that protect steel reinforcements from corrosion. When concrete's surface pH value becomes lower than 8.3 then concrete loses corrosion protective properties for steel reinforcements and when there is enough humidity and oxygen in the environment then steel reinforcements start oxidizing, i.e. rusting. As the rust's volume is up to 10 [12] times more than steel's volume then material's inner stress will causing cracks in the concrete. After time passes this process will accelerate as more water and oxygen will reach the reinforcements and eventually the concrete cover falls off. This will weaken the structure and may lead to breakage of the pole in severe weather conditions.

All kind of defects on the concrete will accelerate the aging of the reinforced concrete structure and hence it is important to identify them. By knowing the most important factor of reinforced concrete's aging, this paper focuses on the different visual signs of the protective layer of reinforced concrete in its life stages. Of course, mechanical damage from heavy transport or agricultural equipment can lead to instant change of technical condition from good to poor, but as the HI is designed to show linear decrease of components' exploitation resources and that means mechanical damages accelerate rapidly the aging process. Example of visual inspection criteria and responding HI values of reinforced concrete poles is presented in Table 1 and sample pictures in Fig.2.

Table I Determination of the HI for reinforced concrete poles

Description of the visual identifier	Health index
There are no defects on the pole (Fig. 2a)	0
Minor defects, but reinforcements are not visible	1
Defects where crosswise-reinforcements are visible	2
Hair-like cracks on the pole (Fig.2b)	2
Defects where cross-reinforcements are visible	3
Micro longitudinal cracks on the pole (Fig.2c)	3
Longitudinal cracks with width 0.3 to 0.6 mm on the pole (Fig.2d)	4
10-20% of passing through defects in the pole's cross section	4
Over 20 % of passing through defects in the pole's cross section	5
Longitudinal cracks with width over 0.6 mm on the pole (Fig.2e)	5
Over the length of 3m longitudinal cracks on the pole	5
Concrete is falling off from reinforcements (Fig.2f)	5

Similar approach as shown for reinforced concrete poles was applied for all components of OHL and as a result of 149 visually identifiable criteria were developed that determine the technical condition of OHLs.

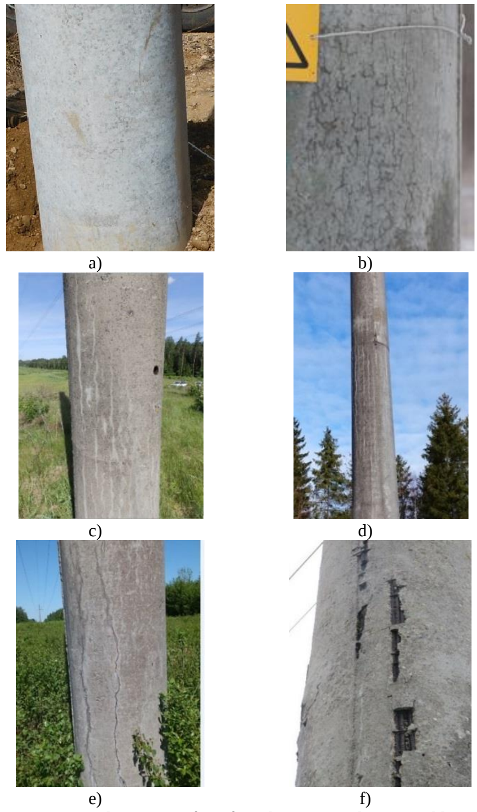


Figure 2. Sample pictures of reinforced concrete poles (see Table 1).

D. Determination of overhead line's general health index

A large amount of data has been gathered after assessing all OHL's components separately that is processed to compare different OHLs and to get general overview of the grid. For more susceptible comparison to extreme values the exponential scale is used to convert HI in the range of 0 to 5 to range between 0 and 100. Equation (2) is used to make the conversion.

$$HI_e = e^{\frac{\ln 101}{5} * HI_n} - 1, \quad (2)$$

where HI_e is the component's HI in exponential scale and HI_n is HI of the same component in linear scale.

After that overall HI of selected tower is calculated according to (3), where conductors and grounding-wires are left out from the calculations because it is nearly impossible to assess their technical condition from the ground level.

$$HI_{et} = \sum_{n=1}^{6} (HI_{e\,n} * w_n), \quad (3)$$

where HI_{en} is the HI of selected component n in exponential scale, w_n weighting factor of component n and n is the number of components.

In order to find summary HI of the whole OHL, (4) is used.

$$HI_{OHL} = \frac{\sum_{i=1}^{n} HI_{et\,n}}{n}, \quad (4)$$

where $H_{et\,n}$ is overall HI of the one single tower and n number of towers.

The main advantage using exponential scale to compare different OHLs is the property of exponential scale that highlights the most critical OHLs. To highlight the advantage two identical OHLs are compared in exponential and linear scales with both consisting of five towers and the HIs for each towers are the following: HI_1=0, HI_2=0, HI_3=1, HI_4=5, HI_5=5.

$$HI_{linear} = \frac{0+0+1+5+5}{5} = 2.2$$

Using linear scale to calculate average HI of OHL gives the result HI=2.2. That means overall technical condition of the OHL is relatively good because the HI of the OHL is in the first half of the scale.

$$HI_{exponential} = \frac{0+0+1{,}51+100+100}{5} = 40.3$$

Using exponential scale [all HI from normal scale to exponential using (2)] to calculate average HI the OHL gives result HI=40.3 and that means overall technical condition of the OHL is in bad condition. It is because using (2) converting HI=4 from linear scale to exponential scale HI=39.1 and therefore average HI of the OHL is determined to be worse than on the line where all towers are assessed with HI =4.

III. Assessment tool for OHL condition monitoring

Specially designed tablet application has been developed for convenient assessment of OHLs and easier data management. The application was created based on the proposed methodology and it includes all assessment criteria in eight component groups with additional group that covers defects in markings of OHLs. Group markings are included because it does not affect technical condition of the OHL, but it is efficient to mark up all defects with one single foot patrol. The designed application user interface is presented on the Fig. 3.

As it is seen from the Fig. 3 then the user interface of the designed application is very intuitive to provide assessors convenient user experience and therefore to ensure the most accurate results. All the aforementioned component groups are in the top row as selectable tabs and under every component group, there are all the assessment criteria that describes the selected component. In order to perform the visual inspection the OHL assessor has to select all identifiable assessment criteria on the field for every component group and specify all the HI assignments. The calculations are performed on the background of the application without line inspector's interventions. This enables to reduce subjectivity of assessors by eliminating the possibility to determine "poor" or "good" technical condition on their own. If there are no identifiable visual signs of the component then the default value of selections is "no visual defects". In addition, there is possibility to add free text comments and take picture of each component.

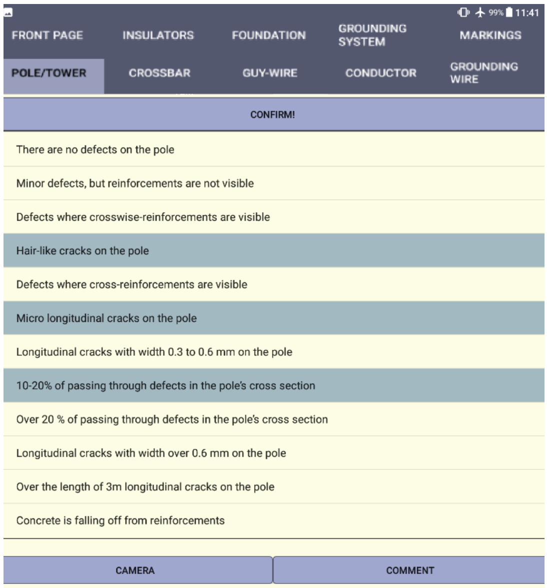


Figure 3. User interface of the designed application

IV. Case study

All Estonian TSO Elering's overhead lines were assessed in 2017 summer with previously described visual inspection method and as a result new information for about 17 000 OHL towers successfully reached the database. 186 rows of new data was generated for each tower, a total of over 3.1 million new data fields that consist of all aforementioned criteria with additional pictures and free text comments.

Summarizing and processing all the data with (2)-(4) enabled to visualize the whole grid as shown on the Fig.4.

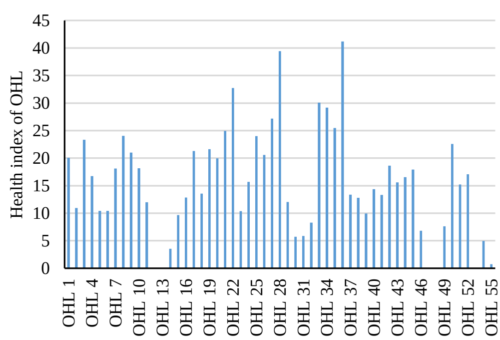


Figure 4. Overhead lines and their overall health indices converted to exponential scale in one part of Estonian transmission network 110 kV- 330 kV

As we can see from the Fig. 4, which includes all assessed OHLs with their calculated health indices in exponential scale. Such an approach does not give specific technical condition of given OHLs but it enables comparison on the same basis and highlights OHLs that are in the worst condition in the grid. For more detailed information about single OHL it is important to represent the whole OHL with all components in linear scale as shown on Fig. 5 because that enables to deduce residual lifetime of components according to (1).

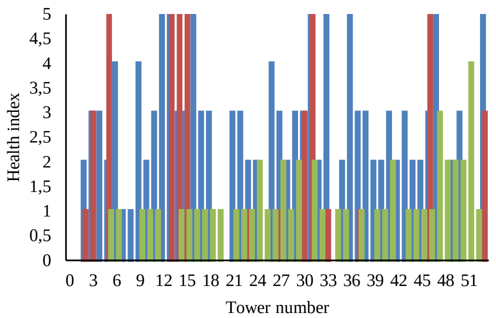


Figure 5. Single overhead line health indices in linear scale (blue – poles, red – foundation and green – insulation)

As seen in Fig.5 the OHL is in relatively poor technical condition due to high HI for 12 towers (HI 5 – end of life criteria). Assuming the OHLs projected lifetime of all components is 50 years then it is possible to plan new investments on the selected OHL because when renovating all components that have HI over four exceeds its predicted lifetime for 10 years. That means changing seven poles and six foundations will give extra 10 years for the OHL lifetime according to (1). Renovating all components with HI over three will exceed predicted lifetime of OHL for 20 years and so on. Knowing the exact technical condition of OHL components will give invaluable information for potential further investments and therefore it is possible to minimize total expenditures by renovating only those components that are in the worst technical condition.

V. Conclusion

Advanced condition monitoring method for high voltage overhead lines (OHLs) based on visual inspection proposes new methodology for OHL condition assessment and health index determination. In order to enhance the quality of visual inspection OHL is divided into components and each component is assessed with predetermined criteria that are developed on the basis of describing different life stages of the components. Implementation of the new method in Estonian transmission grid enabled to gather precise and unambiguous results from inspection patrols. The whole grid was assessed on the same basis and hence it allowed to compare different OHLs across the system for the very first time, (potentially) improving the accuracy and cost-efficiency/effectiveness of future investment plans. Random checks by experts have shown that assessment results are convergent. The proposed methodology will reduce the cost of visual inspections and the subjectivity of results by using component specific and unambiguous assessment criteria without the need of highly qualified personnel to acquire reliable results.

To improve further the methodology and increase the accuracy of the results it is planned to integrate visual inspection results with study on aging of conductors [13] and the GIS database of airborne laser scanning data for all Estonian OHLs. Therefore, the data that determines the need for replacement of components or maintenance would be in same database, which would enable easy access by asset management personnel. Since composed visual inspection method for high voltage overhead lines is developed with the assumption that in near future all Estonian TSO’s OHLs will be assessed with UAVs, then the developed methodology is designed for implementation with automatic picture recognition and machine learning algorithms that could give more homogenous and cost effective results.

VI. References

[1] „Strategic asset management of power networks,“ IEC whitepaper, 2015.

[2] E. Andreini, K. Reich, G. I. Mira, W. Didier, R. Ross, F. Andrea ja T. V. Van, *Facing the Replacement Wave*, ENTSOe, 2015.

[3] H. L. Willis, G. Welch and R. Schrieber, "Aging Power Delivery Infrastructures," Marcel Dekker, New York, 2001.

[4] Working group 22.12, „Technical Brochure 230, Assessment of existing overhead line supports,“ CIGRE, 2003.

[5] Working group 22.13, „Technical Brochure 175, Management of exicting overhead transmission lines,“ CIGRE, 1999.

[6] Working group 37.27, „Technical Brochure 193, Ageing of the System—Impact on the Planning,“ CIGRE, 2000.

[7] S. J. Gusavac, M. D. Nimrihter and L. R. Geric, "Estimation of overhead line condition," *Electric Power Systems Research,* vol. 78, p. 566–583, 2007.

[8] N. Kilroe, „Line Inspections -Eyes and Ears of RCM,“ IEEE Orlando Conference, 2003.

[9] Working group 22.07, „Technical Brochure 141, Refurbishment and upgrading of foundations,“ CIGRE, 1999.

[10] H. Manninen, "Methods for defining the operational lifetime of substations and overhead lines," Master's Thesis, Tallinn: TUT, 2016.

[11] Construction mechanics catheter, "Manual of reinforced concrete manufacturing and inspections (in Estonian)," Tartu: Estonian Agricultural Academy, 1977.

[12] M. G. Grantham, B. Herts and J. Broofield, "The use of linear polarisation corrosion rate measurements in aiding rehabilitation options for the deck slabs of reinforced concrete underground car park," *Construction and Building Materials,* vol. 11, no. 4, pp. 215-224, 1997.

[13] P. Taklaja, I. Kiitam, P. Peetsalu and M. Saarna, "The physical characteristics of aging overhead line conductors," TUT, Tallinn, 2015.

Appendix 2

II

H. Manninen, J. Kilter, and M. Landsberg, “Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices,” in *2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, pp. 1–6, June 2019

Advanced Methodology for Estimation of Value of Lost Load (VOLL) Using Equipment Specific Health Indices

Henri Manninen
Department of Electrical Power Engineering and Mechatronics
Tallinn University of Technology
Tallinn, Estonia
Henri.Manninen@taltech.ee

Jako Kilter
Department of Electrical Power Engineering and Mechatronics
Tallinn University of Technology
Tallinn, Estonia
Jako.Kilter@taltech.ee

Mart Landsberg
Grid Maintenance Department
Elering AS
Tallinn, Estonia
Mart.Landsberg@elering.ee

***Abstract*—This paper presents advanced method to estimate expected VOLL based on equipment specific health index (HI) values for the determination of probability of failure (PoF). PoF calculations are traditionally based on historical failure data and in order to increase the accuracy of the VOLL in this paper, a new method based on equipment specific HI values is introduced. The aim of the proposed methodology is to increase the accuracy of expected VOLL estimation, especially for high voltage overhead lines (OHLs), where there is little to no asset failures in the grid and therefore relying on historical failure data may cause significant errors in the asset management decision-making. This paper includes case study of implementing the methodology in Estonian transmission system.**

Keywords—asset management, health index, value of lost load, probability of failure, transmission grid

I. Introduction

European electricity system has been built mostly in 1960s to 1980s with expected lifetime of 50 to 60 years [1]. That reflects situation where in the following decades there will be inevitable replacement wave of equipment just to ensure the current reliability of electrical supply. Increase in the quality of electrical system affects directly consumers and their market competitiveness because investment funding comes from electricity tariff. High reliability of electricity system implies low outage costs for customers, but high reliability of grid might turn out to be financially irrational due to too excessive investment costs. In addition to aging grid, there are also major transformations in the generation and consumption patterns in European electricity systems that are alternating the direction of investments on large scale. Therefore, it is not cost-effective to renovate all assets that are reaching end of their lifetime because renovating assets that have very low impact on grid's overall reliability is not financially eligible. As the investment decisions are not related only to technical condition of assets then it is essential to use suitable decision-making methods that ensure optimal investments and reliability of electrical system today, tomorrow and in 30 years' time. High reliability of electrical system is one of the most important key factors in investment decision-making because total cost of outages may exceed the cost of electrical energy and assets tens of times.

There have been several large blackouts in North America and Europe that have been analysed in [2]–[4] and they indicate potential risks of large-scale electrical energy interruptions affecting tens of millions of people. For example, North-American blackout in August 2003 affected more than 50 million people and the total economic cost of the blackout varies between 7 and 10 billion dollars [5]. Therefore, it is essential to determine optimum that ensures consumers with sufficient electrical system reliability while maintaining minimal maintenance and investment expenses. Such criteria is illustrated on Fig. 1, where we can see that optimum between failure and outage related costs and expenses to increase asset quality is defined by minimizing total cost of both curves.

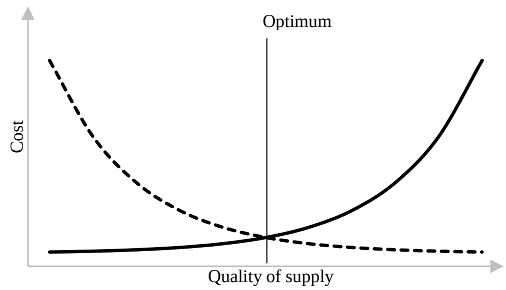


Fig. 1. Simplified optimization principle. Solid line is presenting asset investment and maintenance related costs and dashed line outage and failure related costs [6].

There are various number of methods to estimate financial consequences of outages developed since 1980s to present day that are thoroughly analysed in various literature reviews and publications such as [7]–[11]. The most commonly used method is VOLL and it enables to link financial dimensions with potential outage consequences for asset management decision-making by estimating the price of energy unit that will not be supplied to end-customers. Traditionally outage cost indicator, VOLL, has been used to justify investment decisions by estimating what could be possible consequences when outage in the system happens. Unfortunately, estimating outage cost based only on statistical data and assumptions will cause significant errors for transmission system operators (TSOs) because they are using predictive maintenance and do not have sufficient historical failure data to predict future events. In this paper, it is proposed to use alternative methodology in addition to traditional statistical approach for estimation of expected value of lost load (EVOLL) using asset specific health index (HI) values. HI is fictive parameter that expresses actual technical condition of sophisticated devices in simple manner, usually by the number from 0 to 5. Composition of HI is mostly challenging in technical point of view and it is determined by equipment specifics. By the nature, HI reflects residual lifetime of equipment and therefore determines likelihood of failure. Using HI as a part of VOLL enables to use more advanced PoF calculation methods than

978-1-7281-2650-0/19/$31.00 ©2019 IEEE

statistical probabilities or expert opinions because it has direct link with technical condition of assets.

This paper is divided into five main sections that cover introduction, overview of developed methodology, case study, discussion and followed with conclusions. General overview of methodological explaniation of VOLL assessment is presented in Section II and case study is carried out in Section III. Estonian TSO's data for substation that is connected to electrical system with two OHLs is used for EVOLL estimation.

II. Value of Lost Load and Probability of Failure Calculation Methodology

A. *Value of Lost Load*

The main objective for outage cost assessment is to determine total economic damage due to electricity outage on a common basis regarding different consumers and volumes of energy not supplied. It is relatively sophisticated process because the output is affected by various parameters that are difficult to determine such as consumers profile, expected consumption, duration and time of the outage and even geographical peculiarities of the grid. Outage cost is also affected by customers' risk mitigation measures, frequency of outages and by the state of living in the area under observation. Electricity outages can produce two major types of financial damages where one is directly monetarized such as loss of value caused by direct loss of assets, products or goods, the other is indirectly monetarized such as a loss of time or productivity. There are also damages that are very difficult to monetarize such as social impact on the outage and therefore there are various methods for outage cost calculations based on specific point of view of the methodological focus. Regarding selected method, the output of different approaches remains the same – cost of energy not supplied in €/MWh or $/kWh.

There are mainly two approaches for VOLL calculations and they are based on the collected input data. First approach is using analytical data and second to conduct outage cost functions based on customer surveys. Customer survey methods are most common methods in practice because they enable assessment of outage costs related to parameters that are financially difficult to value in addition to monetarized parameters. The main disadvantage of the method is expensive cost of client surveys and it is time-consuming to gather source data. Analytical methods use economic parameters that are easy to gather from statistical databases and therefore they are easier and cheaper to implement because there is no need for time and money consuming customer surveys but they might not reflect the real behaviour of consumers in the case of electricity outage. Implementations and outcomes using different methods of VOLL calculations is thoroughly analysed in [7] where it is concluded that there is possibility to receive up to tens of times different results for the price of energy unit not supplied just by using different methods, economic environments or consumer categories.

1) *Classification of consumers*

For VOLL calculations, it is difficult to take into account each customer's individual costs because outage costs are depending on various indistinct parameters such as personal consumption, dependence of supply and financial status. As the number of consumers in electricity system is large then it is not practical to involve every single consumer one by one. Therefore, it is reasonable to classify consumers according to rules that enables to compare different consumers on the same basis. In order to achieve comparable results and availability of needed information it is recommended to use international classificatory. CEER (Council of European Energy Regulators) has developed guideline document [12] for VOLL calculations where they recommend using NACE Rev.2 [13] classificatory for electricity outage cost calculations. The NACE groups and sub-categories are explained more detailed in [9]. Based on this approach, CEER recommends the following grouping for a cost-estimation study regarding interruptions.

- Households
- Commercial services (without infrastructure)
- Public services (without infrastructure)
- Industry (without large customers)
- Large customers
- Infrastructure

There is also possibility to develop alternative groupings, depending on the chosen objective of the cost-estimation study or country-specific factors but it is recommended to use international approach for comparable results.

2) *Methodology of value of lost load assessment*

VOLL calculations are usually done based on multi-step approaches that are discussed more in [3], [7], [8], [11], [12], and [14]. Calculation of VOLL reflects the total cost of electricity outage, based on the price of consumer specific energy unit in that consumption point and expected durations of possible outage using (1).

$$\mathrm{VOLL}(r_i) = \frac{C_{L,j}(r_i)}{LF * r_i}, \tag{1}$$

where $C_{L,j}(r_i)$ is cost of energy not supplied, *LF* is load factor and r_i outage duration for consumer *i*.

Irrespective to the selection of methodology, usually outage cost calculation consists of at least three steps that are following.

- Processing of raw collected data (analytical or survey based).
- Developing customer specific interruption cost functions.
- Calculation of VOLL.

Simplified selection process for applicable VOLL method based on input data is proposed in this paper according to Fig. 2 that is composed according to [3] and [8]–[11]. Simple analytical methods are the easiest to implement but relates to the most inaccurate results and detailed customer survey methods produces most accurate results but require also most detailed input data.

Irrespective on that what method is used for VOLL calculation the output of the calculation will be expressed in financial dimensions of possible outage that is determined by consumers, outage duration and frequency. That calculation also covers determination of consumer categories, finding average outage cost values for each category and assumption of the duration of the outage. Because of that, it is possible to reach monetary value that indicates outcome of potential electricity outage in the system.

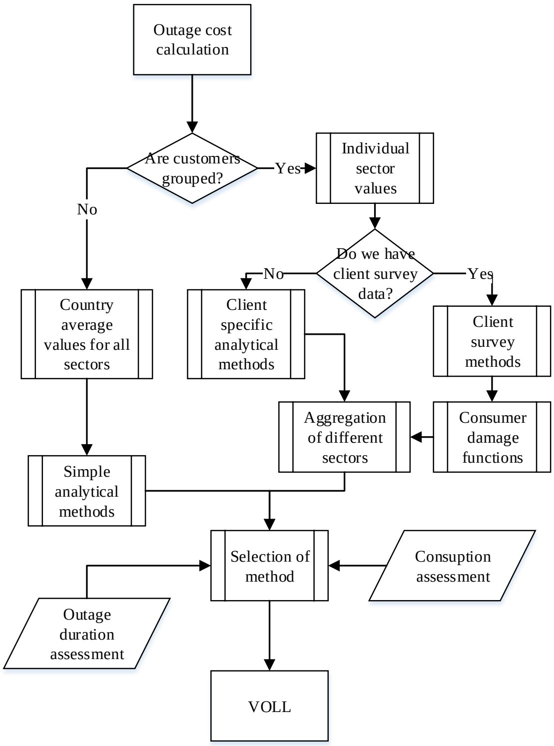


Fig. 2. Simplified selection of VOLL methodology based on input data.

3) Expected value of lost load for investment decision-making

In order to use VOLL in investment decision-making process, it is required to determine additional parameter, PoF that links expected outage cost with probable event in the future. VOLL and PoF relationship is discussed more in [3] and is expressed according to (2).

$$\mathrm{EVOLL} = \sum_{i=1}^{N} \mathrm{VOLL}(r_i) \cdot p(r_i), \qquad (2)$$

where $p(r_i)$ is outage r_i occurrence probability and N is the number of customers.

In addition to improving accuracy of EVOLL calculations, it is essential to specify PoF calculation methodology because inaccurately calculated PoF has potential to influence results on large scale. Without reliable PoF calculations, it is impossible to use EVOLL results in investment decision-making because differences in PoF may alter the results radically.

B. Probability of Failure Using Asset Health Index Values

Usually average historical data is used in system reliability and PoF calculations but it does not take into account the actual technical condition of equipment. At the same time, it must be said that assessment of technically complicated equipment's condition and thereby estimation of PoF in X years is sophisticated challenge. In addition to the usual statistical approach as done in [15], is to use HIs to determine failure rate of assets. There are several publications for PoF calculations based on HIs and failure rates, where most common methodologies are described in [16], [17], and [18]. Intersection of those studies is methodology that enables to link failure rates and the actual technical condition of equipment using historical data with asset specific HIs.

1) Determination of equipment condition score

For this purpose, the equipment is considered as a one and single condition score describing the device is aggregated based on the sub-components condition scores. For example, an entire OHL with all its towers may be under observation, and the overall condition score is calculated as the weighted average of all towers with individual HIs according to (3).

$$\mathit{Condition\ score} = \frac{\sum_{i=1}^{n} w_i r_i}{\sum_{i=1}^{n} w_i} \qquad (3)$$

where w_i is weighting factor of the sub equipment and r_i is normalized HI of the same sub equipment between 0 and 1.

Result of that approach is one condition score for one asset in the range of 0 to 1 where zero means perfect condition and one end-of-life condition.

