Mercuzio Sassetto

Tree High Impedance Fault Modelling

Master's thesis in Energy and Environmental Engineering
Supervisor: Irina Oleinikova
Co-supervisor: Basanta Raj Pokhrel
June 2023

NTNU
Kunnskap for en bedre verden

Mercuzio Sassetto

Tree High Impedance Fault Modelling

Master's thesis in Energy and Environmental Engineering
Supervisor: Irina Oleinikova
Co-supervisor: Basanta Raj Pokhrel
June 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electric Energy

NTNU
Kunnskap for en bedre verden

Abstract

Europe will be subject in the coming years to major changes in the energy sector, according to the plan called "REPowerEu" the goal is to reduce greenhouse gas emissions by 55% by 2030 and achieve climate neutrality by 2050. To meet these goals, it is essential to increase the adoption of renewable energy sources in the areas of power generation, industry, and transportation. This future scenario foresees an increasingly complex power system, which makes it essential to invest in the modernization of the electrical grid infrastructure and its management to ensure proper, stable, and cost-effective operation.

This thesis work is placed under the topic of digitization of the electrical system with an emphasis on supporting the development of artificial intelligence-based methodologies for electrical fault detection. A specific type of fault affecting power distribution lines is studied, the high impedance fault caused by the contact of a live shaft with phase conductors. This fault poses numerous problems for local power distributors due to the difficulty in its detection by protective devices and because of its danger to human health and electrical infrastructure as it can also lead to serious forest fires. The development of these new fault detection methodologies often struggles with the lack of datasets: faults are not very frequent, and measurements are rarely made accessible; field experiments for generating data are very expensive. A way to bypass these difficulties is to create models that simulate the characteristics of faults, this is precisely what was done within this thesis.

A model for the single-phase fault caused by vegetation contact will be exposed, it provides the fault current trend as a function of moisture content and line voltage, thus enabling fundamental data generation for the training phase of fault detection algorithms based on artificial intelligence. The model is shown and developed for a single vegetation species, but the same methodology can be applied to different types of trees.

Table of Contents

List of figures

Fig.1 REPowerEU

Fig.2 Scheme of symmetrical and balanced three-phase network

Fig.3 Scheme of a network during a phase-ground fault

Fig.4 Three-phase network diagram at sequences

Fig.5 Single-phase fault sequence

Fig.6 Single-phase fault vector diagram

Fig.7 path of single-phase fault currents in Petersen-compensated neutral networks

Fig.8 Percentage of relay protections trips for the studied MV system in 2018 depending on the fault type.

Fig.9 Percentage of relay protections trips for the studied MV system in 2019 depending on the fault type.

Fig.10 Percentage distribution of earth faults depending on the line type in 2018.

Fig.11 Percentage distribution of earth faults depending on the line type in 2019.

Fig.12 Number of events reported in the Norwegian 1-22 [kV] grid in 2016 and the resulting energy not delivered (ILE) as function of root cause.

Fig.13 Number of events reported in the Norwegian 33-420 [kV] grid in 2016 and the resulting energy not delivered (ILE) as function of root cause.

Fig.14 Impedance-based method representation

Fig.15 Travelling-wave method representation

List of tables

List of symbols and acronyms

THIF – "Tree-High-Impedance-Fault"

HIF – "High-Impedance-Fault"

AI – "Artificial Intelligence"

V – "voltage"

I – "Current"

R – "Resistance"

Z – "Impedance"

1 Introduction

1.1 Background

In 2019, the European Union set very stringent climate targets for its member states, according to the "European Green Deal" climate neutrality by 2050 is set as a goal. Subsequently, in May 2022, the European Commission published the "REPowerEU Plan" act in which the guidelines for achieving the targets of -55% in greenhouse gas emissions by 2030 and climate neutrality by 2050 are set out [3]. The actions that member states will have to take can be summarised and divided into three categories: diversification in energy sources, save energy and a major acceleration in the transition to clean energy sources. All this will be strongly supported by a massive financial investment [3].
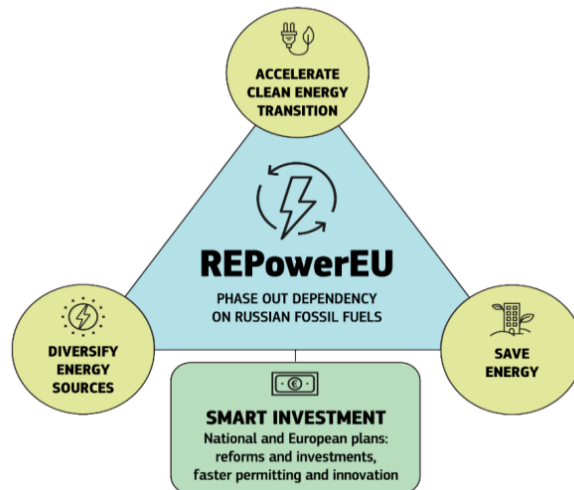
Fig.1 REPowerEU [3]

Renewable energy sources will be the drivers of this change; their increasing deployment will be required in the power generation, industry, buildings, and transport sectors. The target is to have 45% of demand covered by renewables by 2030, which requires a renewable generation

capacity of 1236 [GW]. For solar photovoltaics, being one of the fastest implementable technologies, a target of 600 [GW] by 2030 is set. Another key energy source for the transition will be wind power, especially with offshore installations, a sector in which Europe is a world leader [3]. To overcome the non-programmability of these types of sources, it will be essential to work on increasing energy storage capacity [3]. Extremely challenging targets are also set for the transport sector, for example the obligation to sell only clean cars that will be introduced in 2035 [4].

The integration of variable renewable energy sources and battery energy storage system is a big challenge for the operation of the power system (DSOs), maintain a balance in grid stability in this increasingly complicated multidirectional electricity supply system is very difficult [5]. About 70% of renewables are expected to be electrically connected to distribution networks [5], this represents a radical change in the architecture and organisation of the grid where historically the flow of power has been unidirectional: generation centers are connected to the transmission grids and (passive) users are connected to the medium and low voltage distribution grids. With a high amount of generation in the distribution level, power flows will become multi-directional, it will be possible to see, under certain conditions, power in the direction of the transmission level in high and extra-high voltage.

Aware of the problems summarised above, the "REPowerEU" plan envisages massive investments in the modernization of the electricity network [3]. The digitalization will play a key role in to transform ways grids are managed and operated [5]. The increasing spread of distributed smart grid components such as smart metering infrastructure, communication infrastructure, distributed energy resources, and electric vehicles will allow DSOs to have an incredible amount of network operation data available in their control centers, enabling extremely widespread network monitoring. This massive amount of data can support vast applications:

- Distributed energy management
- Forecasting

- Optimization of network operations and activities
- Resource optimization and strategic business decisions
- Predictive maintenance
- Maintenance and repair
- Security measurement
- Cyber security
- Adaptation/Anticipation to upcoming climate events
- Facilitating active consumers
- Product customization and marketing measures
- Automating processes for metering, billing, and general distribution [5]

A significant computational capacity will be needed to utilize all this information, and this can be provided by the development we are witnessing in the field of artificial intelligence. Many examples could be cited of the application of AI-based algorithms to the electrical system, here I would like to mention the one of the "Use of AI for vegetation management along power lines" as it is closely related to the topic that will be explored within the thesis.

The management of vegetation in power distribution is a mandatory operation to ensure a reliable supply of electricity and public safety. Trees that are too close to the power lines represent a significant hazard, putting human life and the environment in danger and are a leading cause for power outages. Each DSO develops its own "tree management" plan in accordance with legal requirements, the traditional approach to asset inspections is slow, resource-intensive, mostly involving visual inspections and field surveys is often based on fixed annual cycles. DSOs, such as "E.ON", are using new digital vegetation management processes through artificial intelligence algorithms and machine learning algorithms, making this process more efficient [5]. AI helps structure data from a variety of data sources (ground inspections, aerial images, drones, LiDAR, satellites) and extract information by transforming raw data into actionable

information based on vegetation status and forecasts (tree species, weather data, historical patterns, etc.). The use of AI enables to optimize planning of trimming on a condition-based approach built on predictive models which detect potential hazards before they occur [5].

1.2 Problem description

Using tools based on artificial intelligence, it is possible to create new methodologies for detecting faults in power lines. Faults are a major problem for grid operators as their economic impacts can be considerable: they lead to service interruptions on which operators must pay penalties; they can cause damage to infrastructure, to electricity distribution or transmission equipment; in extreme cases, to people's health. It is therefore clear that there is a great deal of interest in developing increasingly effective methodologies to prevent the occurrence of these situations.

Among the various types of artificial intelligence used for fault detection, the most popular in the literature is the "artificial neural network" (ANN) [6]. ANNs are parallel networks with a structure inspired by the human brain's neural network, they have the ability to learn, memorize and generalize. ANNs can be understood as an adaptable system that can learn relationship, and has the ability to generalize new, previously unseen, data [7]. The activation function of a neuron is non-linear, which makes neural networks capable of characterizing the non-linear parameters of a dynamic system. Their ability to capture higher-order approximations higher-order approximations for a given relation, together with their parallel computation, makes them suitable for non-linear analysis of large scale, complex, power system [7].

In general, AI new methods outperform conventional ones offering high accuracy in fault location estimation. However, where they really excel is in fault classification problems [6]. In the fault diagnosis domain the main advantages can be summarised as follows: features can be obtained

automatically by training the collecting data without manual participation; different from the traditional time-frequency analysis, the AI can adapt to data with different periods; data features high-dimension and non-linearity can be effectively processed [8].

In addition to these important advantages, however, it is worth mentioning the major drawbacks encountered by AI models. The development phase of these methodologies requires high computing power and large amount of training data. In order to improve its performance, it is often better to collect more training data rather than improve the algorithm itself [9]. The main problem is that the data needed for this training phase are often difficult to obtain for various reasons: the probability of faults in the power grid is not very high, and measurements are rarely made available; carrying out field experiments on real networks to obtain datasets is extremely expensive [8] [9] [6]. This poses a major obstacle to the creation and diffusion of AI-based methodologies for fault detection.

1.3 Research objective

This thesis work addresses the issue of limited accessibility to datasets during the training phase of fault detection algorithms based on artificial intelligence. To overcome this problem, it is possible to use models that describe the behavior of a fault and to obtain datasets by applying these models in electrical network simulation software [9] [6].

In particular, the main objective is to investigate and analyse a very specific type of fault, the single-phase faults on medium-voltage electrical distribution networks caused by vegetation contact with the phase conductors. This fault, which will be called "tree high impedance fault" (THIF), represents a major challenge for distributors due to its difficulty of detection, as it manifests itself with an extremely low fault current [10]. The analysis and investigation phase are intended to lead to the creation of a model describing the behavior of THIFs. Everything stems from a specific request from a company that is working on the creation of an AI-based tool for fault detection and needs datasets for the design phase.

Experimental data will be used to carry out this analysis and modelling (Appendix 1). In the experiment referred to, approximately 900 tests were conducted in which different species of trees were placed in contact with phase conductors at 22 kV, all voltage and current measurements were then made available [2]. These measurements will be used for the achievement of the research objectives of this thesis, summarised as follows:

- o Analysis and investigation phase of the THIF to identify the parameters that most influence the fault dynamics.
- o Creation of the tree-high-impedance fault model
- o Validation of the model.

1.4 Limitations

It is fair to point out from the outset what the limitations of this thesis work are, as it does not pretend to bring a definitive answer to an extremely complex problem that has been the subject of numerous studies.

- o The major limitation of this thesis certainly concerns the accuracy of the model, which will be described later (6.2 Discussion). This non-high accuracy derives fundamentally from the regression process, the process of estimating the unknown fault current. In this procedure, in fact, a small number of tests were used as a basis, belonging moreover to a single tree species. The procedure, which is considered to be the most important aspect, remains valid and the necessary actions to mitigate this problem will be indicated.