2) Failure rate using condition scores and historical failure data

For the determination of failure rate using condition scores and historical failure data exponential model is used to calculate failure rate for average, perfect and end-of-life condition scores. Exponential equation for failure rate calculations is based on empirical studies and is described according to (4).

$$\lambda\,(x) = Ae^{Bx} + C, \qquad (4)$$

where λ is failure rate and *x* is condition score of asset.

Parameters *A*, *B*, and *C* are calculated according to historical failure statistics λ(0), λ(0.5) and λ(1) that are failure rates for perfect, average and end-of-life assets. While average value for failure rate λ(0.5) is relatively simple to calculate from historical data then for λ(0) and λ(1) it is important to use more detailed statistical analysis or heuristic assumptions to gather reliable input data. Parameters *A*, *B*, and *C* are calculated according to (5), (6) and (7).

$$A = \frac{(\lambda(0.5) - \lambda(0))^2}{\lambda(1) - 2\lambda(0.5) + \lambda(0)} \qquad (5)$$

$$B = 2\ln\left(\frac{\lambda(0.5) + A + \lambda(0)}{A}\right) \qquad (6)$$

$$C = \lambda(0) - A \qquad (7)$$

3) Failure rate calculation example using historical data

In order to demonstrate calculation methods for perfect and end-of-life for failure rates, Estonian TSO's historical failure data is used. In addition to average failure rates for 110 kV and 330 kV OHLs in Estonia, the perfect and end-of-life failure rates are calculated. For both 110 kV and 330 kV OHLs λ(0) and λ(1) values are modified by 10 and 2 times from statistical averages. That means end-of-life condition has 10 times higher failure rate than average asset and perfect condition asset has 10 times lower failure rate than average asset. That is also done with multiplier 2 as recommended in the study [16]. Failure rates depending on modifiers 10 and 2 for 110 kV and 330 kV OHLs based on previously described asset condition relationship are illustrated on Fig. 3.

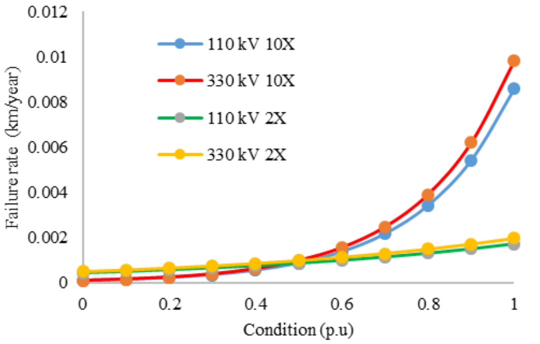


Fig. 3. Failure rates for selected multipliers. Blue line shows 110 kV OHLs λ(0) and λ(1) difference of 10 times from statistical average score and green line 2 times difference from average. Red and yellow line shows 330 kV OHLs λ(0) and λ(1) difference from average values by 10 times and 2 times.

As we can see from the Fig. 3 then modifying failure rates λ(0) and λ(1) changes shape of graphs and failure rates up to 20 times. The larger is the difference from statistical average the steeper is the line on the figure and therefore, technical condition affects failure rate more than where values λ(0), λ(0.5) and λ(1) converge.

4) Probability of failure for overhead liness using health indices of individual components

Based on methodology described in [19] it is possible to determine HI for each tower or component of OHL that represents actual technical condition of selected component. As possible failures of different components of OHLs are causing outages with different duration then additional knowledge can be used when calculating component specific VOLL values. By doing so, it is important to determine expected outage durations for each component group. By going more into the detail then every company can determine specific outage duration for all towers, because there are huge differences in the repair works when comparing towers in forest area or next to road.

For results that are more accurate it is recommended to use component specific historical failure data for each component and determine failure rates for each component group separately. As λ(0) and λ(1) values affect PoF calculations largely then it is crucial to determine reliable values for those parameters as well.

5) Component specific failure rates

In order to use component specific failure rates for each component it is essential to find average failure rates for each component group separately. In order to increase the accuracy of perfect and end-of-life failure rates it is possible to use definition of HI to determine end-of-life failure rate. According to [19] expected remaining lifetime of component is related to HI value and projected lifetime described with (8).

$$A_{predicted} = \frac{HI_{max}-HI_i}{HI_{max}} * A_{projected}, \tag{8}$$

where HI_i is component's HI in the range of 0 to HI_{max} and $A_{projected}$ is the projected lifetime of the same component.

End-of-life condition failure rate λ(1) is described with value of HI and observation period that is defined by maximum value of HI. It should be mentioned that such approach requires linear growth of the HI throughout its life cycle, which again makes HI determination sophisticated. Failure rate for end-of-life condition is calculated according to (9).

$$\lambda(1) = \frac{1}{\frac{A_{projected}}{HI_{max}}} \tag{9}$$

Even if the equipment is in excellent technical condition, there is possibility that there may occur a failure in the event of unexpected conditions. Therefore, λ(0) should not be considered as a zero because there may happen random failures that are impossible to avoid. As an example, there is possibility to use return period of climatic limit load and OHL reliability level values from standard IEC 60826 *Design criteria of overhead transmission lines* [20] as a failure rate for a perfect condition λ(0). When using return period value of 500 years for OHL with reliability level three as perfect condition failure rate value λ(0) then failure rate for one year is according to (9) 0.002.

6) Converting failure rate into probability of failure

PoF is significantly affected by the selection of mathematical methods that are used to determine probabilistic distributions of failures and available statistical data. One possible solution to convert failure rate to PoF is by using (10) according to approach in [20] for the determination of the PoF in observed period.

$$P = 1 - \left(1 - \frac{1}{T}\right)^A, \tag{10}$$

where T is the duration of period when single outage occurs and A is the number of years under observation.

Another possible approach is to use more sophisticated and accurate hazard rate functions for the determination of PoF for OHL components as it is described more detailed in [21] but in this paper simplified method is used.

III. Case Study in Estonian Transmission Grid

For a case study, one substation in Estonian transmission system is used to demonstrate calculation example of proposed VOLL assessment methodology. Selected substation is connected to transmission grid by two OHLs that are sharing same towers, foundations and crossbars but have two different circuits, therefore classical $N-1$ is covered, but in reality, failure of tower's structure will cause outage to that substation. Source data and customers' structure with VOLL values of the selected substation is presented in Table I and Table II.

TABLE I. Source Data of Selected Substation

Excepted outage duration (h)	24
Number of years observed	7
Number of outages (technical) on selected voltage level for OHLs	21
Total number of towers on selected voltage level	12410
Selected return period	500
Number of towers on selected OHL	59
Maximum consumption (MW)	31.56
Average consumption (MW)	15.78

TABLE II. CUSTOMERS' STRUCTURE AND VOLL VALUES FOR SELECTED SUBSTATION

Customer sector	Distribution (%)	VOLL (€/MWh)
Industry	0.36	4030
Commercial services	82.77	5700
Agriculture	0	3750
Households	16.87	3720
Substation	100	5360

To illustrate how different failure rates affect EVOLL values, four different scenarios are modelled using methodology described in Section II.

- New proposed methodology – average failure rate is used for λ(0.5), failure rate based on HI of towers for λ(1) and return period value for end-of-life failure rate λ(0).
- 10X difference – average failure rate is used for λ(0.5), failure rate 10 times higher for λ(1) and failure rate 10 times lower for λ(0).
- 2X difference – average failure rate is used for λ(0.5), failure rate two times higher for λ(1) and failure rate two times lower for λ(0).
- No relation with HI – outage cost is calculated based on VOLL, consumption and outage duration, no relation to HI.

Outage cost calculations are done in following steps:

- Step one – Calculate base failure rates for each scenario using (5), (6), and (7).
- Step two – Calculate failure rates for each HI value and scenario using (4).
- Step three – Calculate PoF for each HI value and scenario using (10).
- Step four – Calculate EVOLL for each tower according to actual HI value and for each scenario.
- Step five – Sum of each individual tower EVOLL for each scenario.

EVOLL values from calculations are presented on Fig. 4 where blue line is EVOLL for new proposed methodology, orange for EVOLL of statistical average with modifier 10, yellow for EVOLL of statistical average with modifier 2 and grey for EVOLL of conventional calculation where HI values are not used. It must be noted that EVOLL value for the proposed methodology using HI 5 for all towers is not shown on Fig. 4 and it is almost 80 million euros.

As we can see from the Fig 4, then the proposed methodology is producing similar values of EVOLL as 10X and 2X until HI 2. After HI 3, the proposed methodology is producing much higher EVOLL than other three scenarios and at HI 5 the difference can be in tens of times. When looking more into detail then the proposed methodology is producing more realistic results than other three scenarios because it is calculating individual PoF for each tower separately and according to actual end-of-life criteria. When all towers of selected OHL are in bad technical condition then most probably more than one tower will fail in case of emergency or in selected observation period and outage duration is longer than 24 hours for single tower replacement. Other three scenarios are not so end-of-life criteria responsive due to relatively low statistical failure rates in transmission grid.

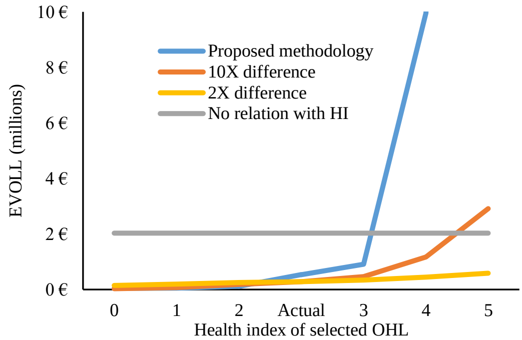


Fig. 4. EVOLL values for four different scenarios.

IV. DISCUSSION

PoF determination is one of the key parameters for VOLL estimation but it has a potential to cause the largest misunderstandings because of its sophisticated nature. As the VOLL calculation has already established framework for decades and dispersion of results is mostly due to different economic backgrounds or selection of implementation methods then PoF is causing inaccuracy in EVOLL due to lack information. It is inevitable to use PoF for investment decision-making, but different approaches such as input from statistics or PoF based on HI is producing scattering results and may alter results on large scale. In order to increase the accuracy of EVOLL it is essential to move to PoF calculation methodology that is based on HI values of assets. This approach is especially useful for TSOs as usually they do not run assets to failure and therefore have little or even no relevant information to predict failures using historical data. In addition, there are always anomalies in the statistical data that can distort historical data sets and therefore produce inaccurate results or unrealistic assumptions in decision-making.

New proposed methodology is especially useful for TSOs that have already determined HI for their assets and it is producing more realistic results than conventionally used methods as it focuses on every asset PoF individually. Therefore, using this methodology enables to have more accurate and timely investment decision by using actual technical condition of assets as the basis of PoF calculations instead of statistical presumptions.

V. CONCLUSION

Advanced methodology for estimation of expected VOLL using equipment specific HIs proposes an alternative to traditional historical failure based approach. New methodology is more suitable for TSOs because it focuses on actual technical condition of assets and lack of failure data does not decrease the accuracy of results for investment decision-making. In order to enhance the accuracy of methodology, VOLL assessment process was thoroughly analysed. In this paper, new PoF method based on HI values of assets combined with statistical failure rates for EVOLL is introduced. That enables to rely more on technical condition of assets and therefore eliminate probable unrealistic assumptions from statistical data.

Case study based on Estonian transmission grid presented that new methodology is functioning at low HI values similarly to methods that are using statistical data, but at high HI values EVOLL price is increasing rapidly due to high HI values that are affecting PoF largely. That outcome is following realistic causes as when all towers for single OHL are in end-of-life condition then most probably more than one tower will fail in case of emergency and outage elimination will take more time than fixing single tower. Therefore, it is possible to conclude that the proposed methodology is following realistic situations because it is using HI values that are the reflection of actual asset condition not pure presumptions.

References

[1] E. Andreini, K. Reich, G. I. Mira, W. Didier, R. Ross, F. Andrea, and T. V. Van, Facing the Replacement Wave. Brussels: ENTSOe, 2015.

[2] A. Berizzi, "The Italian 2003 blackout," in IEEE Power Engineering Society General Meeting, Denver, 2004.

[3] P. Raesaar, E. Tiigimägi, and J. Valtin, "Assessment of electricity supply interruption costs in Estonian Power System," Oil Shale, vol. 22, no. 2, pp. 217–231, 2005.

[4] Y. Makarov, V. Reshetov, V. A. Stroev, and N. Voropai, "Blackout prevention in the United States, Europe, and Russia," Proceedings of the IEEE, vol. 93, no. 11, pp. 1942–1955, 2005.

[5] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, and J. Paserba, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," IEEE Transactions on Power Systems, vol. 20, no. 4, pp. 1922–1928, 31 October 2005.

[6] C. W. g. B3.03, "Guidelines to an Optimized Approach to the Renewal of Existing Air Insulated," CIGRE, 2006.

[7] G. Balzer and C. Schorn, Asset Management for Infrastructure Systems: Energy and Water. Darmstadt: Springer, 2014.

[8] T. Schröder and W. Kuckshinrichs, "Value of lost load: An efficient economic indicator for power supply security? A literature review," Frontiers in Energy Research, vol. 3, no. 55, 2015.

[9] L. Economics, The Value of Lost Load (VoLL) for Electricity in Great Britan. London Economics, London, 2013.

[10] H. Seljeseth, H. Volden, M. Hodmann, and G. Kjolle, "Study on estimation of costs due to electricity interruptions and voltage disturbances," SINTEF Energy Research, Trondheim, 2010.

[11] R. A. o. Engineering, "Counting the cost: the economic and social costs of electricity shortfalls in the UK," Royal Academy of Engineering, London, 2014.

[12] I. Losa, Regulation of Continuity of Supply in the Electricity Sector and Cost of Energy Not Supplied. ERSE, Venice, 2009.

[13] CEER, Guidelines of Good Practice on Estimation of Costs due to Electricity Interruptions and Voltage Disturbances. CEER, Brussels, 2010.

[14] Eurostat, NACE Rev. 2 Statistical Classification of Economic Activities. European Communities, Luxembourg, 2008.

[15] A. Ratha, E. Iggland, and G. Andersson, "Value of lost load: How much is supply security worth?," in 2013 IEEE Power & Energy Society General Meeting, Vancouver, 2013.

[16] R. E. Brown, "Failure rate modeling using equipment inspection data," IEEE Transactions on Power Systems, vol. 19, no. 2, pp. 782–787, 2004.

[17] L. Wenyuan, Risk Assessment of Power Systems: Models, Methods, and Applications. Second Edition, John Wiley & Sons, Inc., 2014.

[18] D. Zhang, L. Wenyuan, and X. Xiong, "Overhead line preventive maintenance strategy based on condition monitoring and system reliability assessment," IEEE Transactions on Power Systems, vol. 29, no. 4, pp. 1839–1846, 2014.

[19] H. Manninen, J. Kilter, and M. Landsberg, "Advanced condition monitoring method for high voltage overhead lines based on visual inspection," in 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, 2018.

[20] IEC, 60826 Design criteria of overhead transmission lines. IEC, Geneva, 2003.

[21] R. Ross, Reliability Analysis for Asset Management of Electric Power Grid. Chennai: IEEE Press, 2019.

Appendix 3

III

C. J. Ramlal, H. Manninen, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automated utility pole condition monitoring: A deep learning approach," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 255–259, 2020

Toward Automated Utility Pole Condition Monitoring: A Deep Learning Approach

Craig J. Ramlal, Arvind Singh and Sean Rocke
Dept. of Electrical & Computer Engineering
The University of the West Indies
St. Augustine, Trinidad & Tobago
{craig.ramlal, arvind.singh, sean.rocke}@sta.uwi.edu

Henri Manninen, Jako Kilter and Mart Landsberg
Dept. of Electrical Power Engineering and Mechatronics
Tallinn University of Technology
Tallinn, Estonia
Henri.Manninen@taltech.ee

***Abstract*—As infrastructure ages, grid operators across the world are becoming more cognizant of the need for monitoring their transmission infrastructure. The geographic extent of the transmission system, however, makes this a difficult and expensive task with thousands of components requiring visual inspection to identify faults which can lead to potentially catastrophic failures. This paper describes the use of Deep Neural Networks to automatically detect areas of concrete damage on utility poles in a European utility from photographs. This is beneficial to reduce time spent in the field as well as variability in between human assessment. The results show that even with a small dataset for training, the network is able to identify new damage with a high level of precision.**

***Index Terms*—Deep Learning, Object Detection, condition monitoring, high voltage, critical infrastructure.**

I. Introduction

The most effective method to transfer high amounts of energy to large distances is by using high voltage overhead lines (HVOHLs). Consequently, a large number of HV OHLs have been built worldwide since the Second World War. The majority of HVOHLs in Europe were built from 1960s to 1980s with expected lifetime between 50-80 years [1], thus a large part of the European transmission system reaching the end of the projected asset life. Since the actual condition of the infrastructure is affected by various parameters such as maintenance policy or climate conditions it is important to base asset decisions on the the actual condition of the infrastructure rather than the age alone.

The most common solution for condition assessment is to use statistical approaches with failure rates of different asset types, with this data it is possible to determine the hazard rate of assets and develop limits for maintenance actions. The approach works well for distribution system operators (DSOs) but not for transmission system operators (TSOs) as they lack relevant failure data. TSOs do not run their assets until failure due to high risks in the electric system and prefer to use preventive maintenance techniques to minimize risks associated with failures in the electricity system. Therefore, TSOs focus on preventive methods that enables the determination of critical defects before failures, such as visual inspection carried out by foot patrols.

Foot patrols are classically the most common method for TSOs to detect defects of HV OHLs and prevent catastrophic failures of the grid by physically examining whole grid. Foot patrol specialists are trained experts that visit towers of OHLs and write down determined defects. This approach is widespread because it gives reliable inputs for maintenance work but it usually lacks the determination of actual technical condition of the OHL and focuses more on critical defects. In addition, major drawbacks for this method are the heavy workload of foot patrols and human factor that enables to assess the same defect differently due to inter-rater variance.

One of the possible solutions to minimize human factor variances and improve technical condition assessment is well explained in [2], where predefined checklists are used to describe visual indicators that determine different life stages of OHL's components. Therefore, foot patrols identify and record predefined visual indicators such as cracks in the concrete and do not give judgement about the technical condition; this is given in the background of the asset management system. Nevertheless, someone still must go to the field and do the assessment of OHLs.

In the last decade, it has become more widespread to perform visual inspections by using airborne vehicles such as helicopters or drones to gather detailed images about OHLs. On one hand, it is relatively easy to gather images about OHLs airborne but on the other hand, assessing technical conditions based on images requires good image quality and a large number of experts that have to review collected images. For example, on average, there are about four towers for each kilometer of 110 kV OHL and solely to assess 100 km of that OHL requires experts to look through at least 1600 images when there is one image taken from each side of the tower. In order to increase the accuracy, the number of images taken from different angles must be increased and therefore, the amount of work to gather reliable information from airborne inspections expands rapidly.

Wooden utility pole condition monitoring mechanisms have been studied in [3]–[7], which give invasive and noninvasive methods of structural health monitoring. There are also practical technical brochures [8]–[10] for metal and reinforced concrete HV OHLs condition assessment by CIGRE that lists most common condition assessment methods of OHLs. These are, however, more focused on the basic framework of the assessment process and provide a list of possible methods instead of specific health index criteria such as magnitude of cracks on reinforced structures that could be used in automatic condition assessment process.

This paper builds on the framework of [2] to determine asset conditions based on visually identifiable defects of OHLs by using machine learning techniques to identify cracks and holes in concrete utility poles to reduce time spent by maintenance personnel in the field and in manually inspecting pictures and minimize variability between human inspections. Several object detection techniques such as fast Region-based Convolutional Neural Network (fast R-CNN) [11], faster Region-based Convolutional Neural Network (faster R-CNN) [12], the region-based fully convolutional networks (R-FCN) [13], single shot detector [14] and you only look once [15]–[17] (YOLO v1, v2, v3), have revolutionized applications such as autonomous driving and face detection through their ability to quickly and accurately classify and localize several classes of objects on images. Foot patrol data has been used in this paper due to data restrictions, however, the potential benefits lie in using ultra high resolution images from fly-by assessments. This would allow thousands of images to be processed and tagged automatically.

This paper is structured as follows, section II presents the description of the data used for training and testing, method of data preprocessing and the networks architecture and training parameters. In order to analyze the performance of the object detection algorithm, a series of tests with popular detection metrics are implemented on the testing data and given and discussed in section III, finally the article concludes in section IV.

II. Methodology

A. Data Description

The full data set comprised 150 images of concrete poles with varying degrees of damage. The samples were taken from pre-existing foot-patrol data so there was no standardization of image requirements. As a result, the data exhibited high levels of variability in terms of backgrounds (*i.e.*, trees, sky, ground), pole orientation and distance, picture angle, shadow and presence of external objects (*i.e.*, transmission lines, insulators and even hands and pens where they were used to give a sense of scale on the pictures). Examples are given in Fig. 1.

B. Data Pre-processing

In order to reduce the computational complexity in training, each image was reduced to a size of 300x300 pixels. Data augmentation was employed to introduce data variability on the features and improve learning [18]. Twenty percent of the training images were randomly selected and augmented by varying the contrast, saturation, hue, magnification, brightness and horizontal flipping. The resulting augmented dataset was 180 images. Bounding boxes were manually placed on each of the degradation artefacts to act as the ground truth for training

[a] [b] [c] [d]

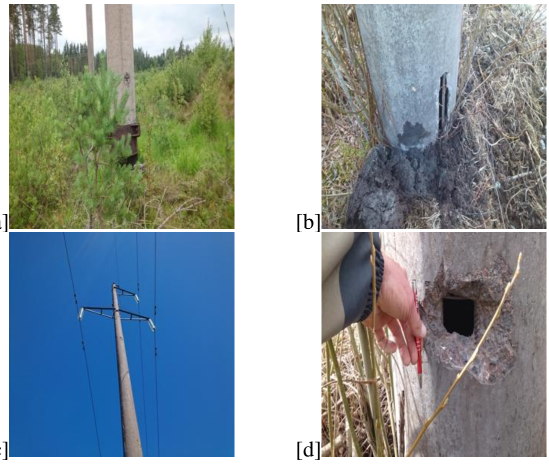


Fig. 1. Input variability (a) vegetation growing across pole (b) earth in the background (c) distant shot with sky background (d) up-close shot with hand in foreground

and validation data. The K-medoids clustering algorithm [19] using the Intersection-over-Union distance metric given by:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

was used to generate the anchor box sizes, where A and B denote the ground truth and estimated bounding boxes, respectively. A holdout cross validation method was used to train and test the neural network. From the data set, 80% of randomly chosen data was used for training and the remaining for testing.

C. Network Architecture

The network architecture consists of a feature extraction network and a classification network. For this study the chosen object detection network was the You Only Look Once (YOLO) V2 object detection model [16]. YOLO frames object detection as a regression problem and unlike Faster R-CNN which uses multiple neural networks to generate potential bounding boxes and classifiers, YOLO unifies the entire pipeline into a single neural network model. The YOLO v2 model extends the original YOLO algorithm by including features such as convolutions with anchor boxes, batch normalization, direct location prediction and dimension clusters etc. These guarantee faster training and a more robust network for object detection. In designing the YOLO system model, the GoogleNet classifier [20] was used as the feature extraction network, with the "inception 4c-output" layer chosen as the feature extraction output layer to the YOLO detection network. The final model comprises 91 layers and 105 connections.

D. Performance Metrics

When an network is fed an image, it returns a set of bounding boxes corresponding to predicted hole locations. These detections can be classified as follows:

- True Positives (TP) - Where the network has correctly identified a defect
- False Positives (FP) - Where the network has detected a defect where none existed
- False Negatives (FN) - Where the network has failed to detect an existing defect

A true positive is determined by the IoU, equation 1, between the predicted bounding boxes with the defined ground truth boxes. If a predicted bounding box does not sufficiently overlap with any Ground truth it is recorded as a false positive. Conversely, if a Ground truth box does not sufficiently overlap with any predicted bounding box it is recorded as a false negative. Note that True Negatves (TN) do not have bounding boxes, either in Ground truth data or predictions and as such they cannot be counted. This study uses average precision benchmark to evaluate the performance of the object detection network. The Average precision value is a numerical metric based on the precision-recall curve, calculated by interpolating and taking the area under the curve. To calculate the average precision the following definitions of precision and recall are needed:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

The AP is expressed as:

$$AP = \sum_{i=1}^{n-1} (r_{i+1} - r_i)\rho_{\text{interp}}(r_{i+1}) \tag{2}$$

where the interpolated precision is given by:

$$\rho_{\text{interp}}(r) = \max_{\{r' \geq r\}} \rho(r')$$

and r is the recall level.

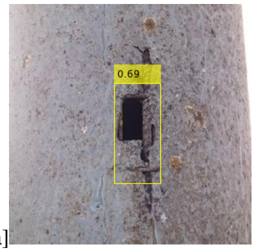
E. Network training

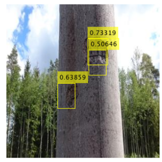
The object detection network designed was trained using the following algorithms and parameters. The stochastic gradient with momentum (SGDM) method is the chosen training technique. (SGDM) accelerates convergence by replacing the actual gradient by an estimate, calculated from a randomly selected subset of the data. The momentum, minibatch size, epochs and initial learning rate were chosen as 0.9, 16, 40 and 0.001 respectively. The training dataset comprised of 144 original and augmented images.

III. Results and Discussion

The testing dataset comprised 37 images of damaged poles with a total of 43 damaged areas. The classifier outputs a vector of bounding box sizes and locations for each image, corresponding to the estimated hole locations. Figure 2 shows example outputs for different backgrounds with the bounding boxes overlaid on the images and their corresponding confidence scores. The threshold confidence score for the detector to assign a positive detection was 0.5.

For the examples shown, the classifier successfully detected all the damaged areas and similarly across the entire dataset, was able to the detect all instances of damage as long as the hole was up close. When images were taken at a distance, so that the damaged areas were very small compared to the image size, several instances of missed detections were noted, e.g. Figs 3–4 . In some cases even when the confidence threshold for detection was reduced to a very low value such as 0.2 the damage could not be detected although several false positives began showing up in other images. This poor performance for small objects is a recognized issue with the YOLO algorithm, and would have to be taken into consideration when specifying guidelines for taking images.

[a] 

[b] 

[c] 

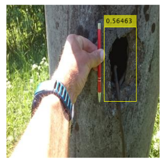
[d] 

Fig. 2. Example outputs from detector with different backgrounds

Fig. 6 illustrates the precision–recall performance for various IoU thresholds for the hole detection system. The final point on each graph represents the Precision and Recall for the entire set of data at the given IoU. As evident, both the precision and recall increase with decreasing IoU till a value of 0.4, below which the graphs overlay perfectly. Fig. 7 illustrates the average precision as a function of the IoU threshold. As seen for the dataset, while decreasing the IoU threshold increases the precision (and average precision) performance of the system, there is a point below which the precision does not improve, or improves marginally. The precision and recall over the whole set of data for IoUs less than or equal to 0.4 is approximately 0.69 and 0.83 respectively. Higher IoUs indicate better detection localization. However, this is not as important as detection for this particular problem. In fact, the IoU should be selected based on highest Recall and Precision and localization dealt with only in as a secondary matter of concern. The present study has shown that the lower IoU thresholds(0.4-0.1) give classifications commensurate with classification by visual inspection. Manual inspection of the boxes indicate that this may be primarily due to the irregularity in hole shape and lack of defined boundaries in damaged areas leading to intrinsically higher ground truth labelling noise. This in turn leads to

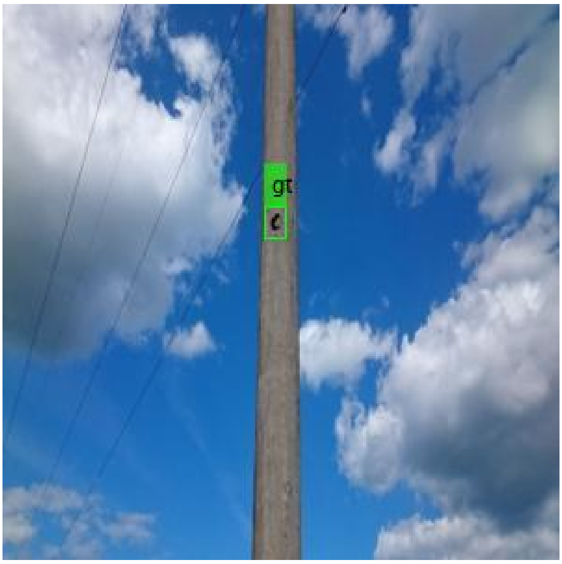


Fig. 3. Far from view Hole missed detection (single hole)

Fig. 5. Multiple detection of single damaged area

Fig. 4. Far from view Hole missed detection (multiple holes)

variations in bounding box sizes or multiple detections of a single damaged area. Fig. 5. Since the Intersection between the boxes is limited to the area of the smaller bounding box, a small predicted bounding box which perfectly overlaps with a larger ground truth box would have a small IoU.

Given the context of the application, more relaxed criteria for detecting a hole may be desirable even though it may increase the number of false positives or multiple detections since it would allow more actual holes to be flagged for maintenance.

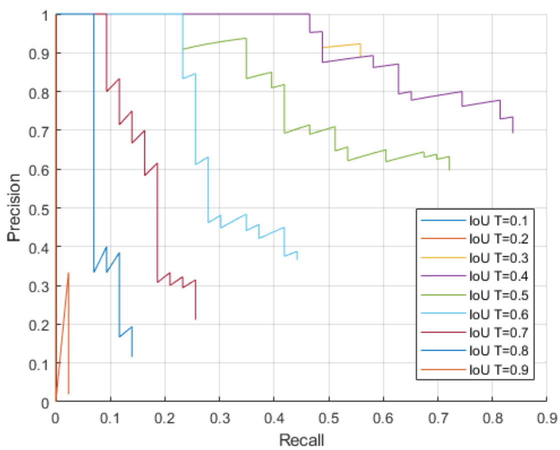


Fig. 6. Precision vs Recall Graph for IoU threshold sweep(0.1-0.9)

IV. Conclusion and Further Work

The object detection method was successful in detecting cracks and holes in utility poles from non–standardized images gained from foot patrols with variations in background, foreground, shadowing and distance. The network performed poorly in instances when the pole was very far and the size of the hole was small compared to the overall image. This, however, can be dealt with by creating photograph guidelines for data collection. In industrial application, Utilities should

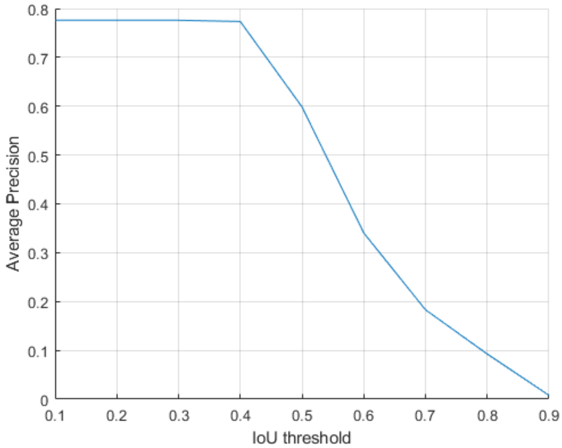


Fig. 7. Average Precision vs threshold IoU sweep(0.1-0.9)

test a small sample of their data to gain insight into the appropriate confidence thresholds used for detection.

Further work should include testing with various feature extraction networks for detecting varying degrees of degradation which may be difficult to differentiate by using multiple classes of objects. Additionally, to determine suitable confidence and IoU thresholds, a framework for cost-based tradeoff analysis should be developed to investigate selection criteria for this application space.

References

[1] ENTSO-e, “Facing the Replacement Wave,” 2015.

[2] H. Manninen, J. Kilter, and M. Landsberg, “Advanced condition monitoring method for high voltage overhead lines based on visual inspection,” in *2018 IEEE Power Energy Society General Meeting (PESGM)*. IEEE, 2018.

[3] K. Tsang and W. Chan, “Non-destructive stiffness detection of utility wooden poles using wireless MEMS sensor,” *Measurement*, vol. 44, no. 6, pp. 1201–1207, 2011.

[11] “Fast R-CNN, author=Girshick, Ross, booktitle=Proceedings of the IEEE international conference on computer vision, pages=1440–1448, year=2015.”

[4] S. Mudiyanselage, P. Rajeev, E. Gad, B. Sriskantharajah, and I. Flatley, “Application of stress wave propagation technique for condition assessment of timber poles,” *Structure and Infrastructure Engineering*, vol. 15, no. 9, pp. 1234–1246, 2019.

[5] Y. Yu, M. Subhani, U. Dackermann, and J. Li, “Novel hybrid method based on advanced signal processing and soft computing techniques for condition assessment of timber utility poles,” *Journal of Aerospace Engineering*, vol. 32, no. 4, p. 04019032, 2019.

[6] U. Dackermann, Y. Yu, E. Niederleithinger, J. Li, and H. Wiggenhauser, “Condition assessment of foundation piles and utility poles based on guided wave propagation using a network of tactile transducers and support vector machines,” *Sensors*, vol. 17, no. 12, p. 2938, 2017.

[7] U. Dackermann, B. Skinner, and J. Li, “Guided wave–based condition assessment of in situ timber utility poles using machine learning algorithms,” *Structural Health Monitoring*, vol. 13, no. 4, pp. 374–388, 2014.

[8] W. group 22.12, “230 Assessment of existing overhead line supports,” CIGRE, Tech. Rep., 2003.

[9] Working group 22.13, “175 management of existing overhead transmission lines,” CIGRE, Tech. Rep., 1999.

[10] W. group 37.27, Tech. Rep.

[12] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Advances in neural information processing systems*, 2015, pp. 91–99.

[13] J. Dai, Y. Li, K. He, and J. Sun, “R-FCN: Object detection via region-based fully convolutional networks,” in *Advances in neural information processing systems*, 2016, pp. 379–387.

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “Ssd: Single shot multibox detector,” in *European conference on computer vision*. Springer, 2016, pp. 21–37.

[15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, real-time object detection,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[16] J. Redmon and A. Farhadi, “YOLO9000: better, faster, stronger,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.

[17] ——, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[18] L. Perez and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” *arXiv preprint arXiv:1712.04621*, 2017.

[19] H.-S. Park and C.-H. Jun, “A simple and fast algorithm for k-medoids clustering,” *Expert systems with applications*, vol. 36, no. 2, pp. 3336–3341, 2009.

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

Appendix 4

IV

H. Manninen, J. Kilter, and M. Landsberg, “Health index prediction of overhead transmission lines: A machine learning approach,” *IEEE Transactions on Power Delivery*, vol. 37, no. 1, pp. 50–58, 2022

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

Health Index Prediction of Overhead Transmission Lines: A Machine Learning Approach

Henri Manninen, *Member, IEEE,* Jako Kilter, *Senior Member, IEEE,* and Mart Landsberg, *Member, IEEE*

***Abstract*—This paper presents an asset health index (HI) prediction methodology for high voltage transmission overhead lines (OHLs) using supervised machine learning and structured, unambiguous visual inspections. We propose a framework for asset HI predictions to determine the technical condition of individual OHL towers to improve grid reliability in a cost-effective manner. The paper focuses on asset HI prediction and the selection of the most parsimonious model. Based on the technical specifications and HI data, our methodology allows for the prediction of a HI for OHLs without HI data, and models asset aging behaviour. Technical specifications and the HI as defined in this paper are taken from the Estonian TSO periodical visual inspections implemented in 2018. The case study successfully demonstrates that the proposed methodology can predict tower HI values for a single OHL with nearly 80 percent accuracy without the need for additional measurements.**

***Index Terms*—Aging, Asset Management, Classification, Health Index, Modelling, Prediction Model, Supervised Machine Learning.**

I. INTRODUCTION

OVERHEAD transmission lines are the backbone of the electricity system, enabling the transportation of large amounts of electric energy across large distances in a cost-effective manner. The majority of European electricity systems are aging [1], and alternative ways are urgently needed to maintain and increase the reliability of this critical infrastructure. It is financially impossible to refurbish all old assets or to build all new assets; investments are made only for assets with a poor technical condition. Assest management has become a major challenge for most transmission (TSO) and distribution (DSO) system operators, creating the need for alternative opportunities to maximise the remaining lifetime of their assets. Historically, the most common approach has been to use interval-based maintenance, where assets were replaced after certain years in service, but that can lead to over-investing. Another approach is to determine the actual technical condition of assets by using condition monitoring techniques and maintenance instead of using age as the key indicator for investment decisions. The age of an asset does not automatically imply a poor technical condition, but is rather dependent on the type of asset, manufacturer, material, climate, weather events, air pollution in the area, etc.

The most common means to describe the technical condition of assets is via a health index (HI) [2] and [3]. Specifically, the HI is a number indicating the asset's technical condition using predefined categories. However, there are a vast number of assets in the actual grid, and few have measurable parameters. Therefore, a HI is rarely used for OHLs or towers because OHLs are considered apriori to have high reliability, a long lifespan, and they cover large areas with limited physical access. Traditionally, condition assessment of OHLs is performed by periodical visual inspection, where critical defects are recorded; a comprehensive overview can be found in [4], [5], [6] and [7]. In [8] for example, foot patrols are used to detect and record all predefined visual indicators found on site. An alternative to traditional human inspections, visual indicators are used to determine HI values via mobile applications instead of relying on a human rater. In the last decade, areal inspections using helicopters and unmannded aerial vehicles (UAV) have gained in popularity due to advances in short range remote sensing, however aerial monitoring remains prohibitively expensive for full network coverage.

An exhaustive assessment of individual assets is therefore infeasible, and methodologies are needed which focus on determining HI without direct measurements, or using statistical approaches. One of the most widely used methods taken from reliability engineering uses a bathtub curve developed for United Airlines in 1978 [9] as a decision-making indicator. Bathtub curves are widespread because of their straightforward implementation. Unfortunately, it has been shown by [3] that when considering high voltage substation equipment, the curves are suitable for use on a small portion of assets with very specific failure modes. In addition to the bathtub curve, electrical equipment condition assessment can also make use of stochastic simulations using the Monte Carlo method as well as machine learning algorithms including artificial neural networks. Usually, these models are used to predict the technical condition of the most expensive assets; transformers [10], electrical machines [11], cables [12] and circuit breakers. These assets usually have condition monitoring systems that provide measurements of critical parameters. Based on the measurements and technical information, the remaining lifetime is predicted using a model. An example of a mathematical approach to overcome OHL's HI determination is proposed in [13] where trends in data, expert opinions, environmental factors and the weighting of different parameters are combined to determine the HI of OHLs. To overcome issues affected by statistical anomalies, we have developed a novel HI prediction methodol based on supervised machine learning combined with structured visual inspection data. The approach is suitable to predict HI in a cost-effective and reliable manner, and can be implemented to improve decision-making and model asset

H. Manninen, J. Kilter and M. Landsberg are with Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Estonia e-mail: henri.manninen@taltech.ee

Manuscript received Month XX, 20XX; revised Month XX, 20XX.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

aging behaviour.

This paper is divided as follows: Section II outlines the proposed methodology, Section III compares the performance of supervised machine learning models, and Section IV presents the findings of our case study and discusses the applicability of the proposed methodology to predict the HI and model asset aging on the Estonian TSO grid.

II. Methodology

Here, we propose a new method using supervised classification algorithms combined with a HI determination framework using predefined visual indicators to predict the technical condition of OHLs on unseen data. The backbone the methodology is the unambiguous determination of an asset's HI using foot patrols to classify OHL condition. First, all OHLs were divided into towers and components, where each component was assessed separately using a tablet application to enter the visual criteria. This enabled the detection of existing defects using a standardized data entry methodology. HI values were then calculated by the application itself to minimize the human factor in the assessment process, to ensure consistency between different inspectors. Table. I and Fig. 1 present an example of a reinforced concrete pole HI determination based on predefined visual indicators. A similar approach was implemented for steel lattice towers, where the HI was determined based on the presence of mechanical defects, bolt condition, rust level and cross-sectional reduction. In [8], nearly 150 different criteria were used to define technical condition of all OHL components. The components included towers, foundations, insulators, grounding systems, cross-bars, guy-wires and conductors. In this paper, a single HI value of the OHL tower was determined using the maximum HI value for the tower and foundation combined. For example, if there is a single defect with a HI value of 5, on the foundation or on the tower, then the HI of the tower is 5.

TABLE I. Description of visual indicators presented in Fig. 1 and corresponding HI values [8]

Visual indicator	Health Index
There are no defects on the pole (Fig.1a)	0
Hair-like cracks on the pole (Fig.1b)	2
Micro longitudinal cracks on the pole (Fig.1c)	3
Longitudinal cracks with width 0.3 to 0.6 mm on the pole (Fig.1d)	4
Longitudinal cracks with width over 0.6 mm on the pole (Fig.1e)	5
Concrete is falling off from the reinforcements (Fig.1f)	5

The asset HI values used in this paper are defined as a set of discrete HI categories from 0 to 5. Classification algorithms were applied to predict the corresponding HI categories in lieu of the remaining lifetime. The asset's HI is defined as a number representing the expected remaining lifetime, which is a standard input variable used in investment decision-making (1). The value therefore describes the asset's technical condition with respect to its projected lifetime. The HI determination indicators used in [8] were developed according to the moment of occurrence in the asset life-cycle, and thereby the following

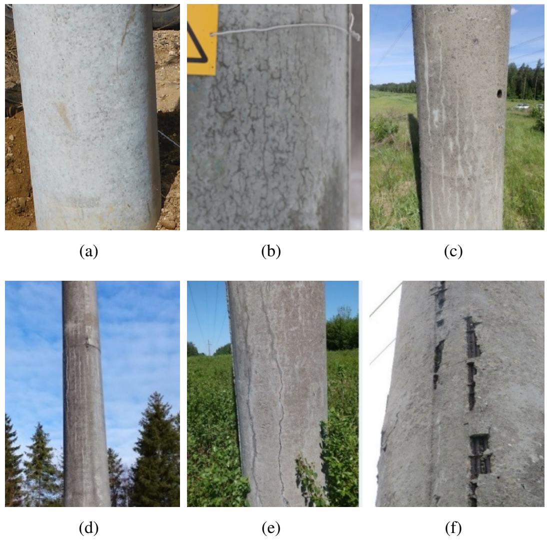


Fig. 1. Examples of reinforced concrete pole visual indicators. Clarifications for the indicators are given in Table I [8].

linear equation (1) can be used to calculate the expected lifetime of a given asset:

$$L_{expected} = \frac{HI_{max} - HI_i}{HI_{max}} * L_{Projected} \quad (1)$$

where HI_{max} is the maximum value of the HI, HI_i is the HI of the selected asset, and $L_{Projected}$ is the projected lifetime of the asset.

In addition to the HI classification, the asset's technical information and specifications are used for HI prediction for assets not in the training or testing data sets. These assests were used for validation because all OHLs in the Estonian TSO database are described with a number of technical features during the asset design and construction phases. Although this information does not change during service, it can affect the speed of the OHL's aging process. Features that influence the technical condition, construction quality or mechanical stresses of the OHLs are presented in Table II. The parameters used in this work are the number of circuits on a single tower, the tower type where the support and tension tower are separated, the nominal voltage level of the OHL (110 kV or 330 kV), presence of bird protection, tower material (reinforced concrete poles, colored steel lattice towers, zinc-coated steel or untreated steel lattice towers). In total, these assets encompass six different manufacturers and 214 different tower configurations. In addition, angled towers are distinguished from vertical OHL structures due to the increased mechanical stresses which they experience. The final feature in our methodology is the current age of the tower, calculated from the installation date.

The framework of the asset HI prediction model used in this paper is presented in Fig. 2. Asset HI prediction starts with the collection and comparison of an asset's technical features

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

TABLE II. Model input data description

Feature name	Number of different features
Number of circuits	4
Tower Type	2
Voltage	2
Material	4
Manufacturer	6
Configurations	214
Angle	2
Bird protection system	2
Age	1 to 67

and HI data. All assets in the database must have technical feature data because that information is collected and saved as a requirement at the beginning of each OHL's life-cycle. Assets that have both technical features and HI data are used as model training data. That means the data of each OHL tower used in the training process is the combination of collected HI values and nine features that describe the technical parameters of the OHL tower. All assets that do not have HI data but have technical features will not be included in the training process; HI values of those assets will be predicted by using a trained prediction model. That approach enables the prediction model to have as much relevant training data as possible from already excising data and therefore improved overall prediction accuracy. On the right side of the chart it is seen that the methodology is divided into two parallel sections after model training data selection. The first branch is the missing asset HI prediction and the second is the asset aging behaviour modeling. The output of both branches in the methodology is HI prediction results for each tower that was selected for HI prediction.

The main difference between the missing asset HI prediction and asset aging behaviour modeling is the step where the selected asset age parameter is modified while all other features remain unchanged. That allows, in terms of sufficient input data, to predict asset HI based on the best knowledge and a similar performance of assets. As long as the modified time intervals are not unrealistic and there is a sufficient number of training samples in the grid, this approach enables different scenarios to be modeled in the near future. For example, if the input data has samples of towers distributed evenly from the age of 1 year to 80 years, then it is possible to model the aging of the selected OHL in those limits. After the limits are exceeded, the model will become inaccurate due to the missing references in the population. The prediction model is trained using all existing data from the facts and measurements similarly as in the missing data approach.

A. Input Data

Input data for this paper is collected from periodical visual inspections in one part of the Estonian TSO grid. The aforementioned HI determination methodology was implemented in the Estonian transmission grid in 2018 and all identified defects were noted using a specially designed tablet application, while the HI assessment was made in the background. That resulted in a determination of individual HI values for each component of each tower separately. For the simplified example, this paper uses HI values that are aggregated to tower

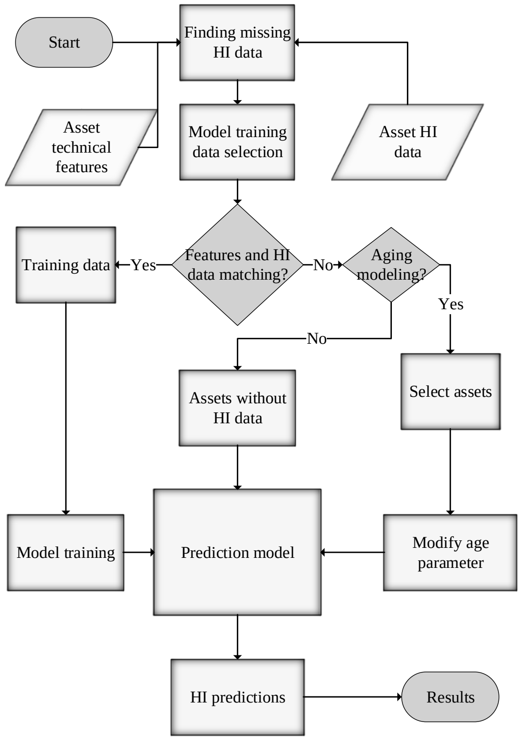


Fig. 2. Flowchart of the asset HI prediction process that enables the prediction of missing asset HI values and model aging behaviour of selected assets based on asset technical features and HI data.

level based on the maximum function of OHL supporting components such as poles, foundations and crossbars. In terms of HI data, it is possible to scale this methodology for a more detailed approach where the HI of each component is predicted separately. After inspection, the HI data was cleaned and errors were fixed in the database manually. It is also possible to use outlier detection algorithms to detect incorrect values in the data automatically as described in [14]. However, as the input data was already pre-processed manually, outlier detection is not used in this paper. The data used in this paper consists of 26,273 rows described in Table III and presented in Fig. 3. As seen in Table III, HI data is concentrated around HI2

TABLE III. Count of different health index values for training and testing data after train/test split

Health Index Class	Training	Testing	Total
HI0	6217	1555	7772
HI1	1799	450	2249
HI2	8964	2241	11205
HI3	2770	692	3462
HI4	1182	295	1477
HI5	86	22	108
Total	**21018**	**5255**	**26273**

with extremely imbalanced classes with 108 samples in the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

HI5 class and 11,205 samples in the HI2 class. For model training purposes, data is divided into training and testing with a ratio of 80:20, using stratification to maintain equal class proportions for each class. That enables the percentage of samples for each class to be preserved and therefore ensures that both data sets include the necessary samples for all classes. As described in [15], stratified re-sampling is easy to implement and has a positive effect both on the variance and bias. It is especially beneficial in the case of a class imbalance as it is present in this data set. If stratification is not used, then there may be a possibility that the test or training sets may not contain any instance of a minority class at all. In Table III, it is also seen that the same imbalanced data tendency is still present after the data is split for training and testing using stratification.

The distribution of data in different HI classes according to the age of the assets is presented in Fig. 3. As seen from the scatter graph (left), there is no strong relationship for OHL assets between the HI and age. There are even a few assets with HI = 5 after only 10 years in service, and some assets have HI = 0 even after 60 years. The regression line on the graph demonstrates that there is light tendency for an asset's HI to increase after a long time in service, but based on that graph it will not exceed HI2 even after 60 years of service. From the data description chart (right) it is seen that a majority of assets in the fleet are 40 to 60 years old with HI2. There is also a larger concentration of assets with an age of around 10 years and HI = 0.

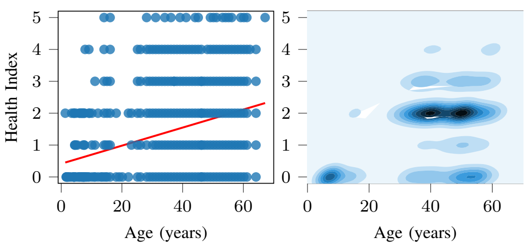


Fig. 3. Description of used data based on HI and age. The scatter graph with linear regression line (red) on the left presents each tower's HI values in relationship to their age. The graph on the right shows the distribution of assets with different HI values and age where the darker color reflects a denser distribution.

B. Class-Imbalance

TSOs have usually implemented interval-based preventive maintenance strategies and do not run their assets until failure. This is shown in Fig. 3, where high HI values, such as HI5 and HI4, are considered as minority classes and where low and medium HI values are considered as majority classes. It is typical for TSOs because a large amount of assets with a critical technical condition have already been replaced or will never reach their end-of-life condition to minimize the risks associated with loss of load. As this paper focuses on the asset HI prediction and not on criticality analysis, there is no need to highlight one class over another, and therefore, all classes will be treated equally.

According to [15] and [16], class imbalance of data occurs when data sets exhibit significant imbalances on the order of 100:1, 1,000:1 and even 10,000:1 between majority and minority classes. When we look at Table III there are 11,205 towers with HI2 and only 108 towers with HI = 5. This means that the ratio of majority vs minority class in this data set is 104:1. In the case of class imbalance, the results from classification prediction models are not the same as using balanced data for model training. Use of imbalanced data in the training process usually causes classifiers to have poor predictive accuracy towards the minority class compared to the other classes and results in a tendency to classify most unseen samples in the majority classes. A decrease in model performance in the case of class imbalance is caused by the model's loss functions, which attempt to optimize error rate or the accuracy of the model without considering the real distribution of different classes. This decreases the performance of the model, and therefore different methods are used to minimize the results of imbalanced data in prediction model learning processes. It is usually made by three methods that are a downsampling of majority classes using random sampling, upsampling of minority classes using random sampling and up-sampling minority classes using the Synthetic Minority Oversampling TEchnique (SMOTE) [17]. SMOTE was first proposed to improve random oversampling by combining two similar linear samples of data from the minority class and therefore producing new data that is similar to the class average but not exactly the same data as already present in the database. Training data for each HI class after up-sampling using SMOTE and down-sampling compared to unmodified data are presented in Fig. 4, where it is seen that there are 86 instances in every class for down-sampling and 8,964 for SMOTE. In addition to data pre-processing, there are other

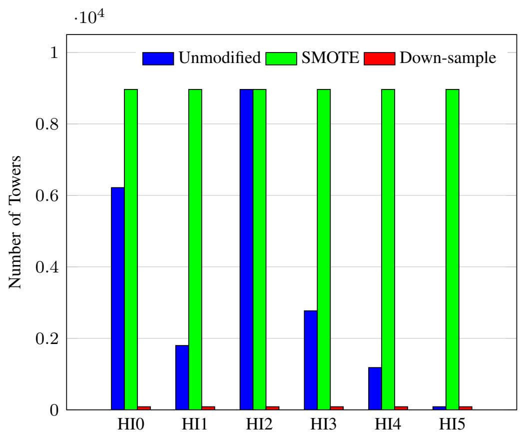


Fig. 4. Model training data for each class after training and testing data split using different re-sampling techniques. Blue is unmodified, green is upsampling using SMOTE and red is down-sampling data.

options for reducing class imbalance issues, which are the balancing of training weights for models such as Logistic Regression (LR) and Support Vector Machine (SVM) or use classification algorithms that handle class imbalance better such as Decision Trees (DT), K-Nearest Neighbor (KNN), Random Forest (RF) and Gradient Boosting (GB) [18]. All those approaches are further tested in Section III.