- o Another critical aspect concerns the fact that the model will only describe the early stages of a THIF. Such a fault after a certain period of time, which will be indicated later (4.3 "Branch on wire" characteristic), may evolve into another type of fault, the arc-ground fault [2]. It would therefore be necessary to associate the proposed model with an arc-fault model.

To present the work carried out, we will start by presenting the theory of faults in electrical networks in "Chapter 2", and then continue in "Chapter 3" with a literature review concerning statistics relating to the frequency of faults in electrical networks; fault detection methodologies; and the analysis and modelling of high-impedance faults. "Chapter 4" will describe the experiment and data structure that was used to create the model, described in "Chapter 5". It will conclude in "Chapter 6" with a presentation of the results and the validation of the model. "Chapter 7" will provide the final considerations regarding possible future research on the subject.

2 Theory

2.1 Faults Theory Introduction

In order to carry out a comprehensive analysis of faults in medium-voltage networks, it is necessary to start with the theory behind this type of event. Let's start by defining a symmetrical and balanced three-phase system.

A three-phase power distribution system is defined as symmetrical if the voltage generators that supply this system generate triads of electromotive forces that can be represented by three rotating vectors of equal amplitude and displaced by ±120°. The same system is defined as balanced if the loads have the same impedance for each of the three phases. Given any cross-section of a symmetrical and balanced three-phase network, under normal operating conditions has on that cross-section a symmetrical triad of star voltages, a consequent symmetrical triad of chained voltages, and a symmetrical triad of line currents.

If short-circuits or interruptions occur in a symmetrical and balanced three-phase network, these are to all intents and purposes non-symmetrical elements, given a generic cross-section, star and chained voltages and line currents no longer constitute symmetrical triads.
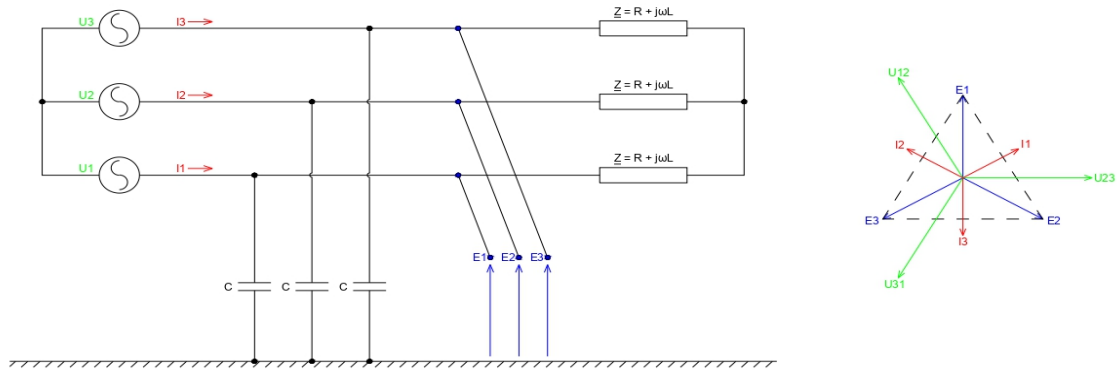
Fig.2 Scheme of symmetrical and balanced three-phase network

Knowledge of the currents and voltages resulting from faults makes it possible to solve numerous problems such as, for example, those concerning the correct choice of circuit breakers intended to interrupt short-circuit currents, those relating to measuring and protection equipment intended to warn of faults and thus intervene appropriately, those relating to the thermal and electrodynamic stresses of the various system elements, and those concerning the stability of synchronous machines. For all these reasons, it is necessary to calculate currents and voltages arising in a three-phase system that has become dissymmetrical due to a fault [11] [12].

Fig.3 Scheme of a network during a phase-ground fault

8

2.2 Sequence theory

The dissymmetrical triad of voltages and currents in the fault section of a three-phase dissymmetrical system can be obtained as a vector sum of three symmetrical triads: specifically, as the sum of a direct symmetrical triad rotating clockwise, a symmetrical inverse triad rotating counterclockwise, and a homopolar triad [11].

The electrical components of a system, i.e. overhead lines, cable lines, loads, synchronous and asynchronous machines change their behavior depending on the nature of the symmetrical triad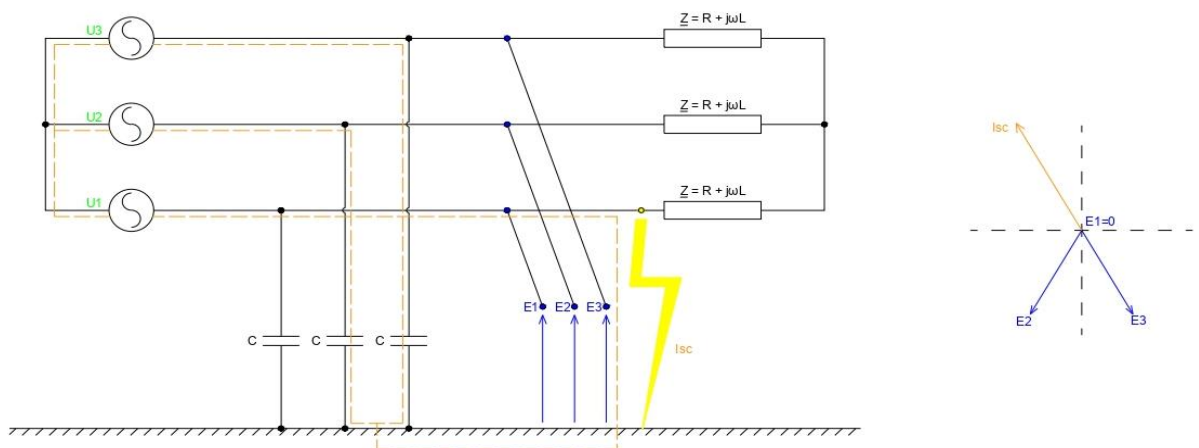 of voltages with which they are supplied. In order to apply the analysis method to the three sequences, it is essential to know the equivalent electrical model of these components in the face of a direct symmetrical triad, an inverse symmetrical triad, and a homopolar triad [11].

Consider a three-phase electrical system consisting of three phase conductors and a neutral (in case it is not present, the ground potential is assumed to be zero). Applying the principle of the equivalent voltage generator, the complex electrical system viewed from a direct symmetrical triad of voltages is equivalent to a real three-phase generator whose electromotive forces are equal to the no-load voltages that would be measured between terminals 1, 2, 3 and the neutral. The internal impedances of the equivalent three-phase generator, on the other hand, correspond to the impedances that would be measured between each phase and the neutral once the complex electrical system is made passive [11].

Reasoning in an analogous manner, it is possible to construct the equivalent voltage generator seen from terminals 1, 2, 3 and the neutral when a symmetrical inverse triad of voltages is applied to them: the electromotive forces are equal to the open-circuit voltages that would be measured between terminals 1, 2, 3 and the neutral. for the impedances, the same procedure as for the direct sequence is followed.

Finally, the equivalent bipole at sequence zero is constructed by applying a triplet of homopolar voltages to terminals 1, 2, 3. The real voltage

generator at the zero sequence will present electromotive forces that correspond to the open-circuit voltages present between ports 1-0, 2-0, 3-0 of the complex electrical system, while the equivalent impedances will correspond to the impedances seen by a homopolar triad of voltages applied to input terminals 1, 2, 3 once the system is made passive [11].
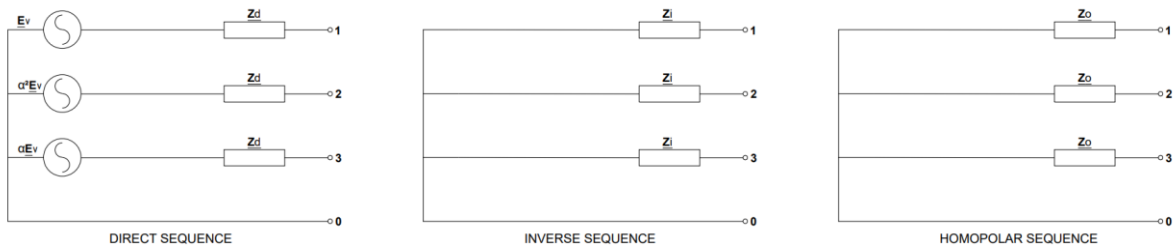

Fig.4 Three-phase network diagram at sequences

Given the bipoles at the different sequences, proceed to see how the mathematical expressions describing a phase-earth fault can be derived. At the single-phase fault, the three sequences are connected in series (Fig.5 Single-phase fault sequence) but consider that as the type of short circuit or interruption changes, the type of connection of the three equivalent bipoles to the different sequences changes. For example, the phase-to-phase fault has direct and reverse sequences connected in parallel.

(1)
$$\begin{cases} \underline{Ev} - \underline{Zd}\,\underline{Id} = \underline{Ed} \\ -\underline{Zi}\,\underline{Ii} = \underline{Ei} \\ -\underline{Zo}\,\underline{Io} = \underline{Eo} \end{cases}$$
"Starting characteristic equations".

(1.1)
$$\begin{cases} \underline{I1} = 3\underline{Id} = \dfrac{3Ev}{\underline{Zd}+\underline{Zi}+\underline{Zo}} \\ \underline{I2} = 0 \\ \underline{I3} = 0 \end{cases}$$
"Phase-earth short-circuit current".

(1.2)
$$\begin{cases} \dfrac{\underline{E1}}{Ev} = 0 \\ \dfrac{\underline{E2}}{Ev} = e_2 = \dfrac{Zo\,(\alpha^2 - 1) + Zi\,(\alpha^2 - \alpha)}{Zd + Zi + Zo} \\ \dfrac{\underline{E3}}{\underline{Ev}} = e_3 = \dfrac{Zo\,(\alpha - 1) + Zi\,(\alpha - \alpha^2)}{Zd + Zi + Zo} \end{cases}$$
"Sustained overvoltages at 50Hz".

10
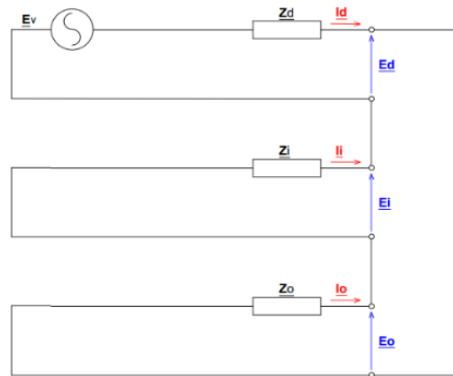
Fig.5 Single-phase fault sequence

The above equations refer to a net earth fault with zero fault impedance. In the following, a high-impedance fault caused by the contact of a live tree with an overhead line phase conductor will be analyzed. Let's see how the equations change in the presence of a non-zero impedance.

(1.3)
$$\begin{cases} \underline{I1} = 3\underline{Id} = \dfrac{3\underline{Ev}}{\underline{Zd}+\underline{Zi}+\underline{Zo}+3\mathbf{Zg}} \\ \underline{I2} = 0 \\ \underline{I3} = 0 \end{cases}$$
  "ph-e current, non-zero impedance"

2.3 State of neutral

As can be seen from the equations (1.1) (1.2), homopolar impedance plays a key role. The value of this impedance in electrical networks is closely related to the state of the neutral [11] [12]. In this thesis work, the focus will be on the analysis of faults in medium-voltage distribution networks; in the following, therefore, the neutral state of medium-voltage networks only will be analyzed.

The two most common ways of operating the neutral in medium voltage networks are isolated neutral and neutral compensated by means of a Petersen coil [12] [11] to limit single-phase fault currents.

11

To derive the value of the single-phase fault current in medium-voltage networks with an isolated neutral, the value of the impedance at the homopolar sequence in this configuration must be derived. The Zd and Zi will depend on the distance of the fault point from the busbar in the transformer substation, the homopolar impedance Zo, on the other hand, is equal to the sum of the capacitances to earth of all lines electrically connected to the busbar in the transformer substation [11].

Considering Zd, Zi << Zo:

(1.4) $Ig = \frac{3\,j\omega C_0 E}{1+j\omega C_0 Z_g}$

In the case of zero fault impedance, the expression can be approximated as follows:

(1.5) $Ig \cong 3\,j\omega C_0 \underline{E}$

This current will be higher in the presence of cable lines which have a higher capacity.

In addition to the calculation of the fault current Ig, the analysis method based on the application of the symmetrical component theory allows the determination of the overvoltages to which the healthy phases are subjected, so that it is possible to draw the vector diagram of voltages and currents present in the fault section following a phase-to-ground fault occurring in an isolated neutral system.

"Es", "Et" are called sustained overvoltages at 50 [Hz]. "V0" is the voltage to earth that is generated on the operating capacitances of the healthy phases (V0=3*E). This overvoltage is quite dangerous as it stresses the insulation of the electrical components affected by it during the time the fault lasts [12].
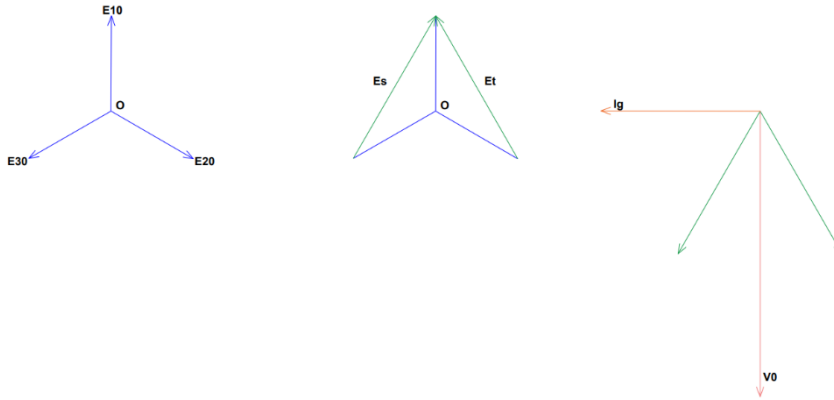
Fig.6 Single-phase fault vector diagram

The other most common type of neutral operation is the compensated neutral. The star point of the transformer is connected to earth by means of a pure inductance (in the transformer substation) realized via a coil known as a Petersen coil. In the event of a phase-to-ground short, the fault current closes through both the healthy phase capacitances and the Petersen coil. If this inductance is perfectly tuned, the earth fault current Ig is extinguished, so the Petersen coil is also referred to as a self-extinguishing coil.

In the single-phase fault situation, phase r to ground, shown in the image above, the short-circuit current takes on the following value [12] [11]:

(1.6) $$Ig = \frac{E_r (1-3\omega^2 LC)}{j\omega L + Zg(1-3\omega^2 LC)}$$

If a perfect tuning is considered, there is a condition of resonance between the inductance of the Petersen coil and the operating capacitance of all lines electrically connected to the busbar in the transformer substation. Resonance is obtained at:

(1.7) $$L = \frac{1}{3\omega^2 C}$$

In such condition there will be zero single-phase current [11] [12].

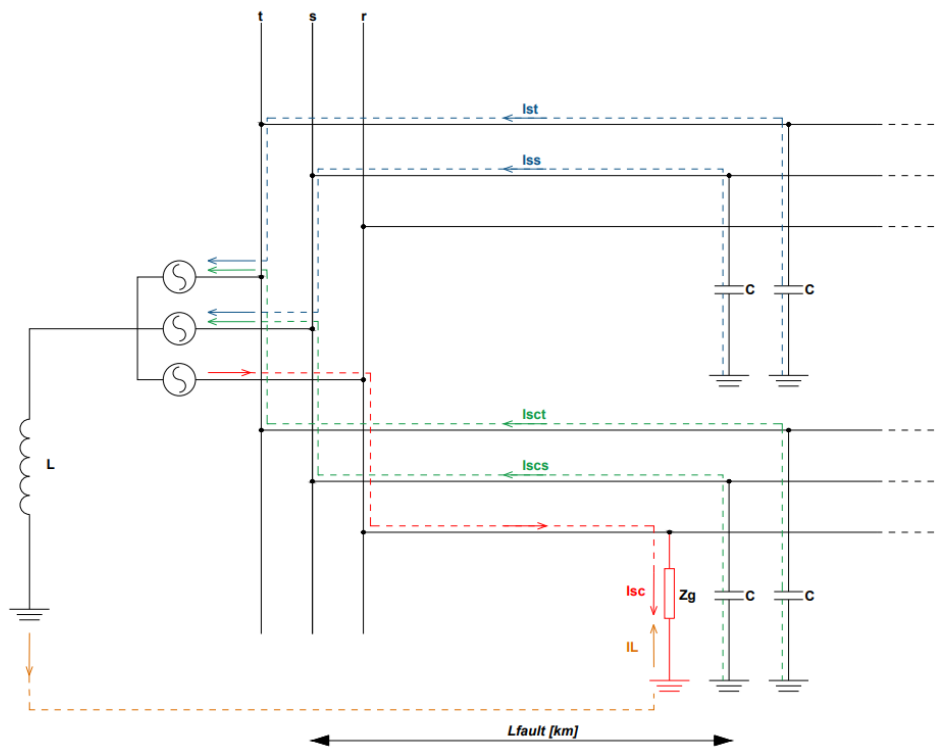Fig.7 Path of single-phase fault currents in Petersen-compensated neutral networks

3 Literature Review

3.1 Statistics on power grid failures

After having set out the basic theory for the study of single-phase faults, it is useful to review which types of faults are most common in distribution networks by looking at statistics, and evaluate which faults are most dangerous and most difficult to detect.

We are particularly interested in assessing the frequency of fault events in medium-voltage distribution networks, and we begin by referring to a study carried out in Bulgaria [1] . A 110/20 kV transformer substation was considered, and the activation of line protections was observed for two years. The used relay protections against earth faults were the "ANSI 50N" and in case of phase-to-phase, the "IP" and the "TOCP ANSI 50". The type and number of power lines connected to the 20 kV supply busbar is summarized in the table below.

Table 1: number of lines analyzed

Line Type	Overhead line (OPL)	Underground cable (CL)	Mixed OPL & CL
N° of lines	7	7	4

After the two-year observation period, it was seen that single-phase earth faults are extremely more common and frequent. The share of phase-to-phase faults is 10%-11% versus the 89%-90% of the earth faults.
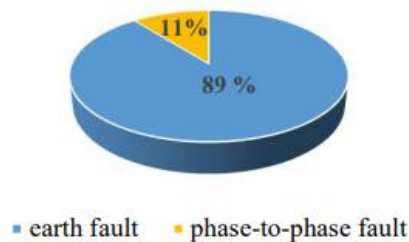
Fig.8 Percentage of relay protections trips for the studied MV system in 2018 depending on the fault type [1].

15
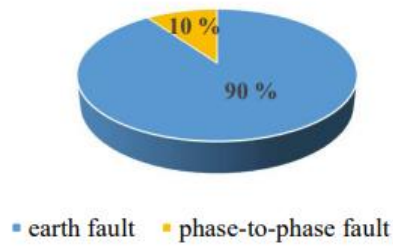
Fig.9 Percentage of relay protections trips for the studied MV system in 2019 depending on the fault type [1].
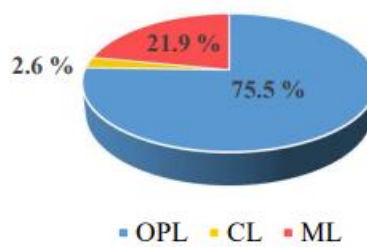
Fig.10 Percentage distribution of earth faults depending on the line type in 2018 [1].
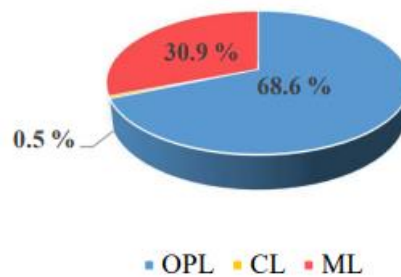
Fig.11 Percentage distribution of earth faults depending on the line type in 2019 [1].

In conclusion, we can state that the most common faults are single-phase earth faults affecting overhead lines, they are more vulnerable to faults than cable lines because they are exposed to numerous environmental factors, lets see what is referred in Norwegian reports [13].

In the Norwegian power grid, there is a mandatory responsibility given to the DSOs and the TSO to report faults and disturbances via a national reporting structure called FASIT. This document specifies the type of incident, the number of affected phases, voltage level, the time duration and plausible root cause and contributing causes. Statistics are divided by voltage levels, 1-22 kV and 33-420 kV, the graphs show the number of failure events and the resulting energy not delivered [13].
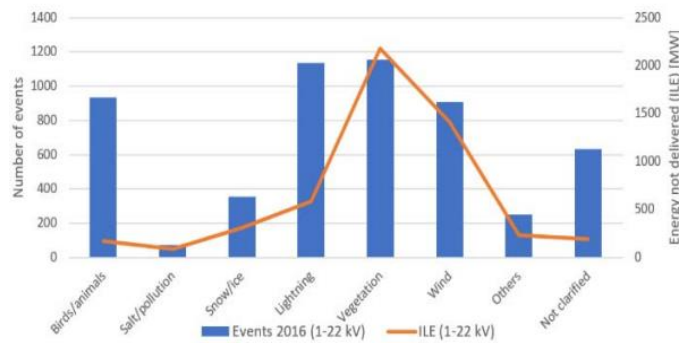
Fig.12 Number of events reported in the Norwegian 1-22 kV grid in 2016 and the resulting energy not delivered (ILE) as function of root cause [13].
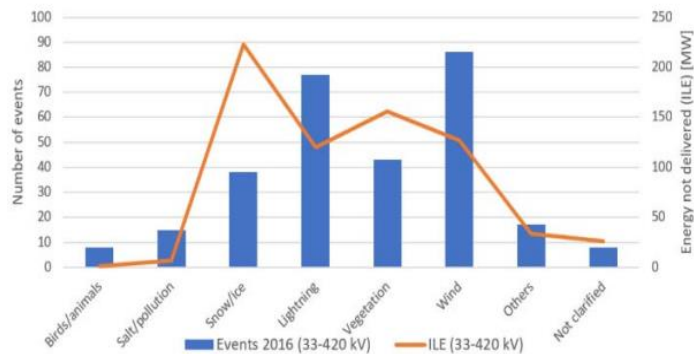
Fig.13 Number of events reported in the Norwegian 33-420 kV grid in 2016 and the resulting energy not delivered (ILE) as function of root cause [13].

As can be seen, the number of events and consequences is far larger in the 1-22 kV grid. Vegetation and wind are two of the major root causes, the events caused by vegetation, which will be discussed in detail in this thesis, are between 70 and 80 per year and have the highest amount of undelivered energy in Norway's distribution systems [13].

3.2 Impedance-based method

In view of the severity and major economic impact of faults in power lines, fault detection, localisation and prediction are of paramount importance. we will now discuss and analyse the most widespread methodologies and those under development:

I. Impedance-based method

II. Travelling waves method

III. AI-based method

The impedance-based methodology applies Kirchhoff's laws and utilizes the voltage and current measurement before and after the fault for locating the point of failure into the grid, is probably the simplest method and one of the first to be used [6].