III. Selection of Prediction Model

Prediction models in this paper are composed using the most common supervised machine learning classification algorithms so that most major approaches are represented. Each algorithm is tested using a number of different hyper-parameters and the three different data sets mentioned in the previous section. All models and data processing are done in the Python 3.7 environment, where models are created using the Scikit-learn module [19]. Classification algorithms used in this paper are Logistic Regression (LR), Support Vector Machine (SVM), Naive-Bayes Classifier (NBC), K-nearest neighbors (KNN), Multi-Layer Perceptron from Neural Networks (NN) and Decision Trees (DT). There are also algorithms that are combined together from multiple methods to convert a set of weak learners to a single strong ensemble model that delivers improved prediction accuracy compared to a single algorithm. In this paper two widespread ensemble algorithms called Random Forest (RF) [20] and Gradient Boosting (GB) [21] are used. The theory and implementation of supervised machine learning algorithms are discussed thoroughly in [22], [23] and [24].

Ensemble methods are used in supervised machine learning to obtain better predictive results from the model by using multiple learning algorithms in a single model rather than using any of the learning algorithms alone. Random forest [20] is an ensemble algorithm that has grown in usage over the last few years because of its great performance. By nature, it is a bagging classification and regression algorithm that is based on decision trees. It is developed on an ensemble of unpruned trees, induced from bootstrap samples of the training data. It uses random feature selection in the tree induction process in addition to bootstraping. A prediction is made by using a majority vote to aggregate the predictions of the ensemble models. As RF loss function is constructed to minimize the overall error rate, it will tend to focus more on the prediction accuracy of the majority class. That will often result in poor accuracy for the minority class result similar to most classifiers when they are trained on imbalanced training data sets. To alleviate the problem, [18] proposes to use balanced or weighted RF models that are both evaluated during hyper-parameter tuning in this paper.

A. Model Selection

For many of the algorithms, there are modifiable parameters called hyper-parameters, such as the number of trees in RF or the regularization strength of an L2 penalty in the loss function of LR, which affect the performance of models drastically. Those parameters and their optimal selection are usually done manually because they rely on specific data and require experimentation to identify appropriate values used for model training. Due to the large amount of possible combinations for each algorithm, this process is extremely computational and time-consuming but essential to increase model performance by selecting best hyper-parameters to maximize the performance of model on actual data. Selection of hyper-parameters is described more detailed in [25].

In this paper model parameter optimization is performed by using random search [26] combined with nested cross-validation to reduce the bias of training data. Nested cross-validation is used to reduce the bias for both hyper-parameter tuning and model evaluation. In terms of computational complexity, relatively simple 5 x 2 setup is used in this paper. That means there is a 5-fold cross-validation implemented in the outer loop and 2-fold cross-validation in the inner loop. The inner loop is responsible for the model selection process, and the outer loop is for estimating the generalization error. The theory behind nested cross-validation and its benefits over simple training and testing split or k-fold cross-validation are thoroughly described in [24] and [25]. According to [26], random search over the same domain is able to find models that are as good or better within a small fraction of the computation time of a pure grid search. This means in terms of computational budgets, a random search finds better models compared to grid and manual searches by effectively searching a larger configuration space. For each model there is a large number of different combinations of hyper-parameters to select the best model for each algorithm. The number of combinations for each models is presented in Table IV, while the implemented random search is limited by 100 combinations for each model.

TABLE IV. Number of hyper-parameter combinations for tested models

Algorithm	Tested models
Logistic Regression	80
Support Vector Machine	400
Naive-Bayes Classifier	-
K-Nearest Neighbor	372
Neural Network	72
Decision Trees	1440
Gradient boosting	576
Random Forest	72

B. Model evaluation

Model evaluation was done through model hyper-parameter optimization using a random search with nested 5 x 2 cross-validation to find the most suitable model for the used data set and algorithm. That enabled the determination of the best hyperparameters for each classification algorithm and therefore their best performance to be compared on actual data. The performance metrics used in this paper for model performance evaluation are accuracy, precision, recall and F1-score, which are calculated using (2) to (5). The One-vs-All approach that enables us to use non-binary classifiers to acquire the Receiver operator characteristic (ROC) [27] and Precision-Recall (PR) curves [28] was left out of the scope because it decreased the performance of the models due to the amplified class imbalance problem.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

where TP, is the number of true positives, FP, is the false positives, TN, is the number of false positives and FN is false negatives from the confusion matrix [28].

C. Performance metrics

An overview of all tested models using four different performance metrics (accuracy, precision, recall and F1) is presented in Fig. 5. It must be noted that precision, recall and F1 scores presented in Fig. 5 are calculated by taking the average of all class values, and therefore they indicate class imbalance better compared to taking the average over all testing samples. It is seen that six models clearly perform better than the other 18 models. Those models are RF, GB, NN, KNN with SMOTE data, RF with SMOTE data and KNN. All those models have an accuracy over 60 percent. The poorest performing models are NBC and LR algorithms, which are the most affected by imbalance of the data. It is also seen that the best overall performance of models was achieved using unmodified data. The SMOTE data set performed better than the down-sample but is comparable with unmodified data in terms of recall and F1 but worse in terms of precision. Down-sampling of training data produced more ambiguous results over all classes, but overall they performed worse than SMOTE and unmodified data. The accuracy of the best performing models was almost 70 percent based on training data. Six models with the best performance are given a more detailed analysis where the performance metrics of each class are presented individually. Fig. 6 presents the precision of the top six models, where it is seen that RF and GB outperform all other models in all classes. It is also seen that NN was not able to detect some classes at all, especially HI = 5. On the recall results graph Fig. 7 it is seen that all models that used SMOTE data sets outperform models with unmodified data in terms of recall values, especially the RF model with SMOTE. RF and GB with unmodified data perform poorly on minority classes compared to RF using SMOTE data and KNN model. The NN model was again unable to detect minority classes. From Fig. 8 it is seen that the best performing models are RF, GB and KNN, where RF with unmodified training data outperformed all other models according to all performance metrics except recall, where the best performing model was RF using SMOTE. Due to the advantages in computational requirements and the facts that RF with unmodified data outperformed other models in nearly all aspects, especially in precision, then it was selected for implementation in the case study. As in precision and recall figures, NN was not able to detect minority classes. The random forest model (hyper-parameters: number of estimators = 100, minimum samples leaf = 1, maximum features = sqrt, class weight = None) with unmodified data is used in Section IV for asset HI prediction and aging behaviour modeling.

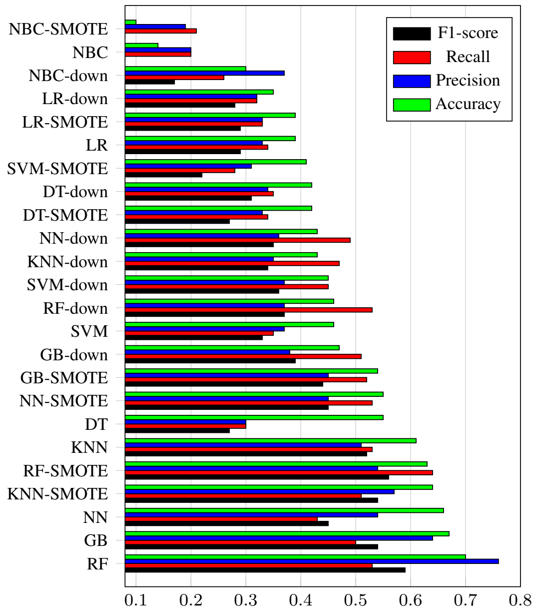


Fig. 5. Performance metrics of each tested model. Black presents accuracy, blue precision, red recall and orange the F1-score.

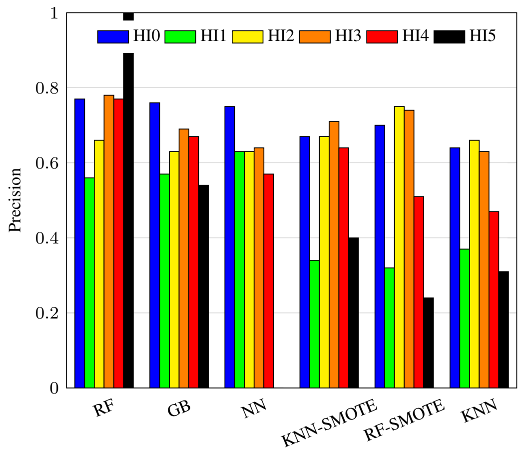


Fig. 6. Precision score for each class of top six models. Each class is presented separately where HI0 is blue, HI1 is green, HI2 is yellow, HI3 is orange, HI4 is red and HI5 is black.

IV. Case Study

The following case study presents two different approaches for methodology implementation. The first case shows the HI prediction for missing assets and the second case the approach for modeling the aging behaviour of assets. For the asset HI prediction, input data is presented in Table V. It can be seen that the training data consists of 26,206 samples and

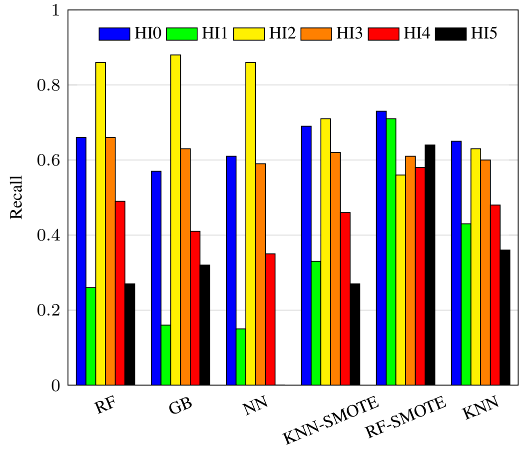


Fig. 7. Recall score for each class of top six models. Each class is presented separately where HI0 is blue, HI1 is green, HI2 is yellow, HI3 is orange, HI4 is red and HI5 is black.

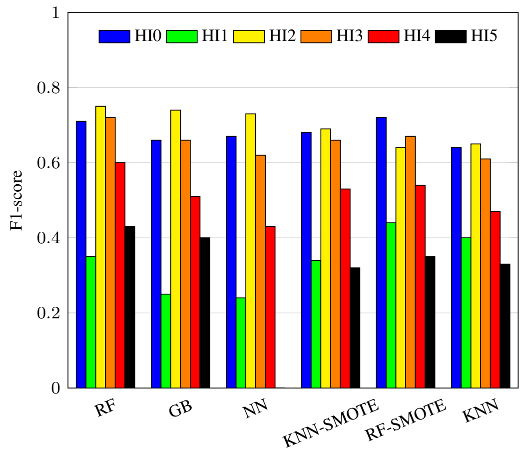


Fig. 8. F1-score for each class of top six models. Each class is presented separately where HI0 is blue, HI1 is green, HI2 is yellow, HI3 is orange, HI4 is red and HI5 is black.

an OHL with 67 towers is tested. The same training data is also used for the asset aging behaviour modeling. The asset HI prediction methodology is tested using the best model with optimal hyper-parameters selected in Section III. The selected model is based on the Random Forest algorithm that will be trained using unmodified data as it provided the best results for the model validation. Hyper-parameters for the model are also selected according to the best results in the model validation process. Input data for the case study is collected in 2018 from periodical visual inspections in Estonian transmission grid using specially designed tablet application based on the methodology described in the Section II. Those inspections produced nearly two million individual defects about OHLs that were aggregated on the tower level. Nearly all towers in Estonian transmission grid were assessed and a single HI value of each tower was generated on the basis of detected defects. That resulted in a situation where nearly all towers had a corresponding HI value in addition to individual asset technical features that were collected from the asset database. For the HI prediction a single OHL with 67 towers that had variety of different tower types was selected. That OHL was constructed and renovated in sections and in different time-frames. To achieve the HI prediction situation, all HI data about the selected OHL was deleted from the HI database while information about technical features remained unchanged. This produced a single OHL with 67 towers with missing HI data as a testing data set and all remaining towers in the grid were used to train the prediction model using the proposed methodology. It is seen from the Table V that there are no towers with HI2, and nearly half of the towers are with HI3. The age of these OHL towers is in the range of 13 to 61 years, and 10 different tower configurations are used.

TABLE V. Training and testing data for case study

Health Index Class	Selected OHL	Training
HI0	15	7757
HI1	0	2249
HI2	2	11203
HI3	32	3430
HI4	16	1461
HI5	2	106
Total	**67**	**26206**

A. *Prediction of Asset Health Index*

A comparison of actual and predicted HI of the selected OHL is presented in Fig. 9, where it is clearly seen that the asset HI prediction model performs well with around an 80 percent accuracy to predict the HI of each tower that it has not seen before. There are a total of 67 towers on that OHL, where 54 of those were predicted correctly and 13 incorrectly. That means the accuracy of the prediction model is much better than randomly classifying towers into six different categories.

B. *Modeling of Aging Behaviour*

For modeling of aging asset behaviour, the age parameter was increased by 10 years while all other technical features remained unmodified. Results from the single OHL example are presented in Fig. 10, where it is clearly seen that for nearly all towers HI values increased compared to the results presented in Fig. 9. It is also seen that for some towers in the range of 50 to 67 years, the HI was increased more than expected in one class according to a linear increase of the HI in the methodology. That is because the model does not just linearly increase the HI of the towers but rather predicts the most probable output based on the data of all towers in the grid. Also it is seen that not a single tower HI decreased in terms of aging and even the maximum HI value 5 remained the same. That reflects a realistic situation where no investments were done in the grid. But there are limitations for the implementation of that methodology. As

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRD.2021.3052721, IEEE Transactions on Power Delivery

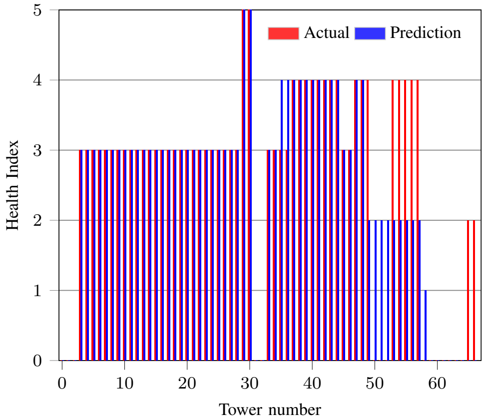


Fig. 9. Comparison of actual and predicted HI for each tower of a real OHL where the blue bars are HI values from field inspections and orange bars predictions from the RF model.

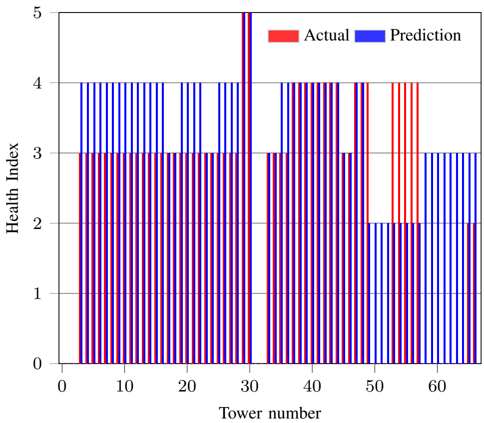


Fig. 10. Comparison of actual and predicted HI for each tower of a real OHL in the case of an extra 10 years of service where the blue bars are HI values from field inspections and orange bars predictions from the RF model.

there are always limits to obtaining a sufficient amount of data and there are impurities in the input data for the grid, this case works well on large data sets where there are a lot of samples from different HI and age ranges. Results of one part of the Estonian TSO's grid is modeled to predict the distribution of HI values in 10 years. That is presented in Fig. 11, where it is clearly seen that towers with HI0 and HI2 have decreased and the number of towers with HI4 and HI5 has increased. The overall number of towers for HI1 and HI3 has remained the same. The age of the grid has increased and concentrations of assets have changed from smaller HI values to larger HI values. That means the overall condition of the grid has decreased and there are more towers with critical or end-of-life than conditions than in present scenario. This reflects the logical aging behaviour of the grid where there are no additional refurbishments or replacements made.

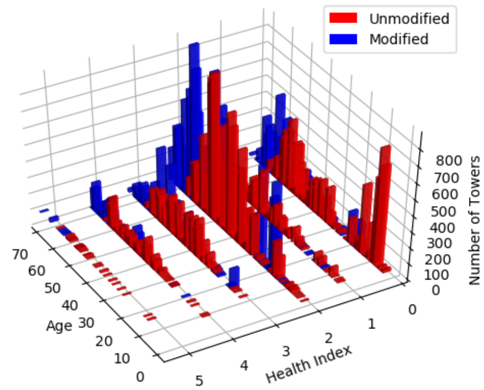


Fig. 11. The 3D chart presents HI and age distribution. Red is the unmodified situation that uses actual HI data periodic inspections and blue the modified situation where HIs in the grid are predicted using modified age parameter with prediction model.

V. Conclusion and Further Work

The health index prediction methodology of overhead transmission lines using supervised machine learning models demonstrated that it is possible to predict missing HI values of high voltage OHL towers based on the asset's technical features and HI results of already existing data. The accuracy here was around 70 percent based on the training data and around 80 percent in the case study. This paper also showed that in addition to missing asset HI predictions, there is the possibility to model the aging behaviour of OHLs using supervised machine learning models on structured technical and HI data. This enables efficiency to be increased in asset management decision-making through the use of more accurate data as an input. It also enables the cost of annual inspections to be decreased because the proposed methodology enables the prediction of the HI values of towers without physically visiting them every year. In addition, if there are a few missing OHLs or towers, then it is possible to predict the HIs of those assets instead of conducting a re-inspection. Even though prediction models have an accuracy of 70 percent, they produce much better results compared to randomly classifying towers into six classes.

For further work, it is recommended to increase the training data of the prediction model by including inspection results from multiple years. Data used in this paper was from a one-year pilot project performed by Estonian TSO. It is also recommended to increase the number of tower features to take into account various parameters that affect aging behaviour of assets. Those parameters can be the distance from roads or sea, soil type or even vegetation in the area. It must be noted that after adding additional features, it is important to re-perform the feature importance analysis to detect features that affect the accuracy of the results the most and to remove features that can be considered as a noise. That enables to use more generalized models to perform more accurate predictions on the data it has not seen before. It is also possible to predict the HI of each OHL component separately by using this methodology but for the simplification purposes this paper used aggregated HI values of towers. Performing the HI prediction of each component separately might produce results where there is an increased prediction accuracy of some components and a decreased accuracy of others because of possible inconsistencies of predefined visual indicators and their corresponding HI values. That approach must be thoroughly investigated in the further studies as there might be a significant potential to increase the accuracy of the methodology but this requires to re-check the basis of HI determination of each OHL component. The component specific HI prediction of OHLs also enables more detailed investment decision-making. Moreover, as there is usually limited to no technical failures in the TSO grid, then all failures and critical defects that cause the tower or component to be refurbished needs to be recorded with corresponding asset technical features at the point where the event happened. That might give better results in predicting aging behaviour and the possibility of the failure of the grid because HI transition phases from one value to another are being recorded. In terms of structural long-term data collection and usage, excellent results may be achieved through the use of the methodology proposed in this paper.

REFERENCES

[1] ENTSO-E, "Facing the replacement wave," Tech. Rep., 2015.

[2] CIGRE WG 37.27, "TB176 ageing of the system—impact on the planning," CIGRE, Tech. Rep., 2000.

[3] CIGRE WG B3.32, "TB660 saving through optimised maintenance in air insulated substations," CIGRE, Tech. Rep., 2016.

[4] CIGRE WG B2.13, "TB175 management of existing overhead transmission lines," CIGRE, Tech. Rep., 1999.

[5] CIGRE WG B2.08, "TB230 assessment of existing overhead line supports," CIGRE, Tech. Rep., 2003.

[6] P. Thongchai, P. Pao-La-Or, and T. Kulworawanichpong, "Condition-based health index for overhead transmission line maintenance," in *10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, May 2013, pp. 1–4.

[7] G. Ricci, R. Bertella, S. Prato, and M. Zippo, "A.ri.el: A structural integrity condition based method for renewal assessment of pylons

[13] Y. Liu, J. Xv, H. Yuan, J. Lv, and Z. Ma, "Health assessment and prediction of overhead line based on health index," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5546–5557, July 2019.

in overhead transmission lines," in *2018 AEIT International Annual Conference*, October 2018, pp. 1–6.

[8] H. Manninen, J. Kilter, and M. Landsberg, "Advanced condition monitoring method for high voltage overhead lines based on visual inspection," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, August 2018, pp. 1–5.

[9] N. F. Stanley and H. F. Heap, "Reliability-centered maintenance. Report Number AD-A066579," United States Department of Defense, Tech. Rep., 1978.

[10] T. Matsui *et al.*, "Development of remaining life assessment for oil-immersed transformer using structured neural networks," in *2009 ICCAS-SICE*, August 2009, pp. 1855–1858.

[11] F. Yang, M. S. Habibullah, T. Zhang, Z. Xu, P. Lim, and S. Nadarajan, "Health index-based prognostics for remaining useful life predictions in electrical machines," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2633–2644, April 2016.

[12] T. Kato and Y. Suzuoki, "Monte-carlo-based modeling of asset management of electric power apparatuses for assessing influence of wide variety of degradation extent and failure probability," in *2009 IEEE 9th International Conference on the Properties and Applications of Dielectric Materials*, July 2009, pp. 232–235.

[14] V. Hodge, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85–126, October 2004.

[15] T. Hoens and N. Chawla, *Imbalanced Datasets: From Sampling to Classifiers*. John Wiley and Sons, Inc, June 2013, pp. 43–59.

[16] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[17] R. Blagus and L. Lusa, "Smote for high-dimensional class-imbalanced data," *BMC Bioinformatics*, vol. 14, no. 1, March 2013.

[18] C. Chen and L. Breiman, "Using random forest to learn imbalanced data," *University of California, Berkeley*, January 2004.

[19] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[20] L. Brieman, "Random forest," *Machine Learning*, vol. 45, no. 5, p. 5–32, July 2001.

[21] J. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis*, vol. 38, pp. 367–378, February 2002.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, 1st ed. Springer, 2007.

[23] E. Alpaydin, *Introduction to Machine Learning*. The MIT Press, 2014.

[24] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, 2nd ed. Springer, 2009.

[25] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning," *ArXiv*, vol. abs/1811.12808, 2018.

[26] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *The Journal of Machine Learning Research*, vol. 13, pp. 281–305, March 2012.

[27] T. Fawcett, "Roc graphs: Notes and practical considerations for researchers," *Machine Learning*, vol. 31, pp. 1–38, 01 2004.

[28] S. Marsland, *Machine Learning: An Algorithmic Perspective, Second Edition*, 2nd ed. Chapman and Hall/CRC, 2014.

Appendix 5

V

H. Manninen, C. J. Ramlal, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques," *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106726, 2021

Electrical Power and Energy Systems 128 (2021) 106726

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

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

Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques

Henri Manninen [a,*], Craig J. Ramlal [b], Arvind Singh [b], Sean Rocke [b], Jako Kilter [a], Mart Landsberg [a]

[a] Tallinn University of Technology, Department of Electrical Power Engineering and Mechatronics, Estonia
[b] The University of the West Indies, Department of Electrical & Computer Engineering, Trinidad and Tobago

ARTICLE INFO

Keywords:
Condition assessment
Deep learning
Health index
Multi-classed object detection
Transmission lines

ABSTRACT

Electrical Transmission System Operators (TSO) are trusted with ensuring the safety and reliability of transmission infrastructure which can span thousands of kilometers. Maintenance of such a geographically expansive system is naturally a matter of concern and companies invest heavily in tracking infrastructure state which still relies predominantly on visual inspection. This paper presents an automated condition assessment methodology for concrete poles supporting overhead conductors based on deep learning object detection networks. Nine defect conditions ranging from incipient to severe are automatically detected from infrastructure photographs and mapped onto established Health Indices used by maintenance personnel. Three different deep learning networks are tested and new metrics, specific to this problem, are defined to evaluate their performance based on asset Health Index (HI) values. Results indicate that deep learning object detection networks hold promise for significantly reducing manual labour associated with visual inspection, especially when combining with automatic asset identification based on image geotag. This paper shows acceptable performance on more severe defect types.

1. Introduction

Overhead transmission lines (OHL) are among the most important assets in the electricity system and usually span thousands of kilometers. They consist of conductors and hundreds of individual poles that support the electrical circuits with expected lifetimes in the order of 60 to 80 years. As the infrastructure ages, it is met with varying environmental conditions which can cause unexpected degradation and compromise the reliability and security of the system. Proper asset management demands regular monitoring of the physical condition of OHLs which is usually done by periodic visual inspections, carried out by trained field crews who visit each pole and OHL's right-of-way. However, it is not economically feasible to carry out inspections regularly. Even annual inspections may be a strain on the TSO budgets. Compounding this problem is the ineffectiveness of statistical methods for lifetime estimation given the paucity of data arising from historically low transmission infrastructure failures.

There are several technical publications about OHL condition assessment and determination of critical defects such as [1–7] which focus on the detection of immanent failure. Others, [8–10], develop a Health Index (HI) system based on a wider set of pole conditions in order to estimate the likelihood of asset failure in a certain time frame. While they developed specific, predetermined criteria lists for each component of OHL, the approach still requires data collection via the foot patrol inspections.