Voltage and current measurements and the line impedance matrix are used to form equations that calculate the overall impedance between the feeder and the fault point, and subsequently the distance of the fault from the busbar.

$$(3) \qquad \begin{bmatrix} U_{Fa} \\ U_{Fb} \\ U_{Fc} \end{bmatrix} = \begin{bmatrix} U_{Sa} \\ U_{Sb} \\ U_{Sc} \end{bmatrix} - d \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} \\ Z_{ba} & Z_{bb} & Z_{bc} \\ Z_{ca} & Z_{cb} & Z_{cc} \end{bmatrix} \begin{bmatrix} I_{Sa} \\ I_{Sb} \\ I_{Sc} \end{bmatrix}$$

The equation (3) is used to calculate the voltage "UF" at fault point in a line without branches. "Us" and "Is" are the voltage and current at the local end respectively; "d " is the distance between the fault and the beginning of the line and "Zxy" is the impedance between phases x,y.

An important parameter of the equations is the fault resistance "Rf", it is a real value and for the elimination of the fault resistance, is consider an unknown parameter. Firstly, the equations are split into their real and imaginary part. The fault distance can then be obtained by the resolution of the equation formed, after the imaginary part is set equal to zero.

$$(3.1) \qquad \begin{bmatrix} d \\ R_f \end{bmatrix} = \frac{1}{M_{1n}I_{Fni} - M_{2n}I_{Fnr}} \begin{bmatrix} I_{Fni} & -I_{Fnr} \\ -M_{2n} & M_{1n} \end{bmatrix} \begin{bmatrix} U_{Snr} \\ U_{Sni} \end{bmatrix}$$

(3.1) is utilized for location of a single-phase fault current occurring at phase n in a line without branches. "r" and "i" refer to real and imaginary part of the variables; "U_{Sn}" is the voltage at the local end; "I_{Fn}" is the fault current; "M_{1n}" and "M_{2n}" are defined as:

(3.2) $\qquad M_{1n} = \Sigma_k \left(Z_{nk_r} I_{Sk_r} - Z_{nk_i} I_{Sk_i} \right)$

(3.3) $\qquad M_{2n} = \Sigma_k \left(Z_{nk_r} I_{Sk_i} + Z_{nk_i} I_{Sk_r} \right)$

Where "k" equals to one of the three phases each time; "Z_{nk}" is the impedance between the phases n,k; "I_{Sk}" is the current at the local end. The calculation can be performed directly or through an iterative process [6].

In the case of a line with laterals (Fig.14), a set of equations describes each line section and if the result of the computations surpasses the length of the line section, the next section's equations are solved. The power flow is calculated each time for the evaluation of the current and voltage on the bus at the beginning of the next line section. If the section contains loads, the phasors preceding the fault are used to calculate the equivalent load impedance on each bus. The final result is the distance "fd" between the fault and the nearest preceding bus, which when added to the cumulative length of the line sections, between the fault bus and the beginning of the feeder, gives the total distance of the fault from the feeder. Different set of equations correspond to each type of fault; therefore a classification of the fault type is usually required before the application of an impedance-based method [6].
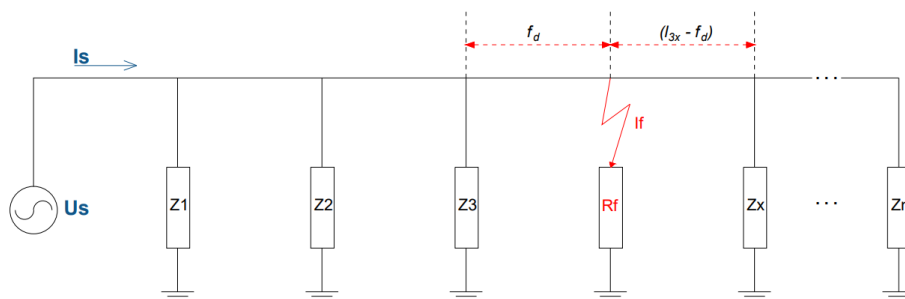
Fig.14 Impedance-based method representation

3.3 Travelling-wave method

Another common and developed method is the one based on the travelling wave theory. Following the occurrence of a fault, a high-frequency waveform appears and travels between the two terminals of the line; the time it takes to reach the end of the line depends on the location of the fault point. Since the speed of the wave is considered known, the distance "fd" from one of the end can be simply calculated by:

(3.4) $\qquad f_d = \frac{l - v(t_a - t_b)}{2}$

Where "ta" and "tb" are the time that the wave takes to arrive at the two terminals; "v" is the velocity of the wave and "l" is the length of the line [6].

It is possible to make measurements on a single terminal by exploiting the travelling wave reflection phenomenon. the wave is reflected at the extremities and the time it takes from one terminal to the fault point and back is evaluated, this allows the location of the fault to be estimated using the following equation:

(3.5) $\qquad f_d = \frac{v(t_2 - t_1)}{2}$

Where "t1" and "t2" are the time that the wave takes to travel to the fault point and back respectively; "v" is the velocity of the wave and "fd" the distance from one terminal of the line.

Fig.15 Travelling-wave method representation

It is now useful to list the advantages and disadvantages of these two methods of fault location. The main drawback of the impe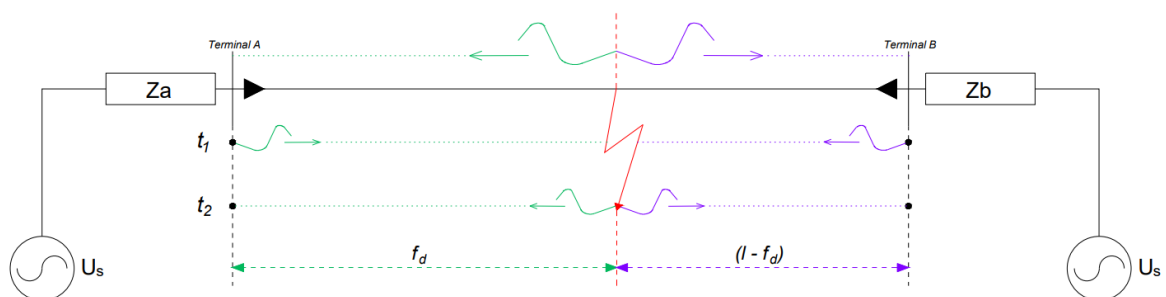dance-based methods is the multiple location estimation in large grids with laterals, the method also depends on the system loading conditions during the fault which can affect its accuracy [6]. Additionally, methods utilizing current measurement may be affected by the saturation of the current transformer, in the end, implementation in grids with bi-directional power flows is a challenge. Nevertheless, it is easy to implement with few requirements.

On the other hand, traveling-wave methods requires very specific equipment able to record the transient waves of voltage and current, and their arrival times. This leads to a significant increase on the cost of this method. Furthermore, in the one-terminal methods, the isolation of the wave of interest can be obstructed by unwanted reflections and noise added by other equipment and junctions of the grid. Moreover, most of traveling-wave methods refer only to single phase faults, as previously mentioned, they are the most frequently encountered. However, single–ended methods are vastly used mainly because they do not depend on any kind of communication and synchronization between devices [6].

3.4 AI-based method

In the last few years, the amount of measuring instruments and sensors in the power grid has increased significantly through, for examples, power quality analysers (PQA), phasor measurement units (PMU), smart meters and other types of sensors (line angle, vibration, wind, humidity, temperature, etc.). this increased of data from the power grid in combinations with the development of machine learning and neural networks gives a lot of new possibilities in terms of fault detection and mitigation [13].

For the future application of fault detection methods based on artificial intelligence real-time monitoring of the electricity system will be required and will be necessary high-resolutions measurements devices capable of

communicating their detections with low latency to a centralized computing unit. Two possible devices offering these properties are PQA and PMU.

Power Quality Analysers, as the name suggest, are used to analyse the power quality in the grid, usually have been installed in important junctions in the grid, such as the busbar on the secondary side of a HV/MV transformer. PQAs give information about voltage parameters, e.g. harmonic distortion, voltage variation and transients. This device has superior sampling rate, compared to PMUs, with a bandwidth of up to 25 kHz and are capable to efficiently reproduce and store continuous waveform for large periods of time. On the other hand, Phasor Measurement Units provide time-synchronized voltage and current phasors at a reporting rate in the range of 10-60 Hz, each reported frame contains phasors constituted by angle and magnitude which correspond to the phase and amplitude of measured quantity, usually three phase voltages and currents [13].

For more comprehensive network monitoring, Smart Meters (SMs) are becoming more widespread, an SM is an advanced measuring instrument that determines real-time consumption of electricity, records, and storages this data at predefined time intervals. SMs are able to define the energy consumption rates by sensing the voltage, phase angle, frequency and securely communicate that data. One of the most interesting things about these devices is that they can be programmed to monitor and to control home appliances and customer's equipment e.g. air conditioners, heaters, lighting, etc. in the European Union Sweden and Finland are leading countries to carry out metering reform and large-scale smart meter deployment, and now have more than 80% of SMs implemented in their power systems [14]. Another good example is Italy, where through the "Decreto Legislativo 4 Luglio 2014 n°102", has been introduced the obligation to install smart meters in order to reduce non-technical losses, this decree implemented provisions of the European Electricity Directive 2009/72/EC, which provides for the adoption of smart meters as a tool for the more efficient management of electricity [15] [16] .

As outlined above, we are seeing a significant increase in the available grid monitoring data, which together with the need for load profile forecast will pose more challenging management for the operators. A great help will come from the evolution in the fields of big-data analysis and machine learning, such as deep neural networks.

Machine learning problems, such as fault diagnosis for the power system, are usually divided in classification and regression. Classification problems are those that require the prediction of a specific label, i.e. the assignment of input to a certain category, while regressions are those that require a quantitative prediction. Both types can be encountered in fault diagnosis methods, for example, in a distribution grid, determining if a fault has occurred or which grid branch or sector is under fault are classification problems because a distinction between healthy and faulty conditions is required. Determining the fault type is another classification problem, as the output can only be one of a group of fault types. On the other hand, predicting the exact point of a fault on the line is a regression problem because the predicted value is a number either in meters or in percentual units [6].

The most popular among artificial intelligence methods are the Artificial Neural Networks (ANNs), they are able to identify patterns among non-linear related variables, simplifying they are a model that given an input x predicts a variable y based on certain parameters. ANNs are trained offline using dataset containing both faulty and normal operation data, and they are tested in an out-of-sample dataset to evaluate their performance. The majority of the proposed ANNs present similar structures with only some design choices differentiating them from each other. The determinants that distinguish the various types of neural networks are the dataset split strategies, the input data, the training algorithms, the activating functions, the number of neurons and number of layers [8] [6]. The Fig.16 shows a basic structure of a neural network.

Other methodologies have also been developed based on artificial intelligence, for example the one with the use of Support Vector Machines

(SVM), those based on Fuzzy Logic (FL) and others that are based on genetic algorithms.
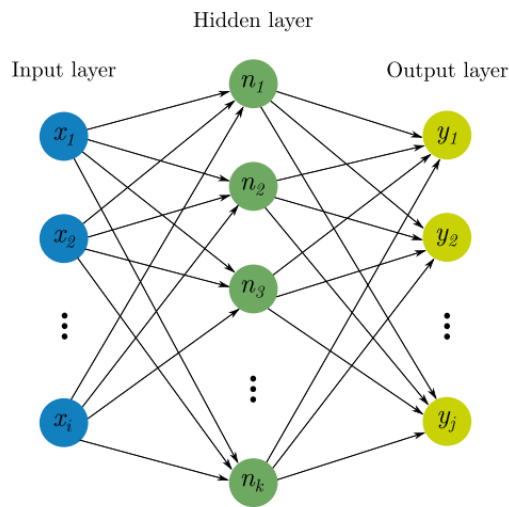

Fig.16 AI-Neural network

3.5 High Impedance Fault Review

Having described the main methods of fault detection in electrical networks, we proceed with a description of an extremely problematic type of fault, this will help us to understand the characteristics of a THIF.

High impedance faults (HIF) are one of the fault types that pose the greatest challenges to electrical grid operators. The HIFs, term normally applied to faults in distribution grids (10-25 kV), are a group of power system disturbance which present a very low fault current, lower of the sensitivity of protections, due to high grounding impedance. Referring to equations (1.3) in Chapter 2 it can be see that high "Zg" leads to a very low current.