In the last decade, there has been an increase in the use of aerial vehicles such as helicopters or drones to carry out inspections. Aerial inspections are usually performed with cameras and LiDAR [11] to acquire 3d models, photos and geographical information such as point cloud of the OHL and its right-of-way as described in [12]. Unfortunately those inspections usually lack the input about OHL defects on the level that foot patrols provide and the main focus is on the vegetation and conductor ground clearance analysis instead of technical condition assessment. In [13] object detection based on video and images from aerial surveys are used to detect line components but no assessment of condition is performed.

* Corresponding author at: Ehitajate tee 5, 19086 Tallinn, Estonia.
E-mail addresses: henri.manninen@taltech.ee (H. Manninen), craig.ramlal@sta.uwi.edu (C.J. Ramlal), arvind.singh@sta.uwi.edu (A. Singh), sean.rocke@sta.uwi.edu (S. Rocke), jako.kilter@taltech.ee (J. Kilter), mart.landsberg@elering.ee (M. Landsberg).

https://doi.org/10.1016/j.ijepes.2020.106726
Received 18 June 2020; Received in revised form 5 October 2020; Accepted 13 December 2020

This paper examines the applicability of deep learning object detection networks in detecting degradation artefacts on tubular type concrete transmission line poles and on steel lattice tower foundations. Recently, deep learning object detection methodologies in the field of computer vision, has become a hot topic of research. Object detection is the technique of both recognizing an object class and predicting the location of the object via a bounding box. These techniques have been applied in the fields of medicine [14], intelligent vehicles [15], agriculture [16] and in damaged aerial power lines [17,18]. Current deep learning object detectors can be categorized into multi-stage region proposal techniques [19–21] and in single-stage global regression algorithms [22,23]. In this study a region proposed object detection network: Faster Region-based Convolutional Neural Network (Faster R-CNN) [24], and two global regression networks: You Only Look Once (YOLOv2) [25] and the Single Shot Detector (SSD) [26] are the chosen detectors for the automatic condition assessment problem. These detection networks find a compromise between training and detection speed and accuracy and are some of the more popular detection networks currently used in academia and industry.

For the present study visual inspection data by foot patrols from the Estonian Power System is used in lieu of aerial survey data to illustrate the efficacy of the methods. The produced base neural network from this study can be adapted using transfer learning as high resolution aerial survey data (>100 megapixel images) becomes more available. Therefore this study is the first step in bridging the gap from manual to automated condition assessment of concrete pole infrastructure. This paper is divided into three main sections where Section 2 explain the background of the methodology, Section 3 describes object detection models used in this paper and Section 4 presents results of used object detection models for individual defect detection and HI assessment on pole level.

2. Methodology

The overview of proposed methodology is presented on Fig. 1 and can be divided into four main processes:

1. Pole selection
2. Image acquisition
3. Image-Asset correlation
4. Defect detection and Health Index Mapping

The process starts with a specific subset of poles selected for inspection based on the maintenance policy of the Utility. For this subset, a set of corresponding images must be taken using UAVs or handheld devices. Geo-tagged images can be automatically correlated to assets but those that are not, must be manually associated with their relevant assets. Once all images are correctly associated with their assets they move to the object detector for defect detection. The detected defects are then classified according to severity using a health index and the results saved in the asset database.

2.1. Pole selection

The set of structures selected for inspection is a function of the time since the last inspection, the last recorded infrastructure state, criticality of the line to system security and any external events that may be cause for concern. Typically, assets with critical condition or high risks are more frequently assessed than assets that have a low impact on system reliability or are in good condition.

2.2. Image acquisition

Clear images of sufficient quality must be taken at multiple angles so that the entire structure can be assessed. In this paper images for the object detection models training and testing data are taken from periodic visual inspections by foot patrols.

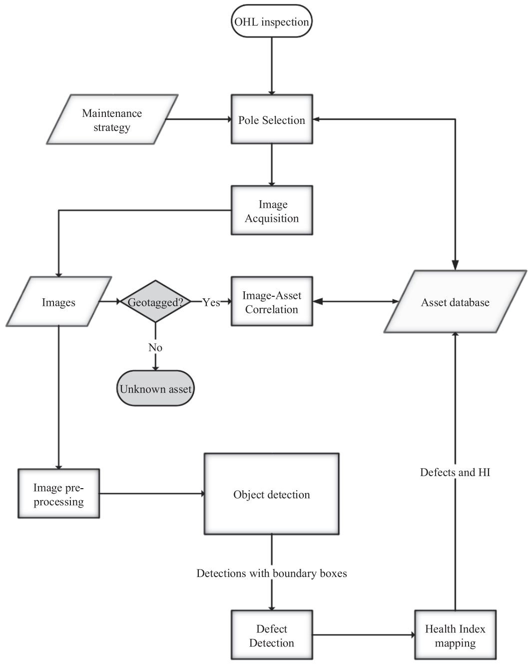


Fig. 1. Principle scheme of automatic condition assessment model where rectangles represent process steps, parallelogram data, rounded shapes start and end of the process and rhombus decision point. The list of poles that are inspected is based on the restrictions from the maintenance strategy.

2.3. Image-asset correlation

This subsection describes the method of linking concrete pole images that are taken via foot patrols with their physical asset, which is stored in a database platform. Geotagging ensures that the latitude and longitude coordinates at the location of the foot patrol are saved with the concrete pole's images. This geotagged data together with the LiDAR information from the asset database can be used to calculate the distances between each image location and each asset location. First a list of potential candidates of concrete pole images mapped to the physical assets are generated, by matching the minimum distance of each image to that physical asset. The distance between two different points on the Earth's sphere can be accurately calculated using Haversine formula (1) [27]:

$$D_{real} = 2 \times r \times \arcsin \times \sqrt{\sin^2 \times} \times \sqrt{\left(\frac{lat_{asset} - lat_{image}}{2}\right) + \cos(lat_{image}) \times} \times \sqrt{\cos(lat_{asset}) \times \sin^2\left(\frac{lon_{asset} - lon_{image}}{2}\right)}, \quad (1)$$

where D_{real} is the distance between two points, lat and lon are the latitude and the longitude and r is the radius of the Earth ($r = 6731$ km). It should be noted that lon and lat should be used in the same coordinate systems. In this example all coordinates in the database are in $EPSG: 3301$ coordinate system and geotags are in $EPSG: 4326$ system. Therefore to use (1), all coordinates are converted to $EPSG: 4326$ first.

Next, to verify each potential mapped candidate, the distances must be less than a specified threshold, D_{max}. Filtering the mapped candidates in this way, reduces errors that may be incurred from erroneous GPS data.

In order to determine an appropriate value for D_{max} three factors must be considered: The distribution of tower spans; the distance that images are being taken from the poles and GPS error. If D_{max} is chosen to be too large compared to the tower spans then more than one asset may be in the radius of the image. If D_{max} is chosen to be too small then no asset may be in the radius due to the distance at which the picture is being taken. The presence of GPS errors place additional constraints on the value of D_{max} since images may appear to be at incorrect locations and may be associated with the wrong pole. D_{max} must be chosen small enough so that a GPS error is not likely to place an image of one pole in the valid space for another pole. This may mean that there are some poles which are not associated with any images and these would have to be manually associated. Practically, the value of D_{max} was determined empirically by looking at the distribution of image-tower distances compared to tower-tower distances and choosing a value which balances the number of poles falling within the distance with the number of pole spans that were smaller than the distance.

High voltage OHLs are considered as open areas for GPS receivers since tall vegetation must be cut in the right-of-way to prevent flash-overs from conductor to vegetation. This is beneficial in terms of GPS position accuracy. Most smartphones and tablets use assisted GPS (A-GPS), this technology uses the mobile networks' signals in addition to satellites to receive the geographical position of the device. This is beneficial in environments where it is difficult to receive a GPS signal but may cause more errors in open fields than regular GPS that uses at least four satellites to determine the device's position. Based on [28], an open field's A-GPS mean error can be estimated at 4 meters, the consumer-grade GPS's mean error is under 2 meters. In both cases, the standard deviation was approximately 2 meters. As state-of-the art smartphones are using dual-frequency GPS receivers that have better accuracy than A-GPS, it is possible to conclude that A-GPS's parameters should be used as the worst-case-scenario as it relates to calculating minimum distances.

2.4. Defect detection and health index mapping

Defects are detected using trained deep learning object detection networks and are discussed in more detail in Section 3. Once the defects are detected, they must be combined and mapped to a single HI for a pole. The theory of concrete pole condition assessment and usage of HIs in this paper are based on [10] which was implemented on the Estonian transmission grid in 2018. Predefined visual indicators correlated to HIs, Table 1 and Fig. 2, reduce the variability associated with unstructured appraisals by assessors typical of traditional inspection. The HI values are based on the impact to overall reliability of the pole and therefore, HI calculation is based on maximum function of all determined defects according to (2).

$$HI = \max\{HI(x) : x = 1 \ldots n\} \quad (2)$$

Table 1
List of visual indicators presented in Fig. 2 and responding HI values.

Visual indicator	Health Index
(Hole Fig. 2: a)	5
Loss of cross section > 20% (Fig. 2: b)	5
Concrete is falling off (Fig. 2: c)	5
Loss of cross section < 20% (Fig. 2: d)	4
Crack (Fig. 2: e)	4
Visible reinforcements (Fig. 2: Fig. 2: f)	3
Micro longitudinal cracks (Fig. 2: g)	2
Hair-like cracks (Fig. 2: h)	1
Other minor Defects (Fig. 2: i)	1
No visible defects	0

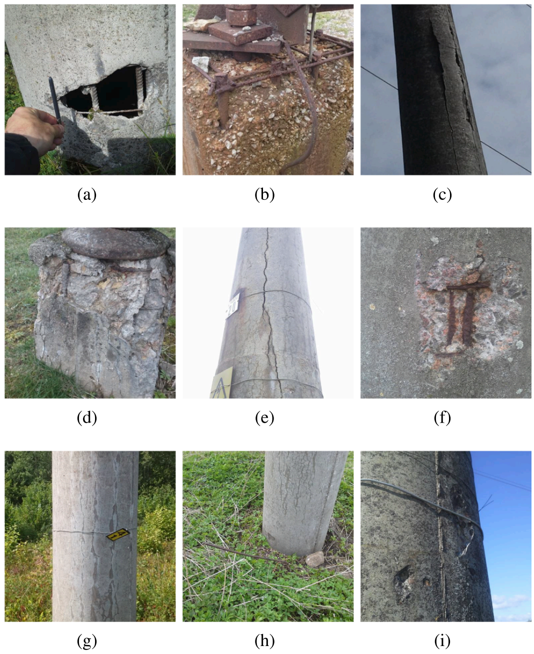


Fig. 2. Examples of reinforced concrete pole and foundation visual indicators listed in Table 1.

where $HI(x)$ is a function that looks for the HI value of selected defects in Table 1 and n is the total number of defects in the object detection model's output. For example, according to Table 1, a "crack" corresponds to HI 4 and "micro longitudinal cracks" to HI 2. By using the maximum function in (2), the overall HI of that asset is 4.

3. The deep learning object detection networks

This section describes the neural network architectures, data, methods of preparing the images for training and testing, and the various performance metrics used in this study. Three object detection networks are used with differing detectors and feature extraction layers. They include the Faster R-CNN network, the YOLOv2 network and the SSD network. For all networks, transfer learning was performed on the same dataset to learn the features of the concrete pole defects and the hyper-parameters were tuned using a Manual Search tuning method.

3.1. Network architectures

3.1.1. The Faster Region-based Convolutional Neural Network (Faster R-CNN)

The Faster R-CNN deep learning algorithm [24] uses a two-stage framework that at first scans the image and then focusses on regions of interest. The faster-RCNN technique advances from fast-RCNN since, it does not rely on additional methods to generate a candidate pool of isolated region proposals, thereby reducing the computational complexity of the algorithm and time of training.

The faster-RCNN architecture comprises of a feature detection network, a Region Proposal Network (RPN) and a classifier. The RPN is a fully-convolutional network that generates proposals and can be trained via supervised learning techniques. In this study the feature detection network chosen is the Inception V2 model [29] which consists of 6 convolutional layers, 2 max pooling layers, 3 Inception modules, 1 linear and 1 softmax layer. The training technique chosen was the stochastic gradient with momentum (SGDM) method. SGDM accelerates convergence by replacing the actual gradient by an estimate, calculated from a randomly selected subset of the data. The momentum, minibatch size and initial learning rate were chosen as 0.9, 1 and 0.0002 respectively.

3.1.2. The You Only Look Once (YOLO) v2

The YOLO v2 technique [25] is a real-time algorithm that accomplishes object detection via fixed-grid regression. Whilst region proposal frameworks such as fast/faster-RCNN, have several correlated stages, that are each trained separately. YOLOv2 is a one-stage framework that maps the image pixels to bounding box coordinates and class probabilities in a single step using a global regression technique. The idea is to make objection prediction on each feature map location without the cascaded region classification step.

The feature extraction network consists of 23 convolutional, 5 max pooling, 2 routing and 1 reorganization layer. The K-means clustering algorithm using the Intersection-over-Union distance metric is given by:

$$IoU_d = \frac{|A \cap B|_2}{|A \cup B|_2} \tag{3}$$

This method was used to generate the anchor box sizes, where A and B denote the ground truth and estimated bounding boxes, respectively. The Adam stochastic optimization technique [30] was chosen for training this network. Adam is an adaptive learning rate optimization algorithm that combines the benefits of RMSprop and Stochastic Gradient Descent with momentum. The momentum, minibatch size, epochs, initial learning rate and decay factor were chosen as 0.9, 4, 160, 0.001 and 0.0005 respectively.

3.1.3. The single shot detector

The SSD object detection system [26] can be classed as a single stage network, similar to that of YOLOv1. SSD however, aims to overcome some of the inabilities of the YOLOv1 algorithm namely detecting small objects in groups, and with certain data sets is more accurate and faster than the YOLOv1 algorithm.

In this study the feature extraction network for the SSD algorithm uses the RetinaNet backbone [31]. The RetinaNet architecture consists of a Feature Pyramid Network [32] on top of a feedforward Residual Network [33]. This topology has the benefit of using a focal loss feature which deals with the class imbalance problem. For the focal loss feature values of $\gamma = 2$ and $\alpha = 0.25$ were chosen for this study. The Stochastic Gradient Descent with Warm Restarts (SGDR) optimization technique [34] was used for training this network. SGDR is a variant of learning rate annealing, that aids in improving the rate of convergence and anytime performance. The momentum, minibatch size, base learning rate, warm up learning rate and warmup steps were chosen as 0.9, 64, 0.04, 0.01333 and 2000 respectively, for 25000 total steps and a cosine decay learning rate.

3.2. Data and performance metrics

3.2.1. Data pre-processing

The data consisted of foot-patrol concrete pole images with no image requirement standardization. Therefore aspects such as resolution, background, lighting, pole orientation, distance, shadow and presence of external objects varied considerably. To prepare the images for use by the object detection networks, ground-truth bounding boxes for each class defect were placed on the original images by an expert. All images were then resized to 512×512 pixels, to reduce computational complexity of neural network training. The input neuron size of the feature extraction layer of all object detection algorithms used in this study accepts images of the resized format. It is important to note that all bounding boxes are saved as relative coordinates on images, such that after object detection, the bounding box edges can be resized to map to the original un-scaled images. This enables ease of follow-up inspections by experts, where they can compare the bounding boxes of the ground-truth and detector on the original image. The dataset comprised of 1008 images with a total of unique 3,544 features. Each feature can be classed into one of 9 classes, and the feature decomposition of the training and test datasets are given in Table 3. A hold-out cross validation method was used to train and test the detection networks. From the dataset, 80% of randomly chosen data was used for training with 2,853 defects, and 20% allocated for testing, making up the remainder 691 features.

3.2.2. Performance metrics

The *Intersection over Union (IoU)* performance measure gives the similarity between the predicted region and the ground-truth region for an object present in the image. It is defined as the size of the intersection divided by the union of the two regions.

$$IoU = \frac{\text{area}(B_p \bigcap B_{gt})}{\text{area}(B_p \bigcup B_{gt})} \tag{4}$$

where B_p is the predicted bounded box and B_{gt} is the ground truth bounding box. The IoU metric is used in the calculations to determine True (False) Positives (Negatives). *Precision or confidence*, is a measure of the proportion of predicted positive cases that are correctly classified as real positives and is given by:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall or sensitivity, is a measure of the proportion of real positive test cases that are correctly predicted positive. Eq. (6) gives the formula for recall.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

A *Confusion Matrix* is a useful metric to give the outcome of classification in image recognition problems. Each row denotes the instances of an actual class and each column gives the instances of the prediction. Confusion matrices for object detection algorithms are evaluated in a similar way, however they utilise multi-class instances where the matrix compiles each object class from the same image on a single table.

In this study, it is more important to analyse the network's performance in relation to estimation of the health indices, which are derived from the true(false) positives(negatives) metrics. We define two general performance metrics for the object detection algorithm as overestimation and underestimation scores, given in (7). Overestimation O_i is a measure of how much the network detects and classifies defects as higher fault indices than they actually are, and underestimation U_i is the converse.

$$O_i = \frac{\sum_{j=i+1}^{5} a_{ij}}{\sum_{j=1}^{5} a_{ij}} \times 100, \quad U_i = \frac{\sum_{j=1}^{i-1} a_{ij}}{\sum_{j=1}^{5} a_{ij}} \times 100 \tag{7}$$

where a_{ij} is the element in the i-th row and j-th column of the confusion matrix. A confusion matrix is then generated with these scores as shown in Table 2. The rows of the table gives the true Health Index, while the column gives the estimated Health Index. The diagonal of the table, λ_i, therefore, represents the correctly estimated indices, the lower triangle

Table 2
Confusion matrix based on estimated and actual Health Index.

		Estimated		
		HI 1	HI 2	HI 3
Actual	HI 1	λ_1	$a_{1,2}$	$a_{1,3}$
	HI 2	$a_{2,1}$	λ_2	$a_{2,3}$
	HI 3	$a_{3,1}$	$a_{3,2}$	λ_3

the under-estimated health indices and the upper triangle, over-estimated health indices.

4. Results and discussion

4.1. Image-asset correlation

For the image-asset correlation, all minimum distances of 16823 individual OHL towers to the nearest physical asset were calculated. Next, 1871 images taken from periodic visual inspections of the Estonian transmission grid were used to determine an appropriate value for D_{max}. Fig. 4 shows the cumulative distribution of the tower to tower distances. The majority (71 percent) of OHL towers in the grid, are more than 100 meters from their nearest neighbour, while 4% are less than 30 m apart and less than 1% are within 20 m of the nearest tower.

Each OHL tower's geographical coordinates were accessed from LiDAR inspections, and each tower's images were taken by foot patrols using tablets that enabled geotagging of images using the A-GPS technology. The distance between each image and all assets were calculated using 1 and the minimal separation used to associate an image to its associated asset, Fig. 3. The majority (89 percent) of images were identified to be in the range of 20 meters of the corresponding assets. 82 percent of images were taken less than 15 meters from the assets and 60 percent of the images were less than 10 meters. Around 9 percent of images were further than 100 meters from the OHL tower (up to 5 km in some cases). The reason for these large offsets is likely due to a GPS drift on tablets or GPS signal lost while taking the picture. Image-asset correlation using tablets and foot patrols can be considered as baseline accuracy when compared to more precise technologies that are widespread with modern drones. However, even using input data from tablets, a threshold value of D_{max} for image-asset correlation in the range of 10 to 15 meters could be used for asset identification with an accuracy of about 90%.

4.2. Performance evaluation of detectors

To evaluate the performance of the network, the test dataset is processed by the detector and compared to the ground truth data (manual inspection). The detections are then classified into three categories:

- True Positives (TPs) - Where the network has correctly identified a defect

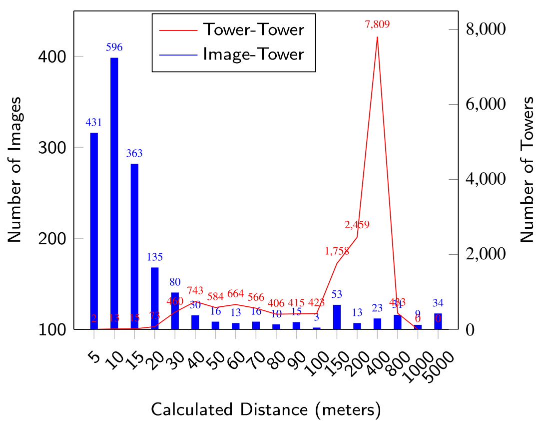
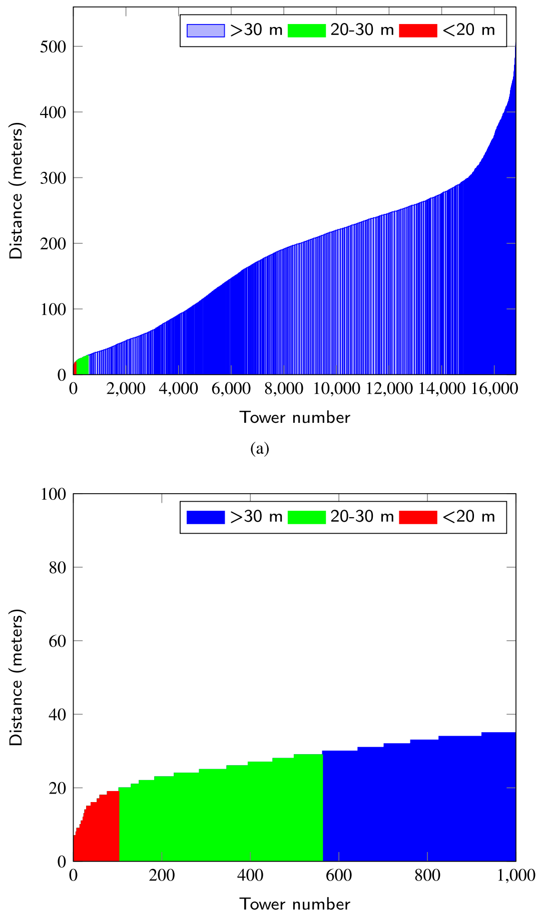


Fig. 3. Calculated minimal distance between taken images and towers in blue with axis on the left and minimal distances between two towers in red with axis on the right.



(a)

(b)

Fig. 4. Cumulative distribution of minimal distances between two nearest towers in the grid (a) and the nearest 1000 towers (b). Blue represents towers that have a distance of 30 m or greater from each other, green towers are within 20 to 30 meters from their nearest neighbour and red towers have a minimum distance of less than 20 meters from each other.

Table 3
Composition of training and testing data by defect type.

Defect	# in Training	# in Testing
Hole	66	18
Loss of cross section > 20%	41	9
Concrete is falling off	289	72
Loss of cross section < 20%	69	20
Crack	356	76
Visible Reinforcements	643	174
Micro-longitudinal cracks	418	89
Hair-like cracks	351	78
Other minor Defects	620	155

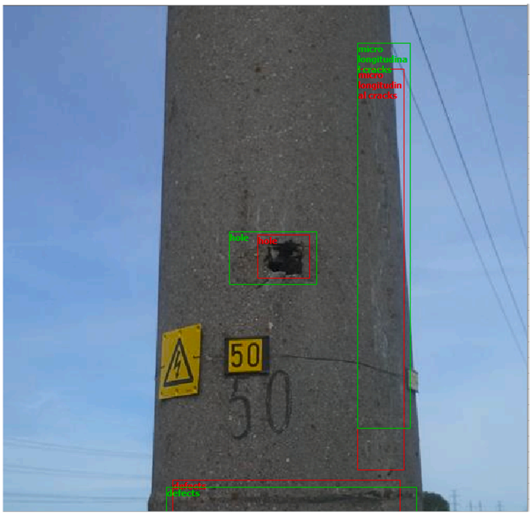


Fig. 5. Bounding boxes overlayed on a concrete pole image, where green is the ground truth and red is Faster-RCNN detector.

- False Positives (FPs) - Where the network has detected a defect where none existed
- False Negatives (FNs) - Where the network has failed to detect an existing defect

The aggregated values in these categories are then used to calculate the precision and recall metrics as given in (5) and (6) respectively, giving an indication of the accuracy of the object detector. Once the detectors are trained, the test data in Table 3 is input into the detection algorithms and results are recorded and compared to the ground-truth dataset. Fig. 5 gives an example concrete pole image, showing the overlays of bounding boxes of the ground truth data represented in green, and the bounding boxes of the Faster-RCNN detector in red. Three different faults namely, holes, micro-longitudinal cracks and minor defects were successfully detected by the Faster-RCNN algorithm for an *IoU* of 0.1. The example, gives a good indication of a detector's performance to an image with multiple classes of features. Several proposal detections may be present in the vicinity of a single ground-truth object. To compensate for this, the proposal with the highest confidence score is usually selected through a process called Non-Max Suppression [35]. Since the number of detections is reduced through this process, it has the potential to reduce the number of false positives and increase the number of false negatives which will affect the Precision and Recall metrics respectively.

To determine whether a detection is a True Positive, False Positive or False negative its bounding box is compared to all ground truth

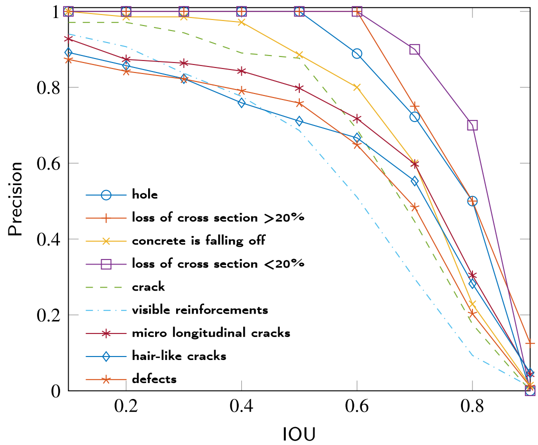


Fig. 6. Precision vs IoU for YOLOv2 detector.

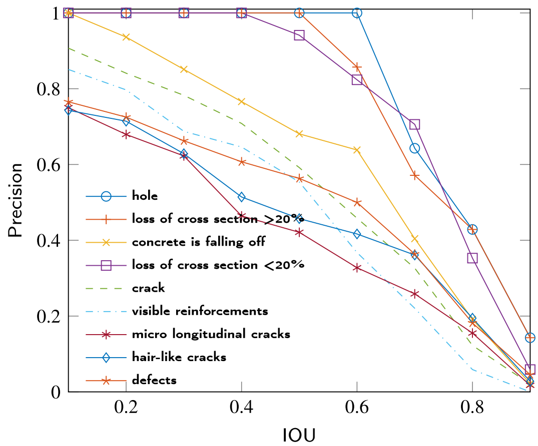


Fig. 7. Precision vs IoU for SSD detector.

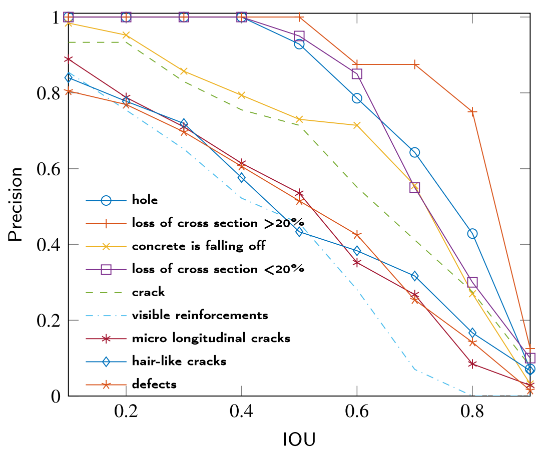


Fig. 8. Precision vs IoU for RCNN detector.

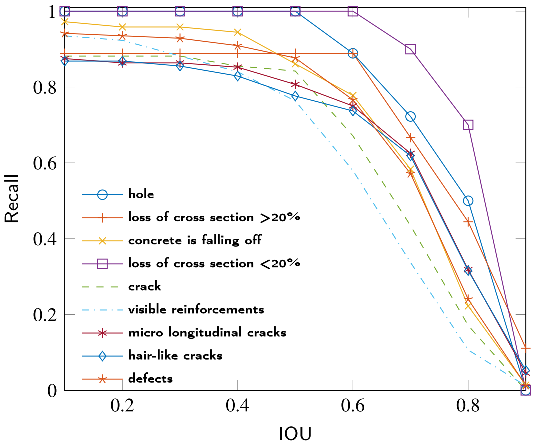


Fig. 9. Recall vs IoU for YOLOv2 detector.

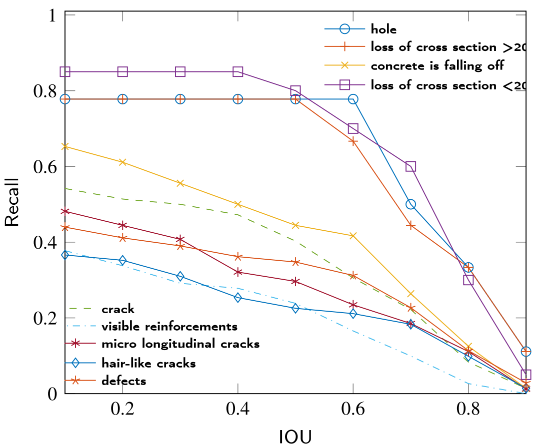


Fig. 10. Recall vs IoU for SSD detector.

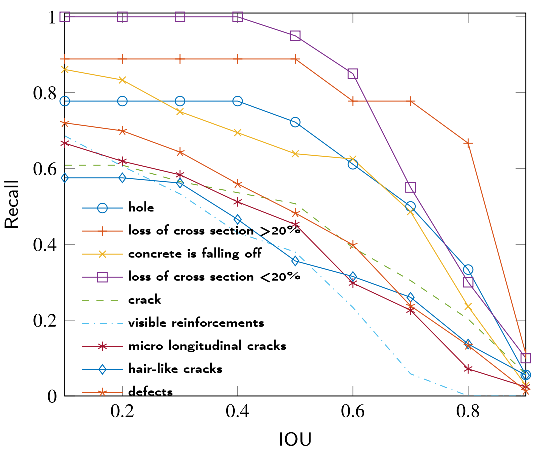


Fig. 11. Recall vs IoU for RCNN detector.

bounding boxes of the same class in the image. An Intersection over Union (IoU) threshold (between 0 and 1) for the bounding boxes is used to determine whether the detection corresponds to the ground truth. For the present study, sweeps of Non-max suppression and IoU thresholds were done between 0.1 and 0.9 for each of the nine classes resulting in a 9x9x9 array for each one of the detector networks (RCNN, YOLOv2, SSD). Based on examination of the arrays, the Non-max suppression was found to have negligible effect. For each of the classes, the graphs of Precision and Recall vs IoU are presented for each of the detector networks, Figs. 6–11.

Figs. 6–11 indicate that the networks gives the best precision-recall performance at low values for IoU thresholds. This is likely due to the irregular shapes and obscure edges of the defects which leads to considerable subjectivity in properly setting the boundaries of bounding boxes of the ground-truth data. This is known as noisy labelling of ground truth data and its effects have been studied for image classification problems in [36,37]. In this study, noisy ground-truth data leads to, among other things, mismatches in ground truth and detected bounding box sizes and centre locations. Defects that are detected may show up with bounding boxes shifted from their ground truth counter parts and with different sizes. While an IoU threshold of 0.5 is usually considered standard for objects with well defined boundaries, a significantly lower IoU of 0.1 was found to perform best for the defect detection problem. Following from [10], the defects detected map directly to health indices for the pole as given in Table 1. These indicies are used by Maintenance Engineers to determine what action should be taken with the Asset. A health index of 4 or above indicates immediate attention is required, while 3 and below indicate that the pole requires only monitoring. Multiple defects may be present on a pole leading to a variety of Health Indices associated with that pole. In such cases the highest index (most severe damage) is taken as the overall Health Index for the pole. By comparing the ground truths and detected defects for each pole, using the precision and recall calculations in (5), (6) and the overestimation and underestimation in (7) a confusion matrix of health indices can be generated for each network, Table 4–6. Table 7 summarizes the overestimation and under estimation scores for each network. Two important details show up through this analysis. Firstly, the networks perform better for more serious defects. This is expected since the more severe faults such as holes show more definite changes in the pole as when compared to minor defects such as micro-longitundinal cracks. Secondly, the networks have a tendency to overestimate instead of underestimate. Although over-estimation leads to additional time spent manually double checking pictures for faults, underestimation may see serious damage if left unchecked. As evident, the networks have a very low under-estimation score which is important to ensure that compromised infrastructure is not ignored.

The YOLOv2 detector shows considerably better performance than both the RCNN and SSD detectors with Precision and Recall over 85% for all categories of objects. Additionally when mapped to the HIs, YOLOv2 underestimates level 4 damage in less than 4% of cases and level 5 damage less than 2% of cases. This can be attributed to YOLOv2 scanning the entire image as opposed to regions therefore extracting more contextual information for each bounding box prediction. YOLOv2 convolutional backbone architecture was pre-trained on higher resolution images from ImageNet and therefore the weights are more sensitive

Table 4
Health Index Confusion Matrix for RCNN.

		Estimated HI				
		HI 1	HI 2	HI 3	HI 4	HI 5
Actual HI	HI 1	1	5	7	4	13
	HI 2	1	5	1	7	5
	HI 3	0	1	9	3	23
	HI 4	1	0	0	14	11
	HI 5	0	0	2	4	83

Table 5
Health Index Confusion Matrix for SSD.

		Estimated HI				
		HI 1	HI 2	HI 3	HI 4	HI 5
Actual HI	HI 1	11	3	2	2	12
	HI 2	0	7	2	6	4
	HI 3	4	2	7	8	15
	HI 4	0	1	2	19	4
	HI 5	7	2	2	9	69

Table 6
Health Index Confusion Matrix for YOLOv2.

		Estimated HI				
		HI 1	HI 2	HI 3	HI 4	HI 5
Actual HI	HI 1	19	2	3	0	6
	HI 2	1	13	4	1	0
	HI 3	2	1	23	1	9
	HI 4	0	0	1	22	3
	HI 5	0	0	1	0	88

Table 7
Health index overestimation and underestimation percentages.

	Overestimation (%)			Underestimation (%)		
	RCNN	SSD	YOLO	RCNN	SSD	YOLO
HI 1	96.7	63.3	36.7	0	0	0
HI 2	68.4	63.2	26.3	5.3	0	5.3
HI 3	72.2	63.9	27.8	2.8	16.7	8.3
HI 4	42.3	15.4	11.5	3.8	11.5	3.8
HI 5	0	0	0	6.7	22.5	1.1

to capturing fine-grained information such as ill-defined defect edges and incipient fault conditions such as micro-cracks.

5. Conclusion

The study indicates that even with modest amounts of data, using geotagging methods to identify towers via images captured near tower locations, and using deep learning for object detection can be quite effective for automatic asset condition monitoring. This problem differs from traditional object detection applications, since there is a high degree of artefact irregularity and an inherent amount of ground truth noise due to subjectivity in bounding the defected objects. New metrics for assessing rank-ordered objects, namely overestimation and underestimation percentages have been defined to give a more accurate picture of the practical applicability of the networks. Although this study presents object detection with foot patrol data, the image resolution is relatively low and therefore, the images can be substituted with those from ultra high-resolution images taken from fly-by missions. The results of this study implies that the combination of both data sources and automatic image detection can revolutionize transmission pole monitoring techniques by significantly reducing manual inspection time and cost.

CRediT authorship contribution statement

Henri Manninen: Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Formal analysis, Investigation, Software. **Craig J. Ramlal:** Validation, Methodology, Writing - original draft, Writing - review & editing, Formal analysis. **Arvind Singh:** Validation, Formal analysis, Writing - original draft, Writing - review & editing. **Sean Rocke:** Supervision, Validation. **Jako Kilter:** Supervision, Validation, Writing - original draft, Visualization. **Mart Landsberg:** Supervision, Validation.

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.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ijepes.2020.106726.

References

[1] CIGRE WG B2.08. TB230 assessment of existing overhead line supports. CIGRE, Tech. Rep., 2003.
[2] CIGRE WG B2.13. TB175 management of existing overhead transmission lines. CIGRE, Tech. Rep., 1999.
[3] CIGRE WG 37.27. TB176 ageing of the system–impact on the planning. CIGRE, Tech. Rep., 2000.
[4] Eyre-Walker R, Howarth G, Ahmed R, Lewin J. Application of advanced condition assessment and asset management techniques on steel tower overhead line electricity networks. IET, 2014.
[5] Havard D, Bissada M, Fajardo C, Horrocks D, Meale J, Motlis J, et al. Aged acsr conductors. ii. prediction of remaining life. In: Proceedings of the 1991 IEEE Power Engineering Society Transmission and Distribution Conference. IEEE; 1991.
[6] Tsimberg Y, Lotho K, Dimnik C, Wrathall N, Mogilevsky A. Determining transmission line conductor condition and remaining life. In: 2014 IEEE PES T&D Conference and Exposition. IEEE; 2014.
[7] Gusavac S, Nimrihter M, Geric L. Estimation of overhead line condition. Electric Power Syst Res 2008;78(4):566–83.
[8] Cardenas LA, Alvarez DL, Rodolfo Garcia S, Herrera FA. Overhead lines assessment using health index. In: 2019 FISE-IEEE/CIGRE Conference - Living the energy Transition (FISE/CIGRE), 2019, p. 1–6.
[9] Yingyi L, Jiahuan X, Haiwen Y, Jianxun L, Zhao M. Health assessment and prediction of overhead line based on health index. IEEE Trans Industr Electron 2019;66(7):5546–57.
[10] Manninen H, Kilter J, Landsberg M. Advanced condition monitoring method for high voltage overhead lines based on visual inspection. In: 2018 IEEE Power and Energy Society General Meeting (PESGM). IEEE; 2018.
[11] Spinhirne JD. Micro pulse lidar. IEEE Trans Geosci Remote Sens 1993;31(1):48–55.
[12] Brunker SW. Aerial lidar for overhead transmission lines – lessons from the usa in realising benefit and value. In: 2016 IEEE International Conference on Power System Technology (POWERCON); 2016. p. 1–6.
[13] Li WH, Tajbakhsh A, Rathbone C, Vashishtha Y. Image processing to automate condition assessment of overhead line components. In: 2010 1st International Conference on Applied Robotics for the Power Industry; 2010. p. 1–6.
[14] Han C, Kitamura Y, Kudo A, Ichinose A, Rundo L, Furukawa Y, et al. Synthesizing diverse lung nodules wherever massively: 3d multi-conditional gan-based ct image augmentation for object detection. In: 2019 International Conference on 3D Vision (3DV). IEEE; 2019. p. 729–37.
[15] Guindel C, Martin D, Armingol JM. Traffic scene awareness for intelligent vehicles using convnets and stereo vision. Robot Auton Syst 2019;112:109–22.
[16] Horng G-J, Liu M-X, Chen C-C. The smart image recognition mechanism for crop harvesting system in intelligent agriculture. IEEE Sens J 2019.
[17] Jalil B, Leone GR, Martinelli M, Moroni D, Pascali MA, Berton A. Fault detection in power equipment via an unmanned aerial system using multi modal data. Sensors 2019;19(13):3014.
[18] Lei X, Sui Z. Intelligent fault detection of high voltage line based on the faster r-cnn. Measurement 2019;138:379–85.
[19] Girshick R. Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision; 2015. p. 1440–8.
[20] He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2961–9.
[21] Dai J, Li Y, He K, Sun J. R-fcn: Object detection via region-based fully convolutional networks. In: Advances in neural information processing systems; 2016. p. 379–87.
[22] Sermanet P, Eigen D, Zhang X, Mathieu M, Fergus R, LeCun Y. Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229, 2013.
[23] Law H, Deng J. Cornernet: Detecting objects as paired keypoints. In: Proceedings of the European Conference on Computer Vision (ECCV); 2018. p. 734–50.
[24] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems, 2015, p. 91–99.
[25] Redmon J, Farhadi A. Yolo9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 7263–71.
[26] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. Ssd: Single shot multibox detector. In: European conference on computer vision. Springer; 2016. p. 21–37.

[27] Nitin R, Mangesh KN. Landmark based shortest path detection by using a* and haversine formula. Int J Innovat Res Comput Commun Eng 2013;1(2):298–302.
[28] Garnett R, Stewart R. Comparison of gps units and mobile apple gps capabilities in an urban landscape. Cartography Geographic Informat Sci 2014;42(1):1–8.
[29] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 2818–26.
[30] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
[31] Lin T-Y, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 2980–8.
[32] Lin T-Y, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 2117–25.
[33] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.
[34] Loshchilov I, Hutter F. Sgdr: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983, 2016.
[35] Neubeck A, Van Gool L. Efficient non-maximum suppression. In: 18th International Conference on Pattern Recognition (ICPR'06), vol. 3. IEEE; 2006. p. 850–5.
[36] Nettleton DF, Orriols-Puig A, Fornells A. A study of the effect of different types of noise on the precision of supervised learning techniques. Artif Intell Rev 2010;33 (4):275–306.
[37] Zhu X, Wu X. Class noise vs. attribute noise: A quantitative study. Artif Intell Rev 2004;22(3):177–210.

Appendix 6

VI

H. Manninen, J. Kilter, and M. Landsberg, “A holistic risk-based maintenance methodology for transmission overhead lines using tower specific health indices and value of loss load,” *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107767, 2022

Electrical Power and Energy Systems 137 (2022) 107767

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

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

A holistic risk-based maintenance methodology for transmission overhead lines using tower specific health indices and value of loss load

Henri Manninen [a,*], Jako Kilter [a], Mart Landsberg [b]

[a] *Tallinn University of Technology, Department of Electrical Power Engineering and Mechatronics, Estonia*
[b] *Elering AS (Estonian TSO), Grid Maintenance Department, Estonia*

ARTICLE INFO

Keywords:
Condition assessment
Health index
Probability of failure
Risk assessment
Survival analysis
Value of lost load

ABSTRACT

European transmission system operators are facing challenging times in the next decade as majority of their transmission overhead lines are reaching the end of their projected lifetime. Traditional maintenance approaches would generate a significant wave of replacements that could be dispersed or postponed with more advanced decision-making methodologies. This paper presents a holistic risk-based maintenance decision-making methodology for transmission overhead lines and its practical implementation. The framework is refined with anomaly detection and health index prediction models that use machine learning algorithms to improve the input data quality. Asset health indices are used to determine the actual technical condition of each transmission overhead line tower separately and to calculate the probability of failure for each asset using survival analysis. The proposed methodology takes into the account transmission grid specific features where usually a failure in a meshed networks will not cause an electricity outage for customers and therefore a novel value of lost load approach is proposed. This paper also presents a case study based on Estonian transmission system where the proposed methodology enables to minimize risks in more cost-effective manner compared to traditional approaches and highlights the most critical elements in the grid.

1. Introduction

Transmission system operators (TSOs) around the globe are facing a significant wave of asset replacements due to the old age of their electricity systems. According to the ENTSO-E [1] it is expected European TSOs have to invest around 53 billion euros to maintain the current level of security of supply, where nearly 80% of the expenses will be for the refurbishment of overhead lines (OHL). There are currently no common approaches to assess the time-frame and urgency of the replacements for overhead lines but various methodologies are used among TSOs. According to [2,3] the most widespread approaches are corrective maintenance (CM), time-based maintenance (TBM), condition-based maintenance (CBM), and risk-based maintenance (RBM). Based on [2–10] the RBM methodology is proposed as the most cost-effective but also as the most sophisticated to implement. In TBM assets are usually replaced once they reach their expected lifetime. It is the most widespread approach among TSOs as it is easy to implement, but may lead to over investments. Decisions using RBM are usually made on the basis of risk matrices [11], where probability of failure (PoF) and consequences of failure (CoF) are combined. That approach is widespread as it enables simplified visualization of PoF and CoF, but unfortunately it lacks transparency on replacement decisions. As stated in [12], many companies use risk matrices without ensuring their efficiency on improved decision-making. For the useful decision-making, it is essential to determine input parameters, PoF and Cof, as accurately and transparently as possible.

PoF of OHLs can be calculated directly using statistical approaches such as Weibull and bathtub [13] curves or through a health index (HI) as done in [14–18]. As in TSOs grid there is usually low number of failures and they may happen randomly then determining PoF using parametric or semi-parametric functions may lead to over- or under-estimations in terms of insufficient data. The use of non-parametric survival analysis [19] is proposed to calculate PoF on the basis of HI data and historical failures. There are various ways to represent HI, but the most common ones divide HI linearly from 0 to 100 [15,17] or 0 to 5 [4,20], to describe the technical condition of complex assets. CIGRE has conducted a study [21] where it is seen that the majority of failures for high-voltage assets are random and [22] illustrates that estimated remaining life of OHL conductors varies significantly when comparing approaches based purely on age against approaches with additional measurements. To improve results, HI determination methodologies should rely on parameters that are reflecting the actual technical condition or deterioration of assets instead of age dependency. It must

* Correspondence to: Ehitajate tee 5, 19086 Tallinn, Estonia.
E-mail addresses: henri.manninen@taltech.ee (H. Manninen), jako.kilter@taltech.ee (J. Kilter), mart.landsberg@elering.ee (M. Landsberg).

https://doi.org/10.1016/j.ijepes.2021.107767
Received 23 May 2021; Received in revised form 6 September 2021; Accepted 28 October 2021
Available online 4 December 2021
0142-0615/

be noted that this paper is not focusing on the development of HI assessment methodology, rather it uses HI as the input in PoF determination from author's previous works done in [20,23], where HI is determined for each OHL tower and its component separately. That enables to determine PoF on the basis of its actual technical condition instead of age.

Value of loss load (VOLL) [24] is widely used to express CoF in monetary terms. It is difficult to determine VOLL in transmission grids as networks are planned using the n-1 requirement, where usually a failure of a single OHL will not cause directly any outage for customers. But in some cases a failure of a single OHL tower may cause an outage for customers, especially when multiple circuits are sharing the same towers or some OHLs are disconnected from the grid due to the maintenance works. To overcome this issue, it is proposed to consider each tower in the grid as an individual element instead of aggregating all towers and circuits on OHL level. OHLs can span up to hundreds of kilometers and its towers can be built on various terrain types such as fields or mountains. To improve CoF results, a novel OHL tower specific VOLL determination methodology is proposed using tower specific expected outage duration that takes into account geographical location and complexity of assets. As the proposed methodology relies on the input data then unsupervised machine learning is used for anomaly detection to improve input data quality and supervised machine learning is used to predict HI for assets without HI data to involve them in the risk assessment process.

In order to overcome of the main limitations of existing RBM approaches, a holistic methodology, based on HI and VOLL, is proposed. It enables transparent asset-management decision-making and presents a clear overview of all individual elements in the grid and how it affects reliability in general. The proposed methodology focuses thoroughly on the determination of RBM input parameters and presents an approach where each element in the grid will acquire a HI and corresponding VOLL values.

This paper is divided into seven main sections. Section 3 describes the overview of the proposed methodology. The framework of PoF determination is in Section 4 and HI determination of OHLs in Section 5. Section 6 explains CoF assessment methodology using VOLL with estimated outage duration. Section 7 presents a case study done in Estonian transmission grid. Conclusions and future work are given in Section 8.

2. Life-cycle management of transmission overhead lines

Life-cycle management (LCM) of transmission OHLs is the process of optimizing the maintenance, investment and condition assessment costs throughout the life-cycle. Transmission OHLs are complex assets to maintain as they cover large distances, consist of large amount of individual towers, thousands of kilometers of conductors and have a lifespan more than 60 years. In addition to maintaining the technical condition of assets it is also important to ensure the safety of OHLs throughout its life-cycle. Fig. 1 gives a brief overview of OHL LCM works that can be divided into three main component class that are right of way (ROW), conductor and grounding wires, and towers.

It is possible to cover almost all ROW LCM activities by using light detection and ranging (LiDAR) technology where data is collected using aerial vehicles and collected point-clouds are processed to acquire precise distance from one object to another. Condition assessment of ROW consists mainly of vegetation analysis and detection of structures in the ROW. The ROW by itself does not affect the technical condition of OHL but it must be maintained to eliminate outage because of vegetation flash-overs. Preventing vegetation-related outages improves also the safety of the OHLs as flash-overs may cause fires or even dangerous step-voltage for human and animals. In addition to vegetation analysis, LiDAR data is also used to determine the minimal ground clearances of each span or crossing with infrastructures by modeling conductor thermal behavior for allowed conductor temperature ranges

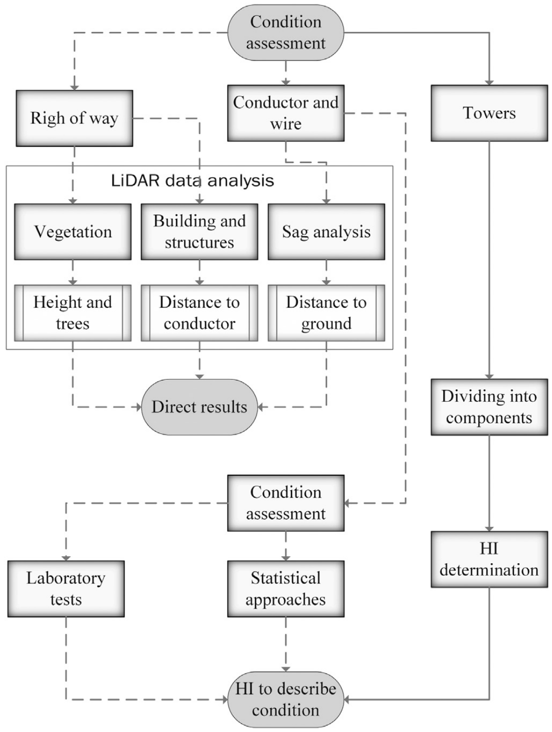


Fig. 1. Principle scheme of OHL condition assessment where dotted lines presents process parts that are also part of OHL condition assessment process but left out of the scope of this paper.

as explained in [25]. That enables to ensure the safety of all agricultural machines, structures in the ROW and crossings with roads by measuring distance from the conductor to the specific point to ensure the required safety margin. Using LiDAR technology has became widespread among TSOs due to the high accuracy and relatively low cost per inspection kilometer. Using LiDAR technology also enables to precisely measure distance from conductors or any point from ROW to structures, ground and crossings. As those results are from distance measurements and not affected by the technical condition of OHLs then they are left out of the scope of this paper.

OHL components that are usually covered with HI in LCM are conductors, grounding wires, insulators and towers with its subcomponents such as foundations, supports and cross-bars. HI is used for those components as they have long lifespan, they are prone to aging effects due to the material's deterioration and there are no direct measurements to assess technical condition of them accurately. The HI of conductors and grounding wires is usually determined through statistical approaches or laboratory tests such as tension and torque tests as there are almost no visual indicators of conductor and grounding wire aging as done in [22,26].

3. Methodology

A general overview of the developed methodology focusing on transmission OHL towers and is presented in Fig. 2. The process is divided into two parallel branches where the first starts with HI determination of OHLs and another one with VOLL determination. PoF of each tower is calculated on the basis of HI data and historical failures as explained in Section 4, where HI determination process is more detailed

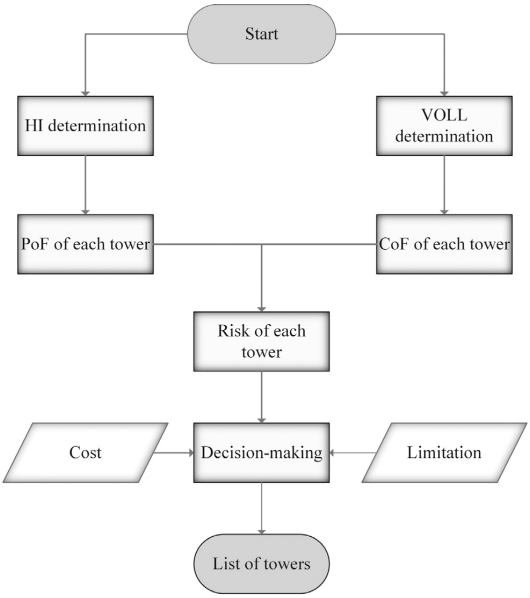


Fig. 2. Principle scheme of an asset management decision-making methodology where the rectangles represent process steps, parallelograms data and rhombuses decision points.

in Section 5. CoF of each tower is calculated on the basis of VOLL and cost to eliminate defects in Section 6. The full model for the RBM of OHLs including all sub-models is composed in Python.