The most prevalent physical signature of the HIF is the fact that it is often accompanied with the AC electric arc, for this reason is usually called high impedance arcing fault. The arc ignition occurs once the magnitude of the voltage of the conductor touching the surface exceeds a specific voltage

limit, the break down voltage "Vbr". Afterwards, the arc extinction occurs once the voltage magnitude becomes lower than Vbr. The Value of Vbr stays at different constant values for each half cycle of voltage [17], in each cycle of voltage the HIF results in at least two arc re-ignitions and two arc extinctions. In addition to the low current magnitude and arc this fault leads in several physical and electrical characteristics:

a. **Asymmetry in current waveform**, due to difference in the break-down voltage against positive and negative voltage value, there is a different between peak value and shape of positive and negative half cycles of the fault current.

b. **Intermittence of the arc**, the arc does not generate current in a steady state pattern, instead it generates a few conductions cycle followed by several non-conduction cycles.

c. **Current buildup and shoulder**, current magnitude gradually escalates until it reaches and maintains a constant value for many cycles.

d. **Non-stationary current**, the current frequency spectrum varies with time.

e. **Randomness**, both the current magnitude and its conduction/non-conductions intervals are random values.

f. **Non-linearity**, voltage and current characteristics of the HIF is non-linear due to the electric arc.

g. **Low frequency components in current and voltage waveform,** in the HIF waveforms can be seen harmonics up to 600Hz for the current and up to 300Hz for the voltage.

h. **High frequency components in current waveform**, arc results in high frequency components in the current.

The HIF, additionally, depends on various factors such as ground surface material, surface humidity, feeder configuration, voltage levels, weather conditions and load type. The two most influential factors are surface humidity and surface material, higher surface moisture results in higher fault current magnitude [17].

Materials that are mostly touched by conductors are: tree branches and vegetation, lawns, gravel, stout gravel, thin gravel, asphalt, concrete, sand, board blocks and cement. These will lead to high impedance. Unlike most electrical system failures, which mainly endanger electrical equipment, HIFs pose a threat to people's safety and could also legal problems. Followed by HIF inception, the generation and accumulation of flammable gas, or the evolution of the arc to close flammable material could results in fire explosion. In addition to these dangers, there is the possibility that line crewman or people would have unintentional contact with the energized exposed conductor or energized surfaces [18].

Certain and comprehensive reports about statistics of the HIFs are missing because only faults that result in relay alarm or trip are reported, however it is estimated that between 5% and 10% of the distribution faults are HIF [19]. For these reasons it is a crucial task to detect this type of faults using algorithms specifically designed to deal with these problems [17].

Create an accurate modelling of high-impedance faults based in realistic measured data is essential in controlling and improving the performance of protections systems for identification and location process. The causes of a high-impedance fault can be of various types as outlined above, so it is not possible to use a single model to describe all HIFs. Most of the HIF model suggested are electrical models, which consist of elements like variable impedances, diodes, DC voltage sources and controlled switches, and thermal models which are used to represent the electrical arc dynamics [20]. In manly of them, similar V-I characteristics were used. Most of the proposed models focus on the study of faults caused by phase conductor breakage [10], in this work the aim is to study faults caused by conductor contact with vegetation.

A Tree-related HIF (THIF) is a complex type that occurs when a live tree comes in contact with a power line, this is difficult to be detected by protections because the fault current draw in the mA regime [10]. Failure to detect THIF early causes line interruptions and consequently possible blackouts (for example some blackouts took place in North America, Italy,

Southern Sweden, and Eastern Denmark when a tree caused critical lines to trip and extended the blackout) [21]. The starting point for the study of this type of failure will be to expose the vegetation conduction ignition exhibiting the development of electrical variables. Using experimental data, a model of the high impedance fault caused by live trees will be constructed.

## 3.6 High Impedance Fault Modelling

Numerous model proposals for high-impedance faults can be found in the literature, so we began with an analysis of the best type of model that could be adapted to the problem of modelling the first two phases of the THIF fault.

The most common types consist of electrical models using elements such as variable impedances, controlled switches, and diodes [10] [22]. Some of these will be described in the following.

MODEL 1 – "Sharaf and Wang" (2003)

This is one of the simplest models; the fault impedance is modelled using a variable resistance Rf and a constant inductance (Fig.17).

The variable resistance is calculated using the following formula:
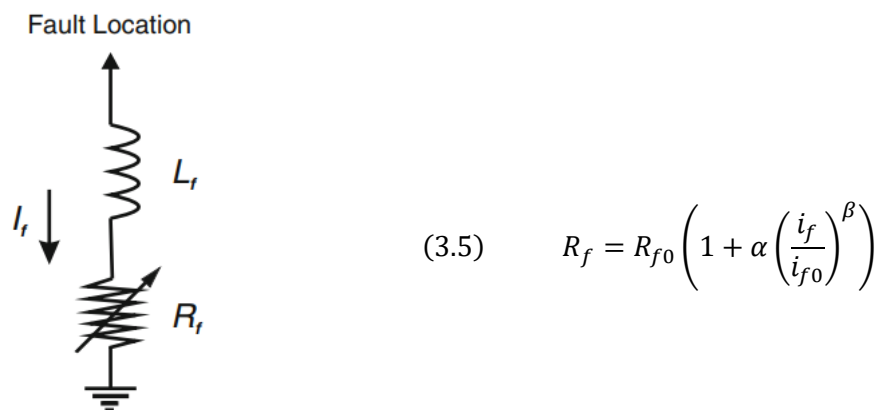
$$(3.5) \qquad R_f = R_{f0}\left(1 + \alpha\left(\frac{i_f}{i_{f0}}\right)^{\beta}\right)$$

Fig.17 "Sharaf and Wang" model [22]

Where R_{f0} is the resistance at the beginning of the fault; alpha and beta are default parameters; i_f and i_{f0} are instantaneous and initial current, respectively. The major limitation of this model is that it only simulates the non-linearity and asymmetry characteristics of the current.

MODEL 2 – "Emanuel et al." (1990)

The model constructed as illustrated in Fig.18 consists of a constant inductance in series with a constant resistance, are then combined with two diodes connected in antiparallel. Each diode is connected with a continuous voltage generator.

The fault current circulates thanks to the voltage sources V_P during the positive half-period, during the negative half-period thanks to the generator V_N. The V_N voltage will be higher than the V_P voltage because during the positive semi-period the current is slightly higher than during the negative semi-period. The presence of harmonic content is controlled by the difference between the two sources ($\Delta V = V_N - V_P$) and the ratio of reactance X_L to resistance. As with the previous model, only the phenomenon of asymmetry and non-linearity is simulated [22].
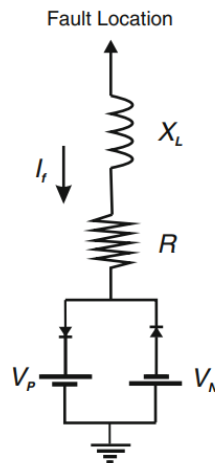

Fig.18 – "Emanuel et al. Model" [22]

MODEL 3 – "Nam et al." (2001)[19]

This model is realized by connecting two variable resistors in series, which are controlled using a Transient Analysis of Control Systems (TACS) routine from ATP.

The resistor R1 simulates the phenomenon of non-linearity and asymmetry of the current, acting in the transient and permanent regimes. The value of this resistor is obtained by the ratio between voltage and current within a cycle in which there are no significant changes in amplitude in relation to subsequent cycles. Therefore, it is assumed that in the selected cycle there was no influence of the accumulation and shoulder phenomena, which allows the division between the voltage at the fault point and the corresponding current to result in the resistance simulating the asymmetry and non-linearity characteristics of the HIF current [22] [23].

On the other hand the resistance R2 simulates the phenomena of buildup and shoulder (period of elevation of the fault current remains constant before beginning to rise).

Its value is calculated by dividing the absolute maximum values of the half-cycles of the voltages by those of the currents. Since the steady-state operation of the system only has the influence of R1, the value at which it has been stabilized (R1) is subtracted from the curve obtained above and the behavior relative to the accumulation and shoulder characteristics (R2)

is obtained. Finally, a regression analysis is performed to approximate the points of R2 with a time-dependent equation. Thus, this model simulates the characteristics of non-linearity, current asymmetry, buildup, and shoulder.
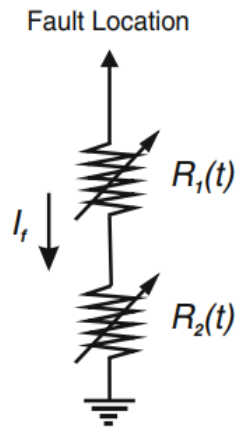

Fig.19 – "Nam et al. Model" [23]

MODEL 4 – "Bahador et al." (2017) [10]

This model differs completely from the previous ones in that it does not use electrical components to reproduce the current pattern.

A mathematical process using experimental data as a starting point is proposed here; the process is divided into four levels: namely the data collection; model structure selection; model parameter estimation and model validation.

The model structure is a mathematical correlation between input and output variables which are, respectively, conductor voltage and fault current. The model has to be nonlinear because the fault current depends non-linearly on conductor voltage. For this reason the input-output correlation is decomposed into three interconnected elements, the dynamics are modelled using a linear transfer function and the nonlinearities are represented using

non-linear functions. This structure is achieved by Hammerstein-Wiener model (Fig.20).

Fig.20 - Hammerstein-Wiener model [24]

Where [24],

- "f" is a nonlinear function that transforms input data u(t) as w(t)=f(u(t)).
- "B/F" is a linear transfer function that transforms w(t) as x(t)=(B/F)w(t). x(t), an internal variable, is the output of the Linear block and has the same dimension as y(t). for "ny" outputs and nu inputs, the linear block is a transfer function matrix containing entries: $\frac{B_{j,i}(q)}{F_{j,i}(q)}$ where j=1,2,…, ny; and i=1,2,…,nu.
- "h" is a nonlinear function that maps the output of the linear block x(t) to the system output y(t) as y(t)=h(x(t)).

This type of model structure is a feed forward block-oriented structure. The following step is the parameter estimation, an iterative process to minimize the error between simulated and experimental data [10].

4 Data and Setup for fault modelling

In Chapter 3.5, the main methodologies for modeling high-impedance faults were described. In the subsequent part of this thesis work (Chapter 5), a methodology will be presented for creating a THIF model, inspired by the "Nam et al." model (MODEL 3). The procedure in Chapter 5 will involve the

utilization of experimental data as a foundation; therefore, it is essential to describe the field experiment that will be referenced.

4.1 Vegetation conduction ignition test

The best way to start constructing a model for the tree-related HIF is to analyse experimental data. This text will refer to data from an experiment carried out in Australia (Appendix 1).

In the "Black Saturday bushfires" in Victoria (Australia) in February 2009 a series of devastating bushfire occurred, some ignited by electrical faults, this resulted in the death of 173 people, more than 4000 injured, and caused $4.4 billion of economic damage to the state [2]. After that the Victorian Government established the "Power Bushfire Safety Taskforce" called for targeted research to improve the state of knowledge on how fires are ignited by power lines faults and how to prevent them. One of the projects, the "Vegetation Conduction Ignition", around twenty different plant species were tested in a hundred experiments, on a real three-wire 22 kV feeder, to study fault behavior and promote the development of protections technologies [2].

The conduction ignition experiment design was relatively simple: different fault configurations generated electric current through vegetation and ignition results were recorded and observed, together with fault signatures. The chosen test configurations were:

  I. **Wire into vegetation** – a conductor dropped into or sitting in earthed vegetation, either grass or a bush.
 II. **Branch touching wire** – branch laid across two conductors, one earthed and one with full nominal phase voltage applied.
III. **Branch across wires** – branch laid across two conductors connected to two separate phases of the incoming voltage supply, i.e. with 22 kV between them.