3.1. Risk-based maintenance

In order to reach optimal decisions in RBM, it is essential to calculate the risk on a common and precise basis. The simplest principle for decision-making to determine whether to replace the asset or not is by using (1), where asset should be replaced once the risk is greater than the cost of repair works. Implementing that approach requires, a well-explained and transparent risk assessment methodology for reliable results.

$$Risk > Cost \tag{1}$$

This methodology relies heavily on the input data, where all errors in the first steps drastically affect final results. Also, compared to TBM or CBM approaches, the RBM approach provides an additional parameter that enables decisions to be made based on the cost of an asset and risk associated to that asset if the failure occurs. It also enables to determine the most critical assets that influence the reliability of grid the most. According to [3], risk in a transmission system can be stated in its simplest form by using (2).

$$Risk = PoF \times CoF \tag{2}$$

where PoF is the probability of failure and CoF the consequences of the failure.

3.2. Risk determination of individual towers

Risk of each tower is calculated based on the (2), but combining together PoF and CoF that are based on HI and VOLL is more sophisticated task. The framework of a risk assessment of each tower in the transmission grid is presented in Fig. 3. Risk assessment starts with grid calculations, where all possible combinations that will cause an outage or a limitation to customers in the grid will be saved. After that, all calculations are reviewed individually and it is checked if the outage was caused by only a single OHL or if there is a combination of more than one OHL. If there is a single OHL that is causing the outage, then all towers of this OHL will acquire the VOLL of the substation where the outage occurs. If there is more than one OHL, then it is checked if there are towers that share circuits of both OHLs. When this happens then all towers that satisfy the criteria are saved with substation VOLL. PoF of each tower is determined in Section 4 and PoF including VOLL determination in Section 6.

As a single OHL can cause an outage in more than on a one combinations, all possible combinations are calculated and saved. The maximum risk across all combinations for each tower is selected for further sections to find the worst-case scenarios for each individual tower. If there is more than one tower involved in the combination with circuits on separate towers, then the PoF of each tower will be calculated by using the joint probabilities for selected tower and the maximum PoF value of OHLs in the combination according to (3). In this paper up to two individual towers are used simultaneously to find possible outage combinations. N-1 and N-2 situations are calculated where up to two simultaneous faults in the grid might occur. The network was modeled until N-2 situation as N-3 and further scenarios increase the computational complexity in a meshed network significantly while increasing the risk assessment only marginally. Increasing the number of towers in an outage combination will increase the computational complexity significantly but improves results marginally as probabilities are unpretentious values.

$$PoF_{JP} = PoF_{Selected} \times \max_{i=1}^{n}(PoF_{HI_i}) \tag{3}$$

where PoF_{JP} is the joint probability of the combination, $PoF_{Selected}$ is the PoF the selected tower according to its HI, n is the number of towers in the combination and PoF_{HI_i} is the PoF of each tower in the combination according to HI.

3.3. Decision-making under budget restrictions

In an ideal world, there is always a sufficient budget to replace all the assets that are required, but that is not always possible. To maximize the efficiency of the proposed RBM decision-making methodology under budget limitations, optimization task is described as a knapsack problem [27] that is mathematically expressed as (4).

$$maximize \sum_{i=1}^{n} Cost_i * Risk_i <= Budget \tag{4}$$

where $Cost_i$ is the cost of replacement and $Risk_i$ risk of ith element.

Solving a small knapsack is relatively simple as after calculating all possible combinations the best one can be chosen. Solving a large knapsack problem becomes extremely computationally expensive as the number of combinations to find the best solution grows exponentially. To overcome this issue, different approaches are used to solve complex knapsack problems using more effective methods than simple brute-forcing, where all possible combinations are tested. In this paper, dynamic programming from [28] is used to solve the knapsack problem.

4. Probability of failure

Probability of failure determination for each asset in the grid is based on the asset HI information and combined failure data of that asset type. Fig. 4 presents an overview of PoF determination, where it is seen that HI data, critical replacements, and historical failures are input for the PoF determination. Based on the data, cumulative hazard functions using survival analysis [19] are calculated for each asset category based on the HI of assets. Output of the PoF model determines a corresponding PoF value for each individual tower according to its categories historical performance and condition.

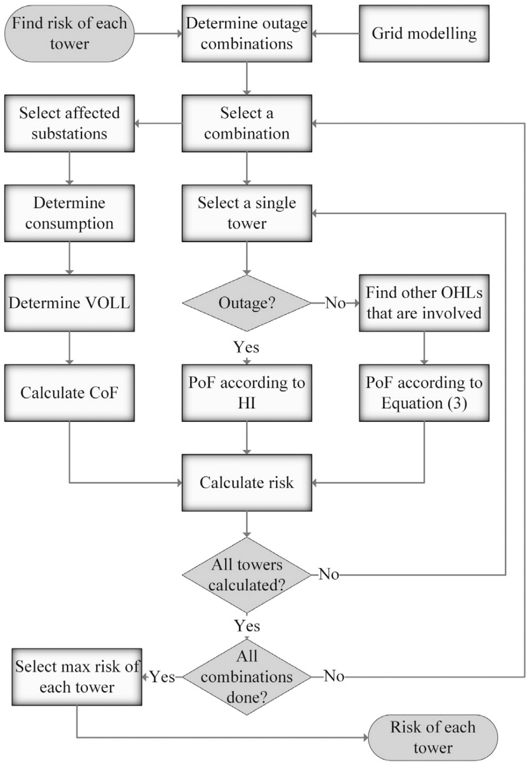


Fig. 3. Flowchart to find maximum risk of each tower in the grid. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes the start and end of the process.

4.1. Historical failures

Historical failure data is an important input parameter for PoF calculations as it reflects the historical behavior of assets. As this paper focuses on TSO level, it is important to take into account not only failures, but also critical defects and just-before-failure replacement of assets. It is important to fix the last recorded asset HI before failure or replacement by using standardized failure reports such as CIGRE recommendation for OHLs [4] or IEC 62271-1:2017 [29] for substation switchgears. The usage of standardized failure reports enables the comparison of different failures on a common basis, and it also enables asset technical condition to be registered. As HI presents the technical condition of the asset, it also enables asset-specific PoF for each HI class to be found.

4.2. Probability of failure using survival analysis

Survival analysis and its application are well described in [19]. In this paper, non-parametric survival analysis is used to calculate the PoF for each tower since no distributional assumptions are known. Different voltage levels are differentiated to acquire more reliable results due to different reliability requirements for 110 kV and 330 kV voltage levels in Estonia. PoF is calculated for each tower individually as is done in [30], where the parametric exponential model based on average, perfect, and end-of-life failure rates was used. Weibull, exponential, gamma or the Cox proportional hazard models can be more detailed compared to non-parametric models once there is a sufficient amount of data involved. This paper focuses on TSO level, where failures in the grid are rare due to preventive maintenance approaches and therefore, there is lack of data to use parametric models, which can be considered as alternatives for the survival analysis in the future. Cumulative hazard functions are composed for each selected asset category using the Nelson-Aalen [31] (NA) estimator. The NA estimator is used to directly estimate the cumulative hazard function given by (5).

$$H_{NA}(t) = \sum_{t_i \le t} \frac{d_i}{n_i} = \sum h_{NA,i}(t_i) \tag{5}$$

where n_i corresponds to the number of towers present at time t_i and d_i is the number of failures at time t_i.

It is also possible to use Kaplan–Meier estimator, but according to [31] NA is slightly superior in terms of increasing failure rates such as aging assets in transmission grid. For better results censoring should be used when assets are taken out of service before failure occurs. Using the NA estimator to compose a cumulative hazard function enables visually to examine distributional model assumptions for reliability data of assets and have a similar interpretation as probability plots. Fig. 10 in Section 7 presents an example of cumulative hazard functions for 110 kV and 330 kV OHL towers.

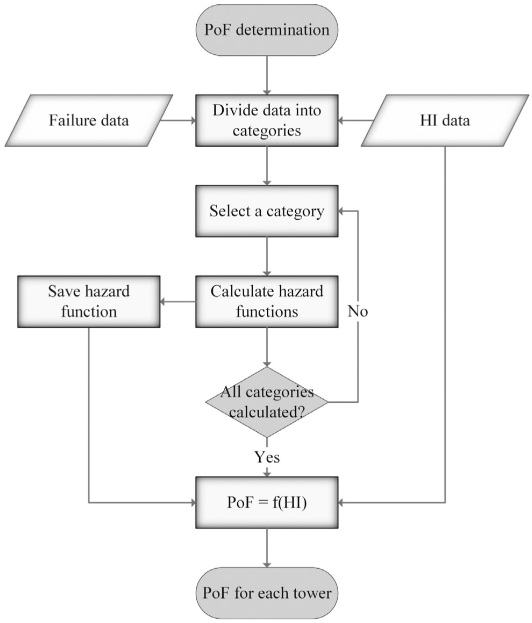


Fig. 4. Flowchart to find PoF of each tower. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes start and end of process.

5. Health index determination

Fig. 5 presents a complete HI determination process that is combined from authors' previous works to overcome the main issues regarding the input data collection for RBM implementation. It must be noted, this section illustrates a HI determination approach for transmission OHLs to determine HI, but various other methodologies that support tower based approach from literature could also be used to implement the general methodology presented in Fig. 2. Two different approaches are used to determine HI of transmission OHLs: visual inspections according to [20] and automatic condition assessment of OHLs using deep learning techniques according to [23]. To increase the data quality, an

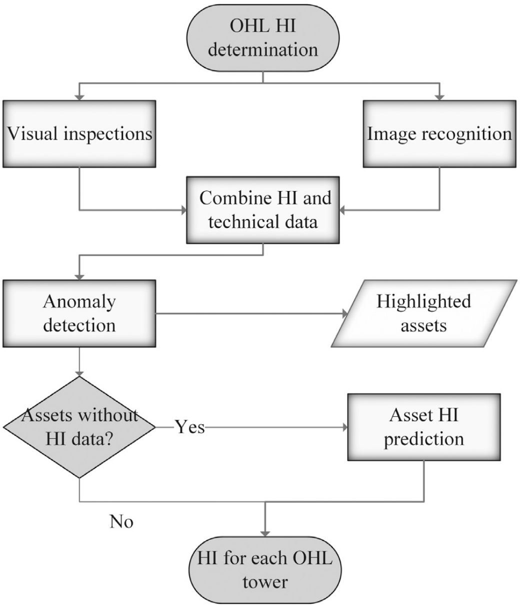


Fig. 5. Flowchart of health index determination process. Rectangles represent process steps, parallelograms data, rhombuses decision points and round shapes start and end of process.

Table 1
Asset technical features used in this paper for anomaly detection and health index prediction.

Feature	Variations
Number of circuits	1 to 4
Tower type	Suspension or tension
Voltage level	110 kV and 330 kV
Tower material	4 options
Manufacturer	6 manufacturers
Tower configuration	214 configurations
OHL direction changes	Angle of deviation
Existence of a perch guard	Yes or no
Age	1 to 67 years

anomaly detection and HI prediction models are integrated in the HI determination process that is described in Section 5.1 and HI prediction model is described in Section 5.2. Anomaly detection model is based on unsupervised machine learning and prediction model is based on supervised machine learning algorithms. Both models use HI and asset technical data to improve input data quality for PoF determination as described in Table 1.

HI determination approaches used in the paper are based on predefined visual indicators that are composed according to material's physical fatigue and visual indicators that describe different life-stages of OHL components. Table 2 presents predefined visual indicators with corresponding HI classes for reinforced concrete poles that are used in [23] to determine HI automatically from images using deep learning. The data collection methodology is different for both branches, but they both use the same backbone where there is almost 150 condition indicators for condition assessment of OHL components according to [20]. Technical condition from excellent to poor is divided linearly between six classes. HI class with a perfect technical condition is HI0 and assets that have end-of-life criteria detected have HI class HI5.

Table 2
List of visual indicators for reinforced concrete poles and responding HI values [23].

Visual indicator	Health index
Hole	5
Loss of cross section > 20%	5
Concrete is falling off	5
Loss of cross section < 20%	4
Crack	4
Visible reinforcements	3
Micro longitudinal cracks	2
Hair-like cracks	1
Other minor defects	1
No visible defects	0

5.1. *Increasing input data quality using anomaly detection*

There are always impurities in the raw data that may lead to inaccurate results even with the most sophisticated models. As this paper presents a methodology, where decisions are based on the input data, it is essential to improve data quality. One option is to detect and fix all anomalies manually, but as soon as the amount of data increases, it will become overwhelming. To improve input data quality, unsupervised machine learning algorithms are used to detect incorrect values in the data automatically. Outlier detection or anomaly detection algorithms described in [32] enable to minimize the risk of incorrect values in further steps by outlining suspicious values in the data for the double check. For outline detection, information about all assets that have technical and HI data will be used to train the model as seen in Table 1. The outline detection model is implemented right after HI data import to minimize errors. The anomaly detection model is composed and integrated into the full framework using scikit-learn [33] toolbox in Python. As the anomaly detection model detects all assets that have the largest deviations in technical features compared to the HI, this does not always mean that the data is incorrect. For example, in some cases there might be relatively new assets with high HI values due to mechanical defects not caused by aging, but those will be highlighted as anomalies in the data by the anomaly detection model. The output of the anomaly detection model is a list of assets that might have data quality issues and the output of HI prediction model is a list of assets with corresponding HI values. As there are possibilities where the model detects correct values as anomalies, this approach requires the results to be validated by experts. That minimizes the data validation workload of experts by only double-checking a short list of highlighted assets, not the full data set.

One of the most accurate and widely used anomaly detection algorithms is Isolation Forest [34], which explicitly isolates anomalies in the data instead of profiling normal points. Isolation Forest is an anomaly detection algorithm that is built on an ensemble of decision trees for a given data set. It is ideal for high-volume data sets due to the low memory requirement and it works well even when there are no anomalies present in the training set. Fig. 6 presents a simple example of anomaly detection using Isolation Forest where straight black lines describe random partitions generated by the model. Detailed description of the Isolation Forest algorithm and its properties is in [34]. Similarly to other anomaly detection models, Isolation Forest uses anomaly scores for decisions. Decisions are done based on following rules:

- Anomaly score close to 1 indicates anomaly
- Anomaly score close to 0 indicates normal data point
- Anomaly scores for all data points are close to 0.5 indicates there are no anomalies in the data

The selection of hyper parameters for Isolation forest:

- Number of estimators - 100
- Contamination - 0.01
- Bootstrap - enabled

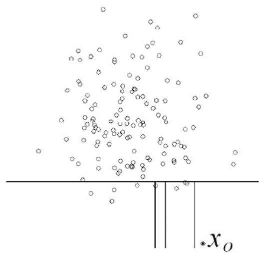


Fig. 6. An example of the anomaly detection (X_0) using Isolation Forest [34].

- Number of features - 9 as in stated in Table 1

The contamination parameter is affecting the results of using anomaly detection model the most. It is describing the expected proportion of outliers in the data and by selecting a certain value, the model will output a selected amount of data that is closest to anomalies. For example, in the implemented model contamination is set 0.01 and therefore, the model output 1% of towers that are the most suspicious based on their features. That approach still requires an expert to double-check the results, but it eliminates the danger of automatically marking correct data as anomalies. Unfortunately there are always possibilities where a really new tower has a significantly bad technical condition or vice-versa. By using the anomaly detection model with a fixed and relatively small contamination parameter, it is possible to minimize the workload of experts by validating only a small partition of the data. Other hyper parameters used by the model are selected according to manual hyper parameter tuning where the model detected the most suspicious values in the data set such as 5 years old tower with HI 5 that was hit by an agricultural equipment.

5.2. Health index prediction

The HI prediction methodology was developed in [35] to predict HI of assets without HI data, where the most widespread supervised machine learning algorithms were analyzed. Random Forest [36] using unmodified data provided the most accurate results and it is also implemented in the framework of HI determination to predict HI of missing assets. A detailed analysis, performance metrics and hyper-parameters of tested models are presented in [35]. That approach enables maintenance decision-making to be implemented on the full grid and not only on assets that have HI data. It also decreases the inspection requirement and budget as it enables HI values to be gotten for towers with a relatively high accuracy without physically visiting them. Fig. 7 presents an example of HI prediction of OHL towers that did not have HI data associated with them. According to [35] it is possible to predict the HI of missing OHL towers with 70% to 80% accuracy in six different HI classes without additional measurements or site visits in an accurate and cost-effective manner. The implementation of the missing HI prediction is based on the technical parameters of the assets combined with the corresponding HI values acquired from periodic visual inspections. To predict the missing values of assets, it is important first to determine all assets in the data set that do not have corresponding HI values. After detecting assets with missing HI values, all assets that have HI values will be used to train the prediction model. The HI prediction for assets without HI will be done based on the technical features of those assets and the HI value is acquired.

6. Consequences of failure

Consequences of failure are based on VOLL and direct costs associated with the repair of the failure. All other widespread parameters

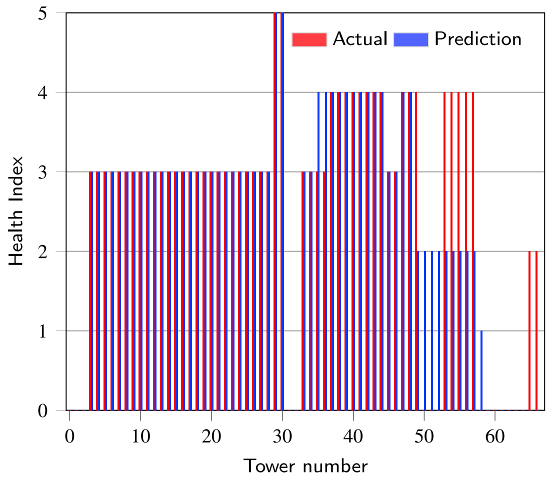


Fig. 7. An example of health index prediction model results of OHL without health index from [35].

to assess CoF such as safety, environmental issues or loss of reputation are not considered in this paper as they are not comparable between different companies. Total consequences of an outage can be calculated according to (6), but only direct cost and VOLL parameters are used in this paper and therefore $\sum CoF_i = 0$.

$$CoF = VOLL + CoF_{Direct} + \sum CoF_i \tag{6}$$

where CoF_{Direct} are costs associated directly with the outage such as cost of repair works, CoF_i is consequences related to safety, environment, and publicity for the company or even political pressure.

For further studies, it is possible to add additional parameters to assess all possible CoF, but as the main aim of this paper is to present a full methodology to combine HI and VOLL for decision-making, this item is not studied further here. The main reason why all other parameters for the failure consequences assessment are ignored is that they are based on non-comparable values and each utility has its own risk mitigation strategies. For example, if some TSOs include loss of human life in the consequences with great value and the other TSO does not, then the results are significantly different. It is possible to include those consequences in this methodology as well, but they have to be first monetized and must be calculated on a standardized basis.

6.1. Value of lost load

VOLL is used as a monetary indicator that expresses the cost associated with an interruption of electricity supply. VOLL is determined through multi-step approaches that usually start with dividing consumers into predefined categories and assessing cost of energy not supplied (CENS) for each customer separately. A comprehensive comparison of VOLL values and different methodologies in various countries is presented in [24]. It is seen that there is a large gap between macroeconomic and willingness-to-pay approaches. In this paper, the VOLL of a single substation is determined based on the consumer profiles in the Estonian transmission system where VOLL values for each main sector are determined in [37]. Using the substation's historical load for the previous years, it is possible to calculate weighted average VOLL for a substation when there is sufficient data for each consumer connected with that substation.

VOLL can be calculated using (7) where it reflects the total cost of electricity outage for a single substation based on the price of the consumer-specific energy units, consumption, and expected duration

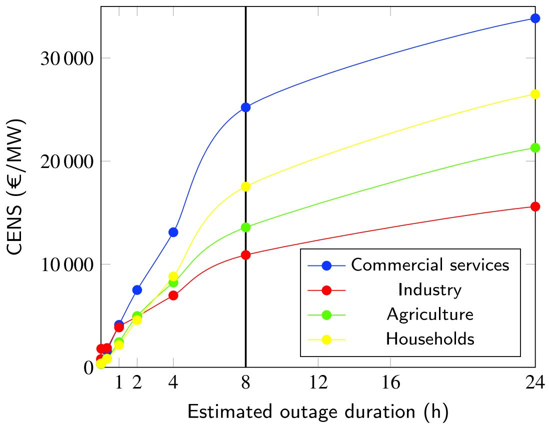


Fig. 8. CENS for different estimated outage durations. Blue line represents commercial services, red industry, green agriculture, yellow households and black vertical line outage with duration of 8 h. [37]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the possible outage. It should be noted that it is impossible to predict failure occurrences with an exact time frame and therefore average consumption is used in this paper. For the worst-case scenario, maximum consumption of the substation could be used as it is reflecting the maximum possible VOLL in the substation, but it may result in over-investments.

$$VOLL_{substation} = \frac{\sum CENS(t)_i \times load_i}{\sum load_i} \times t \tag{7}$$

where $CENS(t)_i$ is cost of energy not supplied for the sector i at expected outage duration t.

CENS is time dependent and can be expressed as functions in Fig. 8. Those functions are composed according to surveys conducted in Estonia [38] and converted to today's value in [37]. As seen in the figure then it is essential to define the estimated outage duration as accurately as possible as it affects CENS significantly. The black vertical line represents an estimated outage duration of 8 h. For further VOLL calculations CENS values should be used as function of estimated outage duration.

6.2. Estimated outage duration

Duration of the outage is usually determined by the type of the failure, the complexity and time of repair works of assets. The distance from the nearest road is crucial parameter for the outage duration estimation as there are towers that are next to easily accessible infrastructure and in the middle of nowhere. For the simplicity, this paper assumes the worst-case-scenario, where the most critical component of the tower fails. Voltage level and material are used to distinguish assets by their complexity as for example steel towers require additional foundations and reinforced concrete poles does not. Usually higher voltage levels refer to more complex assets as higher voltage level require more safety margins and therefore are larger. Values in Table 3 present estimated repair times of OHL tower failures based on the best practice from the Estonian TSO. Estimated outage duration values are heavily influenced by TSOs that defines those times due to asset types, country-specific geographical features and maintenance policies. The estimated outage duration is also significantly affected by organizational aspects such as the availability of manpower and weather, but for the simplification purposes the most pessimistic approaches are used. It is also possible to expand Table 3 with more options and decision points to increase the accuracy of the estimated outage duration while taking into the account more parameters. Table 3 presents simplified estimated outage duration times that characterize average values according to tower type, voltage level, and distance from the nearest road. Those values are determined empirically based on the best knowledge about historical outages in the grid.

Table 3
Estimated outage duration for OHLs according to tower type, voltage level and distance from roads.

Distance (m)	Estimated outage duration (h)			
	110 kV		330 kV	
	Steel	Concrete	Steel	Concrete
<100	12	8	16	12
100–1000	24	12	24	24
1001–10 000	36	24	48	36
>10 000	72	72	72	72

As OHLs usually cover significant distances and are in remote areas, the repair time to eliminate the outage is also significantly influenced by terrain type. Outage duration of OHLs is also affected by the geographical location as there are some areas where it is impossible to reach due to the large distance from the nearest roads or they are located on difficult terrain. Estimated outage duration can be calculated by using (8). It must be noted that $C_{Terrain}$ is empirical constant that should be selected according to countries geography. As it is significantly more challenging to replace a tower in the mountainous or swamp area compared to field then $C_{Terrain}$ can vary up to multiple times. But as Estonia is relatively flat and has high infrastructure coverage all around the country, the distance from the nearest road in combination with tower type is used as an indicator of possible outage duration and therefore $C_{Terrain} = 1$.

$$T_{outage} = C_{Terrain} \times T_{estimated} \tag{8}$$

where $C_{Terrain}$ is constant that takes into the account terrain type and $T_{estimated}$ the estimated outage duration according to Table 3.

To acquire a minimal distance from an OHL tower to the nearest road, the tower's geographical location is used with road information from OpenStreetMap [39]. For the simplicity, each individual distance from the nearest road to towers is not used but is divided into four proximity zones around the tower. Proximity zones can be generated based on country-specific features. In this paper, four zones are used and the radius of each is determined based on the best knowledge by the Estonian TSO. The first proximity zone has a radius of 100 m around the tower, the second has 1000 m, the third 10,000 m and the fourth 10,000 m or more.

The methodology is explained in Fig. 9, where are two OHL towers, two roads and proximity zones around the towers. It must be noted that only single towers from both OHLs are used and OHL spans with other towers are ignored to present a clear and simplified example. Proximity zones are marked with colors that represent the following: green — <100 m, yellow — 100 m to 1000 m, red — 1001 m to 10 000 m. It is seen that for the *Tower*1 only *Road*1 collides into that tower's proximity zones. To be more precise then *Road*1 collides with *Tower*1 green proximity zone that is less than 100 m. It is clearly seen that it is possible to select proximity zone <100 m for *Tower*1. In case of *Tower2*, it is seen that there are no roads in the green and yellow zones, but there are two roads in the red zone. That means it is possible to select 1001 m to 10 000 m proximity zone for that tower.

7. Case study

The case study presents an example of implementing the proposed methodology in the Estonian transmission grid for OHLs. For the case study, three approaches are compared based on the rules from the corresponding strategy. The first approach is time-based maintenance (TBM), where all assets will be replaced once the expected lifetime is

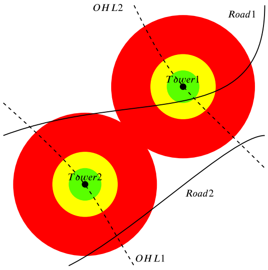


Fig. 9. Simplified example of the OHL tower's distance determination from the nearest roads. There are two OHLs (*OHL*1 and *OHL*2) marked with dashed lines where just individual towers (*Tower*1 and *Tower*2) with corresponding proximity zones (green, yellow and red) and two roads (*Road*1 and *Road*2) are used in this example. Green represents a proximity zone with a radius of 100 m, the yellow zone a radius of 1000 m and the red zone a radius of 10 000 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Input data for a case study.

Explanation	Count
Number of OHLs	200
Number of towers	16 071
Number of substation	139
Number of towers with HI data	15 971
Number of towers without HI data	100
Number of failures in grid	277

exceeded. In this case study, 50 and 60 years will be chosen for replacements as those values are commonly used by TSOs. The next approach is condition based maintenance (CBM), where assets will be replaced after reaching a certain condition from the condition assessment. In this paper, HI values are used for the replacement decisions, where, in the first case, all assets with HI5 will be replaced and in the second case all assets with HI4 and HI5 will be replaced. The third approach is using the proposed risk-based methodology (RBM) where all assets that have higher risk than the cost of replacement will be replaced.

7.1. Input data

Input data about towers technical condition in this case study is collected in 2018 from periodical visual inspections in Estonian transmission grid. Specially designed mobile tablet application were used based on the methodology described in [20] to acquire HI data about 16 000 individual OHL towers. Further more, inspections produced nearly two million individual defects about OHLs that were aggregated on the tower level. For the simplicity of the proposed methodology HI data was aggregated on tower level and a single HI value of each tower was generated on the basis of detected defects. For the simplification in the decision-making, a full tower replacement is chosen as a maintenance decision once a critical tower is detected from aggregated data. Asset technical features, grid topology, substation historical consumption and failures is from Estonian TSOs databases.

Input data for the case study is presented in Table 4. It is seen that there are 200 different OHLs connecting 139 substations. The total number of OHL towers with HI is around 16 000 and without HI is just a single OHL with 100 towers that were deleted from the HI data base to illustrate the HI prediction model. Historical data about critical replacements and failures is recorded for 277 towers with HI and other technical parameters.

The anomaly detection model described in Section 5 was used on the whole dataset, where 10 of the most abnormal values were highlighted. Those values included an 11-year-old tower with HI5 that has critical

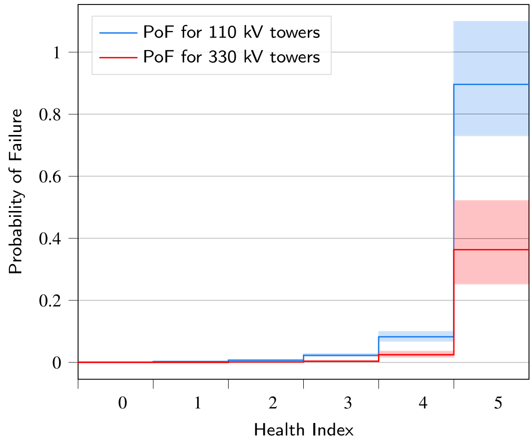


Fig. 10. Probability of failure of 110 kV and 330 kV towers based on the health index values. Light-colored areas around the lines represent 95% confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

defects. As all the data has been previously cleaned then all HI values remained unchanged.

For the towers without HI values, an HI prediction model was used. As the model used exactly the same input data as in [35] then the results were also identical and can be further studied in [35]. Using HI prediction model enabled missing HI values to be predicted for 100 missing towers in the grid and for further calculations all towers in the grid had a HI value to find PoF.

7.2. Probability of failure

Survival analysis using Nelson-Aalen [31] estimator is used to calculate asset PoF based on asset HI and historical failures and critical replacements in the grid. PoF is calculated for each tower and all towers are divided into two categories based on 110 kV and 330 kV voltage levels. For both voltage level a separate cumulative hazard curve is composed. In Fig. 10 it is clearly seen that the PoF of 330 kV towers is greater than for 110 kV towers on all HI classes. Also by increasing the HI of the asset, PoF increases, especially for HI5 where it reaches almost 0.89 for 110 kV and 0.38 for 330 kV towers. PoFs for all other HI classes are much lower compared to HI5. That reflects realistic situation where assets with critical technical condition are prone to failures and assets with good condition are extremely reliable. It is interesting to see that PoF for HI0 to HI3 increased only marginally while HI4 and especially HI5 have much higher PoF compared to other HI classes. That can be explained with the exponential nature of degradation processes where mechanical strength decreases rapidly at the end of their lifetime.

7.3. Consequences of failure

7.3.1. Estimated outage duration for each tower

Estimated outage duration is calculated on the basis of the distance from the towers to the nearest roads. Proximity zones for each tower in the Estonian TSOs grid are calculated using tower geographical coordinates and comparing them against the OpenStreetMaps road network. In addition to distance, different tower types are used to determine realistic outage durations. As Estonia is relatively flat with accessible ROW then $C_{Terrain} = 1$ in (8). The results of estimated outage duration in this case study are presented in Fig. 11. It is seen that the estimated outage duration for a majority of towers is either 24 h (8947) or 12 h (4829). There are only 187 towers for which it is expected to have repair time longer than 24 h.

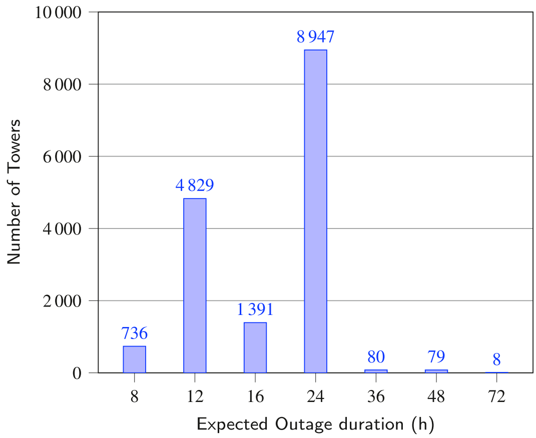


Fig. 11. Summarized outage duration of each tower based on the rules from Table 3.

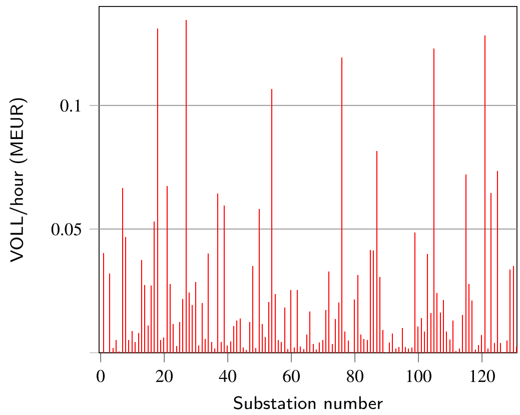


Fig. 12. Calculated VOLL/hour values of all substations for 8 h estimated outage duration.

Table 5
Customer's structure and VOLL values for the selected substation according to Fig. 8 for estimated 8 h outage.

Customer sector	Distribution (%)	VOLL (EUR/MWh)
Industry	0.36	10 890
Commercial services	82.77	25 210
Agriculture	0	13 570
Households	16.87	17 520
Substation	**100**	**23 861**

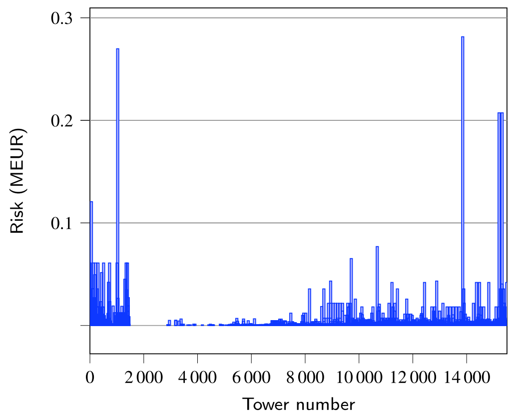


Fig. 13. Calculated risk values of each tower in the grid.

7.3.2. Value of lost load

VOLL values for each sector in Estonia used in this paper are from [37]. An example of a single substation's VOLL values for each connected customer sector and substation in general are calculated using (7) and are presented in Table 5 for a 8 h outage according to Fig. 8. Similar tables are calculated for each substation and estimated outage duration from Table 3 resulting in nearly 1000 possible VOLL values.

Fig. 12 presents all substations with calculated VOLL values. VOLL/hour is calculated according to (7) for all substations with the outage duration of eight hours as done in Table 5 based on five year consumption data for each sector. Results are showing large variances between different substations where the maximum VOLL/hour is nearly 140 000 EUR and minimum ones just around 300 EUR. This is because there are some substations with large consumption and some substations have nearly no consumption at all. Those VOLL values for substations are further used in combination with expected outage duration to calculate CoF.

7.4. Results

Risk determination results of all individual towers are presented in Fig. 13. It is seen that some towers have significantly higher risk values than others, especially towers 1032, 13 854, 15 207 and 15 319 where risk values exceed 0.2 MEUR compared to towers with risk of only few euros. That illustrates clearly that the proposed methodology enables to determine the most critical towers in the grid for the effective maintenance decision-making.

A comparison of different asset management strategies is presented in Fig. 14. It is seen that TBM scenarios produced the highest cost of replacement, replacing, in particular, all assets after 50 years of service with the cost of 161 MEUR. To replace all assets that are older than 60 years requires around 41 MEUR. For the CBM scenarios it is seen that replacement costs are less, but to replace all assets that have higher HI than 4 requires around 25 MEUR compared to CBM5 with 1.9 MEUR where only assets with HI5 will be replaced. The proposed risk-based approach requires around 1.7 MEUR to replace all critical assets. When looking into the cost of the remaining risk, TBM60 has around the same risk level as when doing nothing. CBM5 has remaining risk around 7.8 MEUR and TB50 6.4 MEUR. CBM4 and RBM produce similar results in terms of remaining risk, respectively 4.9 MEUR and 4.8 MEUR, where RBM has significantly lower cost compared to CBM4. It is clearly seen that the proposed risk-based approach produces the lowest cost of replacement with the lowest cost of remaining risks compared to other strategies.

8. Conclusions and further work

This paper presents a holistic RBM methodology for transmission OHLs that enables to overcome the main issues related to classical RBM implementations such as the transparency related to input parameters and decision-making. The main drawbacks of classical RBM methodologies are eliminated by using tower specific PoF determination based on the actual technical condition of assets and CoF based on precise VOLL determination. The proposed methodology increases substantially decision-making efficiency compared to TBM and CBM approaches and clearly determines the most critical towers in the

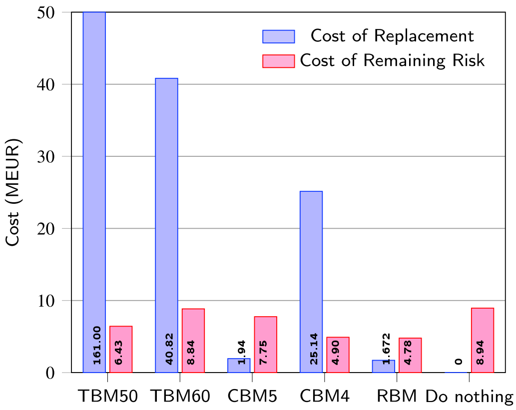


Fig. 14. Comparison of different asset management strategies. TBM50 and TBM60 are time-based scenarios where assets will be replaced in 50 and 60 years, and CBM5 and CBM4 are condition-based scenarios where assets with corresponding HI will be replaced and RBM is the proposed risk-based approach.

grid. That enables well-argumented and transparent asset-management decision-making as each element in the grid can be observed separately. The case study presents that the proposed methodology determined the most critical towers in the grid while outperforming all other approaches in terms of decision-making efficiency. The total cost of replacements was up to 100 times lower using the proposed methodology compared to TBM and up to 14 times lower compared to CBM approach where only towers with the end-of-life criteria were replaced. That indicates the efficiency of the proposed methodology as it enables to reduce substantially risks in the grid by detecting the most critical towers.

In further work, it is recommended to scale up the methodology by including conductors, grounding wires and insulators into the framework for effective LCM of OHLs in a single decision-making model. Non-monetized parameters can be added into the CoF determination process, but it requires common grounds for safety, environment, etc. risk assessment. It is also recommended to increase the quality of asset HI and failure data by collecting them for several years in a standardized manner. Decision-making process can be further improved by using component specific HI and defects in combination of optimization algorithms such as linear programming [9] or knapsack optimization [27] to find the most cost effective solutions based on the data. It must be noted that using mathematical optimization will lead to better results but majority of the preciseness lies on the assessment of model input parameters.

CRediT authorship contribution statement

Henri Manninen: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Formal analysis, Investigation, Software, Visualization. **Jako Kilter:** Supervision, Validation, Methodology, Writing – original draft, Writing – review & editing, Formal analysis. **Mart Landsberg:** Supervision, Validation, 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.

References

[1] Andreini E, et al. Facing the Replacement Wave. Tech. Rep., ENTSO-E; 2015.
[2] Balzer G, Schorn C. Asset Management for Infrastructure Systems - Energy and Water. 2015.
[3] Ross R. Reliability Analysis for Asset Management of Electric Power Grids. Wiley - IEEE; 2019.
[4] CIGRE WG B213. TB175 Management of existing overhead transmission lines. Tech. Rep., CIGRE; 1999.
[5] CIGRE WG B306. TB576 It Strategies For asset Management Of Substations - General Principles. Tech. Rep., CIGRE; 2014.
[6] CIGRE WG B332. TB660 Saving Through Optimised Maintenance in Air Insulated Substations. Tech. Rep., CIGRE; 2016.
[7] Hughes D. Condition based risk management (CBRM) — enabling asset condition information to be central to corporate decision making. In: CIRED 2005 - 18th International Conference and Exhibition on Electricity Distribution. 2005, p. 1–5.
[8] Raghavan S, Chowdhury B. State diagram-based life cycle management plans for power plant components. IEEE Trans Smart Grid 2015;6(2):965–72.
[9] Mkandawire BO, Ijumba NM, Whitehead H. Asset management optimization through integrated systems thinking and N-1 contingency capability for refurbishment. IEEE Syst J 2011;5(3):321–31.
[10] Pokorádi L. Fuzzy logic-based risk assessment. In: Academic and Applied Research in Military Science, Vol. 1. 2002.
[11] Shakoor W. Qhse risk assessment matix. In: Integrated Management System (IMS) Project. 2018.
[12] Thomas P, Bratvold R, Bickel J. The risk of using risk matrices. SPE Econ Manage 2013;6.
[13] Stanley NF, Heap HF. Reliability-Centered Maintenance. Report Number AD-A066579. Tech. Rep., United States Department of Defense; 1978.
[14] Jahromi A, Piercy R, Cress S, Service J, Fan W. An approach to power transformer asset management using health index. IEEE Electr Insul Mag 2009;25(2):20–34.
[15] Dong M, Nassif AB. Combining modified Weibull distribution models for power system reliability forecast. IEEE Trans Power Syst 2019;34(2):1610–9.
[16] Blackmore P, Leeprechanon N. Improving network reliability through effective asset management. In: 2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA). 2015, p. 1–9.
[17] Ayu K, Yunusa-Kaltungo A. A holistic framework for supporting maintenance and asset management life cycle decisions for power systems. Energies 2020;13:1937.
[18] Hashim R, Usman F, Baharuddin I. Determining health index of transmission line asset using condition-based method. Resources 2019;8:80.
[19] Emmert-Streib F, Dehmer M. Introduction to survival analysis in practice. Mach Learn Knowl Extract 2019;1:1013–38.
[20] Manninen H, Kilter J, Landsberg M. Advanced condition monitoring method for high voltage overhead lines based on visual inspection. In: 2018 IEEE PESGM. 2018, p. 1–5.
[21] Janssen A, Makareinis D, Sölver C. International surveys on circuit-breaker reliability data for substation and system studies. IEEE Trans Power Deliv 2014;29(2):808–14.
[22] Tsimberg Y, Lotho K, Dimnik C, Wrathall N, Mogilevsky A. Determining transmission line conductor condition and remaining life. In: 2014 IEEE PES T&D Conference and Exposition. 2014, p. 1–5.
[23] Manninen H, Ramlal CJ, Singh A, Rocke S, Kilter J, Landsberg M. Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques. Int J Electr Power Energy Syst 2021;128:106726, [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0142061520342708.
[24] Kuckshinrichs W, Schröder T. Value of lost load: An efficient economic indicator for power supply security? A literature review. Front Energy Res 2015;3.
[25] CIGRE WG B242. TB601 Guide for Thermal Rating Calculations of Overhead Lines. Tech. Rep., CIGRE; 2014.
[26] Gusavac S, Nimrihter M, Geric L. Estimation of overhead line condition. Electr Power Syst Res 2008;78(4):566–83.
[27] Shaheen A, Sleit A. Comparing between different approaches to solve the 0/1 knapsack problem. Int J Netw Secur 2016;16:1–10.
[28] Trick M. A dynamic programming approach for consistency and propagation for knapsack constraints. Ann Oper Res 2001;118.
[29] IEC 62271-1:2017 High-voltage switchgear and controlgear - Part 1: Common specifications for alternating current switchgear and controlgear. Standard, International Electrotechnical Commission; 2017.
[30] Manninen H, Kilter J, Landsberg M. Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices. In: 2019 PQ and SEEM. 2019, p. 1–6.
[31] Colosimo E, Ferreira F, Oliveira M, Sousa C. Empirical comparisons between kaplan-meier and nelson-aalen survival function estimators. J Stat Comput Simul J Stat Comput Sim 2002;72:299–308.
[32] Hodge V. A survey of outlier detection methodologies. Artif Intell Rev 2004;22:85–126.
[33] Pedregosa F, et al. Scikit-learn: Machine learning in python. J Mach Learn Res 2011;12:2825–30.

[34] Liu FT, Ting KM, Zhou Z. Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining. 2008, p. 413–22.

[35] Manninen H, Kilter J, Landsberg M. Health index prediction of overhead transmission lines: A machine learning approach. IEEE Trans Power Deliv [Manuscr Sec Rev] 2020.

[36] Brieman L. Random forest. Mach Learn 2001;45(5):5–32.

[37] Sokk K. Methods of power outage damage assessment and their applicability in the Estonian electricity system Outage cost assessment methods and application in Estonian electricity system. (Master's thesis), Tallinn University of Technology; 2019.

[38] Raesaar P, Tiigimägi E, Valtin J. Assessment of electricity supply interruption costs in estonian power system. Oil Shale 2006;22:217–31.

[39] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org. 2017, https://www.openstreetmap.org.

Curriculum Vitae

1. Personal data

Name	Henri Manninen
Date of birth	14 July 1992
E-mail	henri.manninen@taltech.ee

2. Education

2016–...	Tallinn University of Technology, School of Engineering, Electrical Power Engineering and Mechatronics, PhD studies
2014–2016	Tallinn University of Technology, Faculty of Power Engineering, Electrical Power Engineering, MSc *cum laude*
2011–2014	Tallinn University of Technology, Faculty of Power Engineering, Electrical Power Engineering, BSc

3. Professional employment

2019– ...	Elering AS, Project Manager of Asset Management Digitalization
2016–2019	Elering AS, Operation and Maintenance Specialist

4. Voluntary work

2016–2021	Member of CIGRE WG B3.48 (TB858)

5. Honours and awards

- 2019, Scholarship for Full-time Students
- 2019, Kristjan Jaak scholarship to study periods abroad (Trinidad and Tobago)
- 2017, Scholarship for Full-time Students
- 2015, Elering's Energy Scholarship

6. Defended theses

- 2016, Methods for defining the operational lifetime of substations and overhead lines, MSc, supervisor Paul Taklaja, Tallinn University of Technology, Faculty of Power Engineering
- 2014, Overhead Transmission Lines Capacity and Mathematical Methods for its Determination, supervisor Jako Kilter, Tallinn University of Technology, Faculty of Power Engineering

7. Supervised dissertations

- Tauri Kalmet, Master's Degree, 2021, (sup) Jako Kilter; Henri Manninen, Eesti ülekandevõrgu alajaamade SF6 võimsuslülitite seisundi hindamise metodoloogia arendamine ja analüüs (Development and analysis of condition assessment methodology for SF6 circuit breakers in Estonian transmission network), Tallinn University of Technology School of Engineering, Department of Electrical Power Engineering and Mechatronics

- Kristen Sokk, Master's Degree, 2019, (sup) Henri Manninen; Jako Kilter, Toitekatkestuskahjude hindamise meetodid ja nende rakendatavus Eesti elektrisüsteemis (Outage cost assessment methods and application in Estonian electricity system), Tallinn University of Technology School of Engineering, Department of Electrical Power Engineering and Mechatronics.

- Keyt Auner, Master's Degree, 2018, (sup) Jako Kilter; Henri Manninen, Alajaama lülitusseadmete online mõõtmised ning nende rakendatavus seadmete tehnilise seisukorra määramiseks (Online measurements of the substation switching devices and their applicability to determine the technical condition of the equipment), Tallinn University of Technology School of Engineering, Department of Electrical Power Engineering and Mechatronics.

8. Scientific work
Papers

1. H. Manninen, J. Kilter, and M. Landsberg, "Advanced condition monitoring method for high voltage overhead lines based on visual inspection," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, Aug 2018

2. H. Manninen, J. Kilter, and M. Landsberg, "Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices," in *2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, pp. 1–6, June 2019

3. C. J. Ramlal, H. Manninen, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automated utility pole condition monitoring: A deep learning approach," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, pp. 255–259, 2020

4. H. Manninen, J. Kilter, and M. Landsberg, "Health index prediction of overhead transmission lines: A machine learning approach," *IEEE Transactions on Power Delivery*, vol. 37, no. 1, pp. 50–58, 2022

5. H. Manninen, C. J. Ramlal, A. Singh, S. Rocke, J. Kilter, and M. Landsberg, "Toward automatic condition assessment of high-voltage transmission infrastructure using deep learning techniques," *International Journal of Electrical Power & Energy Systems*, vol. 128, p. 106726, 2021

6. H. Manninen, J. Kilter, and M. Landsberg, "A holistic risk-based maintenance methodology for transmission overhead lines using tower specific health indices and value of loss load," *International Journal of Electrical Power & Energy Systems*, vol. 137, p. 107767, 2022

7. H. Manninen, C. J. Ramlal, A. Singh, J. Kilter, and M. Landsberg, "Multi-stage deep learning networks for automated assessment of electricity transmission infrastructure using fly-by images," *Electric Power Systems Research*, no. Under second review, 2022

Technical reports

1. CIGRE WG B3.48, "TB858 Asset health indices for equipment in existing substations," tech. rep., CIGRE, 2021

Conference presentations

1. H. Manninen, J. Kilter, and M. Landsberg, “Advanced condition monitoring method for high voltage overhead lines based on visual inspection,” in *2018 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, Aug 2018
2. H. Manninen, J. Kilter, and M. Landsberg, “Advanced methodology for estimation of value of lost load (VOLL) using equipment specific health indices,” in *2019 Electric Power Quality and Supply Reliability Conference (PQ) and 2019 Symposium on Electrical Engineering and Mechatronics (SEEM)*, pp. 1–6, June 2019

Elulookirjeldus

1. Isikuandmed

Nimi	Henri Manninen
Sünniaeg	14.07.1992
E-post	henri.manninen@taltech.ee

2. Haridus

2016–...	Tallinna Tehnikaülikool, Inseneriteaduskond, Elektroenergeetika ja mehhatroonika, doktoriõpe
2014–2016	Tallinna Tehnikaülikool, Energeetika teaduskond, Elektroenergeetika, MSc *cum laude*
2011–2014	Tallinna Tehnikaülikool, Energeetika teaduskond, Elektroenergeetika, BSc

3. Teenistuskäik

2019– ...	Elering AS, Varahalduse Digitaliseerimise Projektijuht
2016–2019	Elering AS, käiduspetsialist

4. Vabatahtlik töö

2016–2021	CIGRE töörühma B3.48 liige (TB858)

5. Autasud

- 2019, Täisajalise Üliõpilase stipendium
- 2019, Kristjan Jaagu välisõpingute stipendium (Trinidad ja Tobago)
- 2017, Täisajalise Üliõpilase stipendium
- 2015, Eleringi Energeetikastipendium

6. Kaitstud lõputööd

- 2016, Alajaamade ja õhuliinide kasuliku eluea määramise metoodika, MSc, juhendaja Paul Taklaja, Tallinna Tehnikaülikool, Elektroenergeetika Instituut
- 2014, Õhuliinide läbilaskevõime ja selle määramise matemaatilised meetodid, BSc, juhendaja Jako Kilter, Tallinna Tehnikaülikool, Elektroenergeetika Instituut

7. Juhendatud väitekirjad

- Tauri Kalmet, magistrikraad, 2021, (juh) Jako Kilter; Henri Manninen, Eesti ülekandevõrgu alajaamade SF6 võimsuslülitite seisundi hindamise metodoloogia arendamine ja analüüs, Tallinna Tehnikaülikool, Inseneriteaduskond, Elektroenergeetika ja mehhatroonika instituut.
- Kristen Sokk, magistrikraad, 2019, (juh) Henri Manninen; Jako Kilter, Toitekatkestuskahjude hindamise meetodid ja nende rakendatavus Eesti elektrisüsteemis, Tallinna Tehnikaülikool, Inseneriteaduskond, Elektroenergeetika ja mehhatroonika instituut.

- Keyt Auner, magistrikraad, 2018, (juh) Jako Kilter; Henri Manninen, Alajaama lülitusseadmete online mõõtmised ning nende rakendatavus seadmete tehnilise seisukorra määramiseks, Tallinna Tehnikaülikool, Inseneriteaduskond, Elektroenergeetika ja mehhatroonika instituut.

8. Teadustegevus
Teadusartiklite, tehniliste raportite, konverentsiteeside ja konverentsiettekannete loetelu on toodud ingliskeelse elulookirjelduse juures.

ISSN 2585-6901 (PDF)
ISBN 978-9949-83-791-5 (PDF)