The ignition test area contained the test rig that supported the test sample on two parallel sections of a typical aluminum power line conductor (19/3.25AAC). The conductors could be connected to medium-voltage busbars mounted on the ceiling by means of flexible drop cables or they could be earthed depending on the desired operating configuration for the test. With both conductors connected to the high voltage busbars, the test rig operated in the phase-to-phase configuration to simulate "branch across wires" faults. An earthed trolley was used to support vegetation samples for grass and bush tests to simulate "wire into vegetation" faults [2].
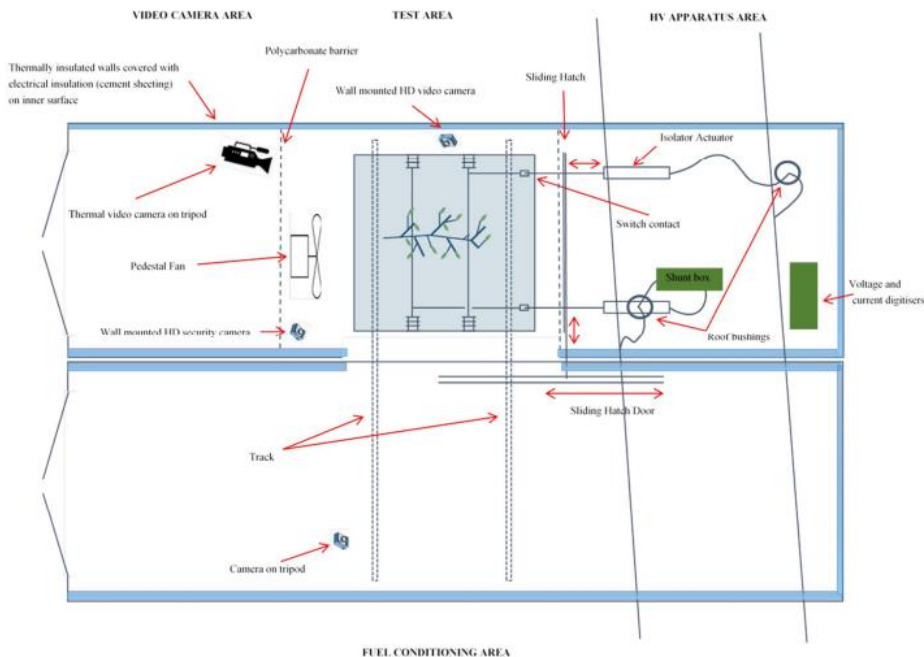

Fig.21 22 kV conductors and vegetation in contact

Fig.22 Test layout

4.2 Electrical measurement system features

For current measurement was used a HiLo Test 0.2Ω coaxial shunt, to enable wide-band and low-noise recording. The voltage across the shunt was fed to digitizers via two separate channels:

- Low-frequency channel – the shunt output signal was sampled at 100kS/s with suitable anti-alias filters. This provided a cut-off frequency of 50kHz and a record of currents up to 7Arms with 14-bit (2.4mA) resolution; wide-band noise levels below about 3mA; Narrow-band baseline noise in computed spectra was about 1μA [2].

- High-frequency channel – the shunt output signal was fed through a three stage RC high-pass filter with 10kHz cut-off frequency. It was sampled at 2MS/s with suitable anti-alias filters. Was provided a wide-band (10kHz – 1MHz) records of high frequency current components up to 180 mArms with 14-bit (30μA) resolution and wide-band noise levels below about 300μA. The narrow-band baseline noise levels in computed high-frequency voltage spectra were observed to be around 200nA [2].

For the voltage measurement two Omicron 24kV 1.1nF coupling capacitors were combined with bottom-end capacitors to form a dual channel capacitive voltage divider (CVD). Two layers of over-voltage protection were included in each channel; a 125V bidirectional voltage limiting diode embedded in the coupling capacitor itself and a 350V spark gap contained in the bottom-end termination box. The two channels of the CVD were recorded differently:

- Low-frequency channel – a 2.2μF bottom end provided a 2000:1 ratio; the output signal was sampled at 100kS/s with suitable anti-alias filter. This provided a cut-off frequency of 50kHz and a record of voltages up to 14kVrms with 14-bit (2.44V) resolution and wide-band measurement system noise level below about 50mVrms [2].

- High-frequency channel – a 110nF bottom-end capacitor with 220Ω shunt resistor provided a ratio of 100:1 at high frequencies and a high-pass characteristic with a corner frequency of 10kHz. The output signal was processed by a Frequency Devices active 4-pole Butterworth filter with a 10kHz corner frequency to further eliminate 50Hz signals and low order harmonics. It was then sampled at 2MS/s with suitable anti-alias filters. This provided wide-band (10kHz-1MHz) records of high-frequency voltage components up to 1.4Vrms with 14-bit (0.24mV) resolution and wide-band measurement system noise below 20mVrms. The narrow-band baseline noise levels in computed high-frequency voltage spectra were observed to be around 50µV [2].

The image below (Fig.23) shows the low-frequency voltage and current waveforms of the test number 646 as an example (high-frequency waveforms for voltage and current are also available). The test relates to a phase-ground fault via the species of tree "Pinus Radiata" (Appendix 1), the duration time of the test is 64563ms; Ilimit = 2A; a moisture content of 53.0%wt.; conductivity of the sample 96µS/cm and it did not result in a fire.
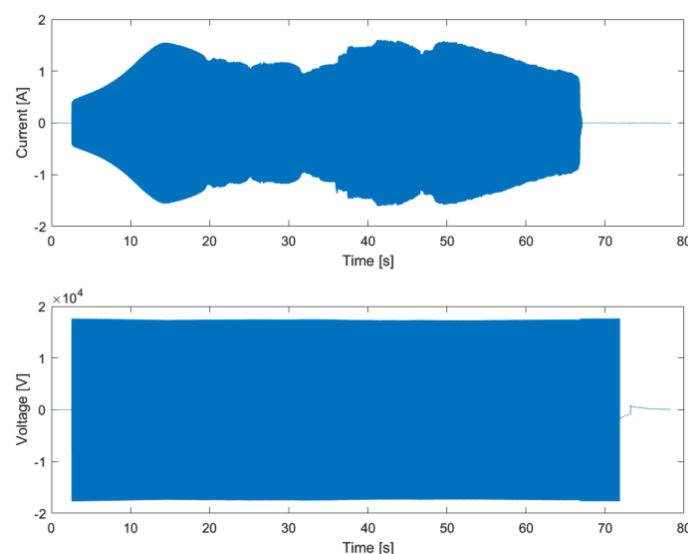

Fig.23 Low frequency voltage and current test n°646

4.3 "Branch on wire" characteristic

Having seen the data structure and an example of the voltage and current waveforms, it is necessary to explain the phases of a powerline "branch on wire" vegetation fault which will explain the development of the fault current value. It is possible to divide this phenomenon into four different components, the development of conductor-vegetation contact; the expulsion of moisture; progressive charring extending from the thinner end; flashover.

Fig.24 branch on wire test

Phase 1: "Conductor-vegetation contact"
When the vegetation sample comes into contact with an energized conductor, for the experiment under analysis 22kV, sparks can be seen at the contact point(s). Sometimes the sparks are visible for a short distance (typically not more than 100mm) as they are usually carried away by the wind. After an interval of time that can last from a few seconds to tens of seconds, each contact point is progressively wrapped in a ball of plasma or flame that extend almost right around the branch above the contact point. Since plasma and flame are conductors, the current has a larger area available to enter the branch. As the effective contact area increases, the fault current increases until it reaches the first maximum value. The current

usually takes 10 to 15 seconds to reach the maximum point, but in extreme cases it could take more than a minute [2]. The Fig.25 below considers test 407, shows the increase of current (low frequencies channel) in the "conductor-vegetation contact" step.

Test n°407: A.Pycnantha "Golden Wattle"; duration 31645 ms; Ilimit = 2 A; a moisture content of 32%wt.; conductivity of the sample 236µS/cm and it did not result in a fire.
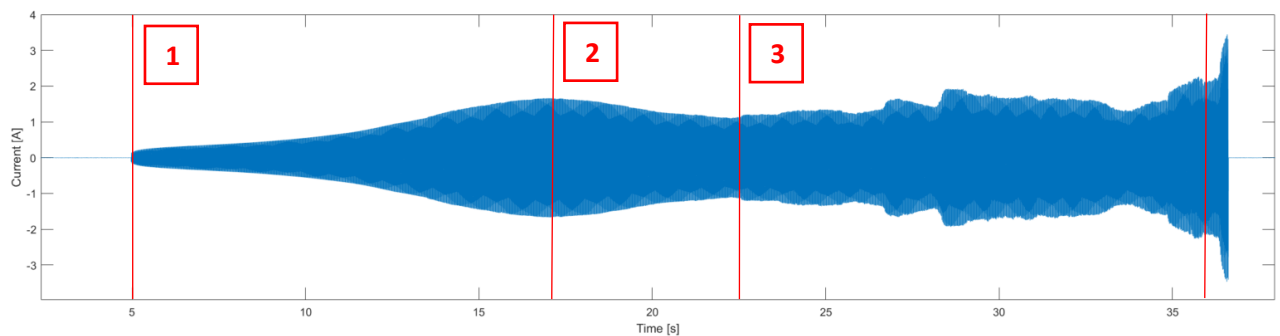

Fig.25 Stages of current evolution Test n°407

Phase 2: "Expulsion of moisture"

After the first current peak, the most relevant phenomenon of the majority of tests is the expulsion of moisture from the sample in the form of water and steam, often accompanied by a very loud continuous squealing or whistling noise.

During the second phase the current stabilizes and remains almost constant and even show a substantial decline from its first peak before the star of the next phase (Fig.25) [2].

Phase 3: "Progressive charring extending from thinner end"

At the end of the moisture expulsion process, flame slowly spread along the branch from one or both ends. In the Fig.25 is illustrated the progression of Phase 3 and the associated current waveform in a phase-to-earth test (test n°407) of a "Golden Wattle" sample.

Presumably the flame starts from the thinner end towards the thicker because of factors such as the higher resistance of thinner sections

generated more heath from the same current; lower mass of thinner sections meant a higher temperature increase for the same amount of heath.

Phase 3 is associated with progressive charring of the bark. In many tests this charring did not seem to penetrate the core wood of the branch, but often small branches and twigs were completely burnt out. As this phase progresses, small pieces of burning bark and twigs can fall into the ground creating potential fire risk.

As more and more the branch became covered in flame, intermittent arcs would sometimes appear in the flame. These would short-circuit the burning section of branch causing large fluctuations in the current [2].

Phase 4: "Flashover"

When the flame extends from conductor to conductor in an unbroken path, flashover occurs and the current go to the maximum value allowed by the 200Ω resistors in each phase of the supply – 45A in the phase-to-phase test and 65A in phase-to-earth tests.

In the majority of tests there was applied a current limit, usually 0.5, 1, 2, 4A. Such tests often ended during phase 1 or 3, flashover was only considered in the tests performed to study the entire fault development process [2].

4.4 Visualization of experimental results

Before proceeding with the modelling process, it is essential to analyse the characteristics of the fault current resulting from the contact of a shaft with a phase conductor.

The fundamental characteristics of a THIF do not differ much from those of other high-impedance faults. The amplitude of the current is very low, less than 2A in the totality of the tests in (Appendix 1), which has already mentioned is a major issue because it can easily be confused with other

phenomena such as energizing and de-energizing of loads and capacitor bank switching [22]. Following the contact of the branch with the conductor, the amplitude begins to rise (phase 1), reaches its maximum and begins to fall as moisture is being expelled and this corresponds to an increase in impedance (phase 2). Subsequently, we enter the carbonization phase, and this can result in an arc, so there is the possibility of a THIF fault evolving into an arc-ground fault. In this paper, we will concentrate on modelling the phases prior to carbonization for several reasons:

- Firstly, these two phases recur with similar trends for almost all the tests analyzed and it is therefore possible to compare them with each other in order to derive some kind of relationship and identify the parameters that most influence the evolution of the phenomenon, the only similarity between the various tests that is found in the later phases is that of phase 4, meaning the important increase in current amplitude following flashover (Fig.26).
- It is desirable that a fault detection methodology be developed to identify the presence of a THIF in the early stages, because at the end of the moisture expulsion phase, the charring phase begins which carries a high risk of fire ignition.
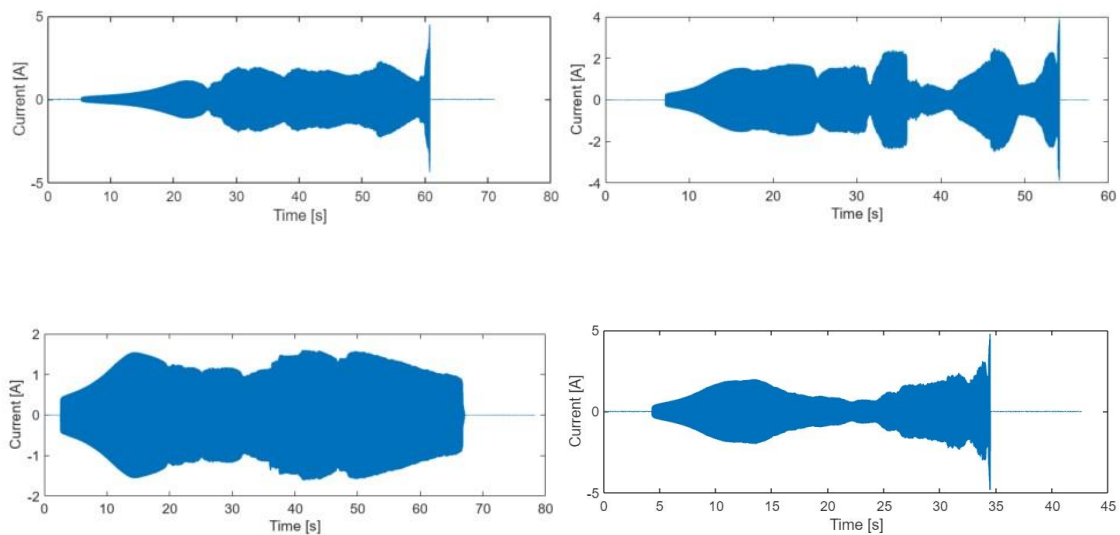

Fig.26 Similarity between various tests phase 1 and phase 2

As explained above, reference will be made to the data resulting from the project "Vegetation Conduction Ignition Testing" (Appendix 1). Following the

2009 Black Saturday bushfires in Australia, a series of devastating bushfires with some ignited by electrical faults, the commission responsible for investigating the fires recommended research in the field, this resulting in a $750 million founding [25]. One of the projects was the investigation on electrical faults caused by the contact of trees with power distribution lines.

In the tests around twenty species were tested in slightly less than one thousand experiments on a real 22 [kV] feeder to study fault behavior. The data was made public by the Victorian Government and contains video, logs, signal recording graphs and analysis of the tests.

Data are available in the format "H5DF" (Hierarchical data format) accompanied by a visualizer scrip in MATLAB. In total there are 53 background and 983 fault tests [25]. within the final report, a summary table can be found where the main characteristics such as species under analysis, test duration, sample diameter, configuration (earth phase or two-phase) are indicated (Table 2) [2].

Let us now take test number 646 with the characteristics below as an example, enter this number into the MATLAB visualization script (Appendix 1) and see what information is displayed.

Once executed, the script provides 5 graphs, low-frequency voltage and current; high-frequency voltage and current; and trigger signal (Fig.27).
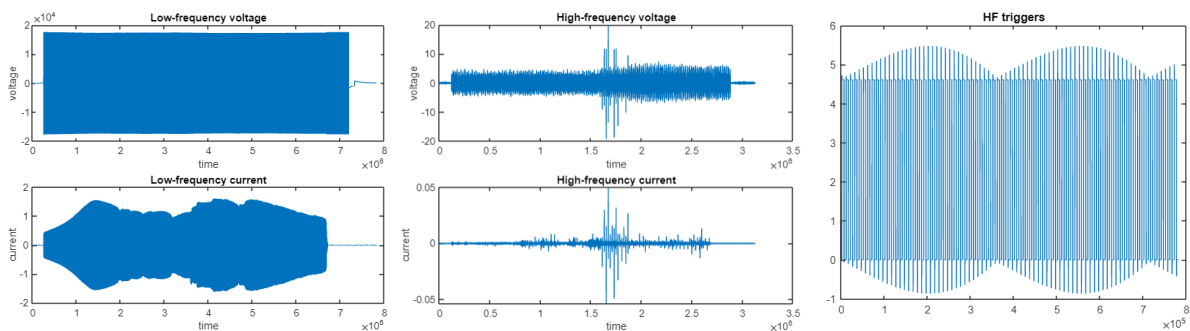

Fig.27 Test n°646 visualization

Table 2: test 646 characteristics

Test	Type	Iinit[A]	Duration [ms]	Ilimit[A]	Flashover	species	Moisture [%wt.]	Conductivity [µS/cm]	Diameter [mm]	Fire result
646	Ph-e	0.289	63563	2	N	Pinus Radiata	53	96	25	0

Table 3: test characteristics [2]

Date (2015)	Test	Type	I_init (amps)	Duration (ms)	I_limit (amps)	Flashover	Species	Conditioning time (hours)	Moisture (%wt.)	Conductivity (µS/cm)	Diameter (mm)	Fire result (1=fire)
	75	ph-e	0.22	4000	0.5	N	Eu. Viminalis	16	39.7	249		0
	76	ph-e	0.24	2000	0.5	N	A. Pycnantha	16	28.7	142	45	0
	77	ph-e	0.13	6000	0.5	N	A. Pycnantha	16	27.9	212	30	0
	78	ph-e	0.19	5000	0.5	N	Eu. Baxteri	16	44.0	193	40	0
	79	ph-e	0.1	72000	0.5	N	Eu. Baxteri	16	35.5	210	30	1
	80	Bush	0.12	14000	0.5	N	K. Ericoides	16	33.7	153	35	0
	81	ph-e	0.04	31000	0.5	N	Eu. Viminalis	16	30.5	342	20	0
	82	ph-e	0.07	22000	0.5	N	Eu. Baxteri	16	38.6	140	35	0
	83	Bush	0.07	22000	0.5	N	K. Ericoides	16	33.6	178	20	1
	84	Bush	0.05	36000	0.5	N	K. Ericoides	16	30.4	276		1
	86	ph-e	0.13	60000	1	N	Eu. Viminalis	16	32.5	230	30	1
	87	ph-e	0.04	31000	1	N	A. Melanoxylon	178	32.1	373	30	1
12 Feb	88	ph-e	0.07	58000	1	N	Eu. Baxteri	178	36.5	203	30	1
	89	Bush	0.14	18000	1	N	K. Ericoides	178	36.2	181	30	1
	90	ph-e	0.09	72000	1	Y	Eu. Baxteri	178	35.8	163	25	1
	91	ph-e	0.19	7000	1	N	A. Pycnantha	178				0
	92	ph-e	0.06	58000	1	N	Eu. Baxteri	178	34.7	189	30	1
	93	Bush	0.16	16000	1	N	K. Ericoides	178	33.7	160	35	0
	94	ph-e	0.08	23000	1	N	Eu. Baxteri	178				0
	95	ph-e	0.08	45000	1	N	Eu. Viminalis	178	38.4	195	30	0
	96	ph-e	0.04	84000	2	N	Eu. Baxteri	178	31.8	173	35	1
	97	ph-e	0.04	49000	2	Y	A. Pycnantha	178	29.8	275	30	1
	98	Bush	0.11	60000	2	Y	K. Ericoides	178	35.0	170	25	1
	99	ph-e	0.1	69000	2	Y	Eu. Baxteri	178				1
	100	ph-e	0.08	67000	2	Y	Eu. Baxteri	178	40.4	210	35	1
	101	Bush	0.16	37000	2	N	K. Ericoides	178	41.8	234	40	1
	102	ph-e	0.08	67000	2	Y	Eu. Viminalis	178	33.1	215	25	1
	103	ph-e	0.12	76000	2	N	Eu. Baxteri	178	41.7	185	45	1
	106	ph-e	0.08	13000	0.5	N	C. Glaucophyllus	24	37.1	229		0
13 Feb	107	ph-e	0.07	25000	0.5	N	C. Glaucophyllus	24	36.8	225		0
	108	ph-e	0.15	17000	1	N	C. Glaucophyllus	24	34.2	167		0

The graphs of Fig.27 were left in their original form without modification to show the visualization in the original script, after which it was changed to obtain a graph as in Fig.26/27. In the MATLAB workspace following execution, the variables of interest, the low-frequency voltage and current, are saved in vector format 7830008x1 double.

5 Tree High Impedance Fault Modelling

The field experiment described in the previous chapter, with its multitude of voltage and current records relating to single-phase and two-phase faults caused by the contact of vegetation with phase conductors, served as a basis for the following tasks:

- o Develop an understanding of the phenomenon: through the analysis of various tests, it was possible to comprehend the parameters that influence the dynamics of this type of fault and the ways in which they influence it.
- o low-frequency voltage and current measurements are used for a regression and interpolation process: This allows fault currents to be estimated for unknown moisture values (as moisture is one of those key parameters influencing the dynamics of a THIF).

- o Finally, the available data is used to validate the model.

The creation of a reliable model describing the dynamics of a tree-high impedance fault is desirable because it would allow a large number of datasets to be derived without recourse to expensive and time-consuming field tests such as the one described in Chapter 4. The model could also be applied to electrical network simulation software, obtaining fundamental data for the training phase of AI-based fault detection methodologies.

A procedure similar to that set out in section 3.5 (MODEL 3 "Nam et al. Described in 3.5 [23] [20]) was followed for the creation of the model. The image below summarises the procedure followed, which will be explained in detail later in this chapter.
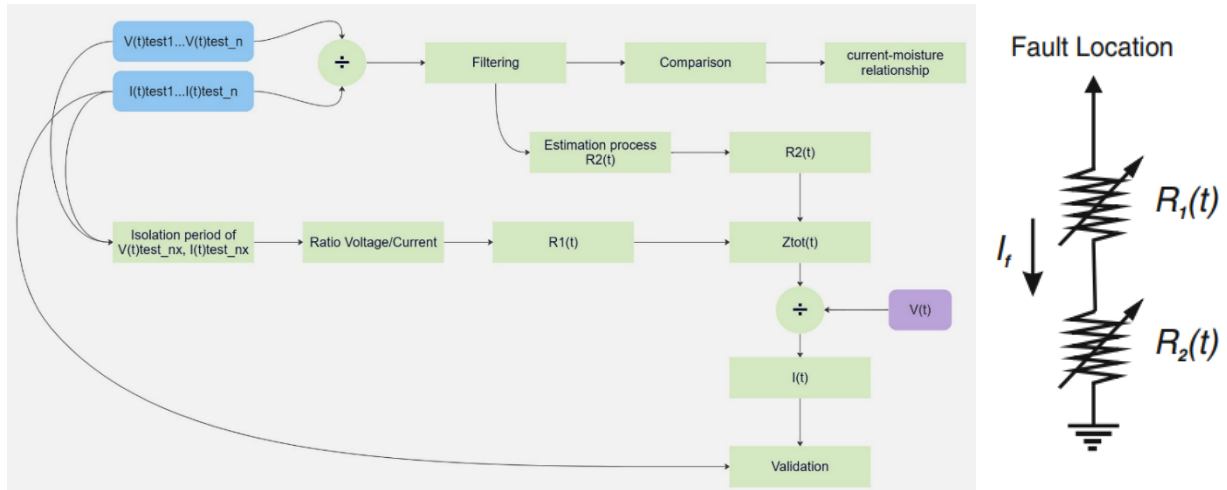
Fig.28 summary modelling process

5.1 Modelling R1(t) component

To create a model describing the behavior of a THIF, a methodology similar to Nam et al.'s 'MODEL 3' described in 3.5 [23] [20] is used. It is therefore required to derive two variable resistors R1(t) and R2(t). The first will model the asymmetry and non-linearity of the current, while R2(t) will simulate the increase in current of the contact formation and moisture expulsion phases. In this thesis, however, it was chosen to derive a complete impedance without splitting into two components.

To begin the procedure for evaluating R1(t), a MATLAB script was created which receives as input the vectors with the voltage and current data provided by the visualization script described above. At this point, it was decided to perform a linear interpolation of the input data to make it possible to find the zeros and thus isolate the various periods. Subsequently, the portion of the graph between the beginning of the

conductor-branch contact formation and the end of the moisture expulsion phase was isolated (Fig.29).
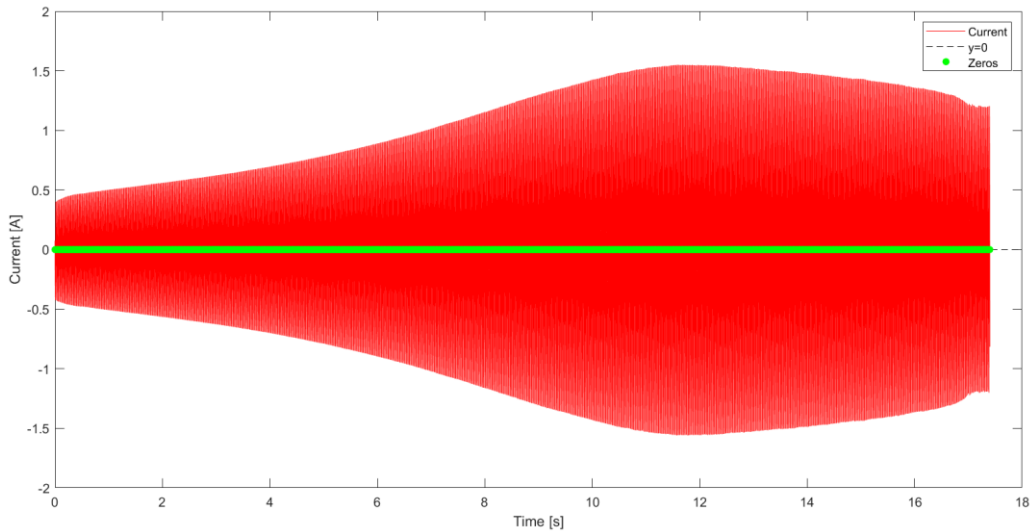
Fig.29 Current interpolation test n°646 - phase one and two isolation

Using interpolation, it was possible to isolate the individual cycles of current and voltage, and thus to evaluate the non-linearity by creating the V-I characteristic. A cycle in the beginning of phase two, at 11s from the start of contact formation, and an internal cycle in phase three of carbonization were therefore taken. The V-I characteristics were then plotted by placing the voltage values as a function of current (Fig.30).
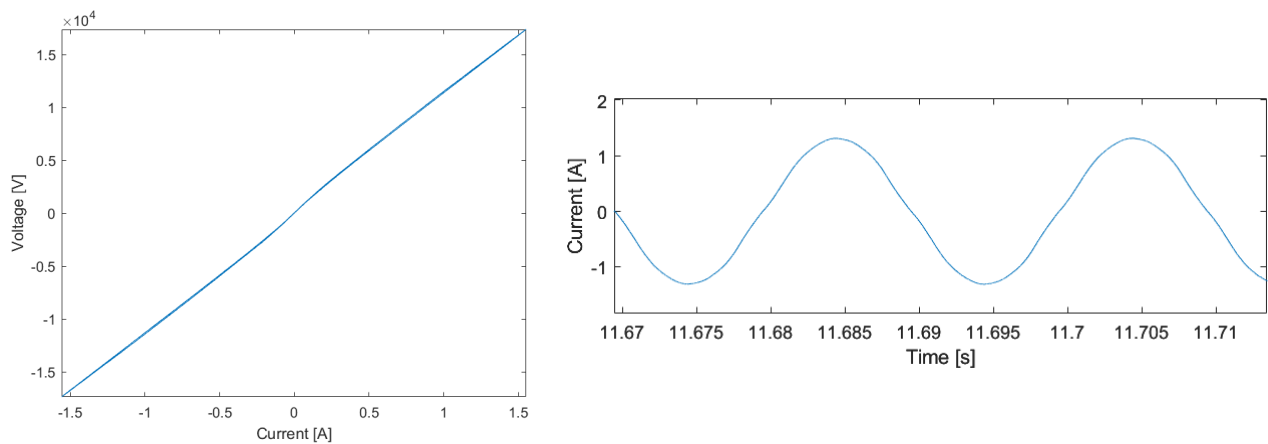
Fig.30a V-I characteristic in the end of phase 1
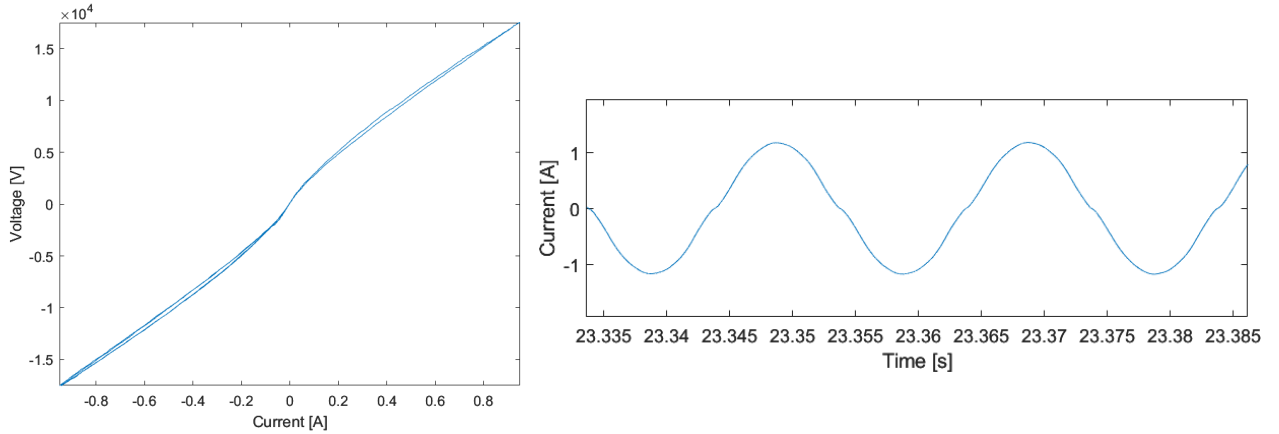
Fig.30b V-I characteristic phase 2

It was found that the non-linearity and discontinuity in the proximity of the zero crossings increases significantly in the carbonization phase compared to the first two phases. This behavior was found for the entirety of the tests analyzed, in particular, the species Pinus Radiata was taken, and the V-I characteristics of the test beacons were confronted; end-of-phase one cycles (near the point of maximum current) were used for all the tests compared and no significant differences were seen.

Following this constatation, it was decided to use the cycle in Fig.31 to derive a basic impedance describing the non-linearity and discontinuity characteristics of the current for the model. This process of constructing R1(t) is therefore similar to that proposed in [23]. The expression of R1(t) was derived as a simple ratio between voltage and current:

(5) $R1(t) = \dfrac{V(t)}{I(t)}$

V(t) and I(t) are both two cycles chosen near the end of the contact formation phase Fig.31. 154 points of both voltage and current were used in order not to overload the memory in the final stage of the model where the complete impedance will be constructed.
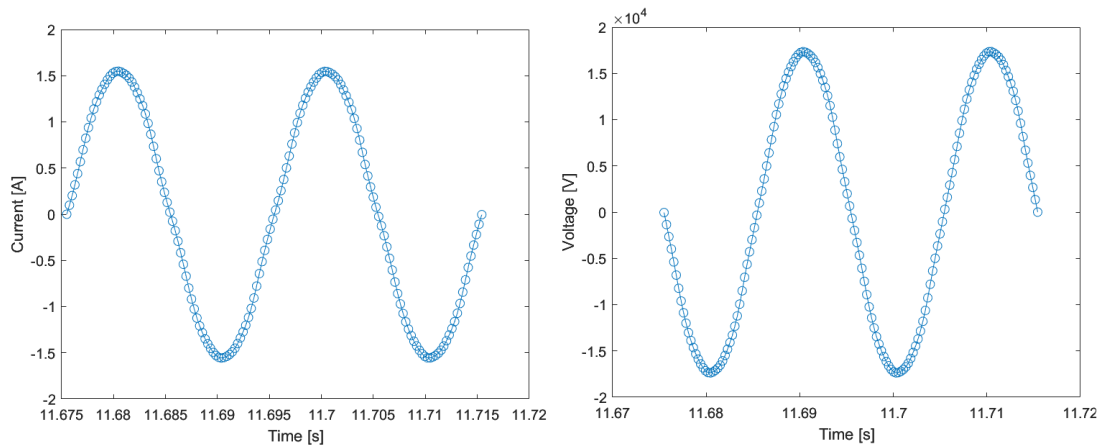
Fig.31 Current and voltage used for R1(t)

Once the ratio between these two quantities was obtained, the impedance trend had to be eliminated, some points, all in the proximity of the zero crossing, which appeared as very high peaks. The mean value between the maximum and minimum was also calculated, which will be used later to derive the complete impedance curve (Fig.32).

It is possible to see the non-symmetry of the current, the current value in the positive half-period is always greater than in the negative half-period, consequently the impedance has inverse behavior with a greater value in the negative half-period.
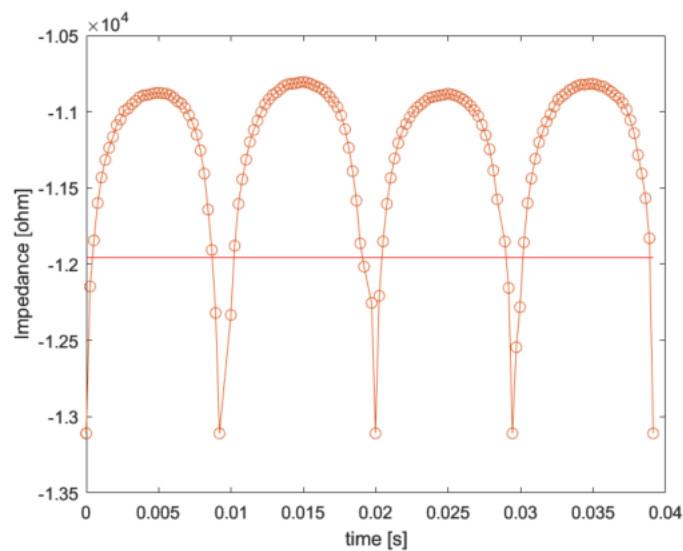
Fig.32 R1(t)

5.2 Impedance Filtering

With the construction of R1(t), it was intended to describe the behavior of THIF in the order of magnitude of "ms" in order to model the current waveform. It is now necessary to move into the order of magnitude of seconds to try to model the increase in the value of the current.

As a first step, different tests belonging to the same species (Pinus Radiata) were taken in order to compare the current trend in the first two phases. The intention was to compare fault impedances rather than using currents, and therefore the relationship between the vectors (7830008x1 double) of voltage and current provided by the visualization script was performed. The result Fig.33a was then plotted, this was initially surprising as it did not seem to follow the expected pattern considering the shape of the current. Once zoomed in to the "ms" order of magnitude, it was seen that the signal exhibits very high peaks at each zero crossing of current and voltage, then at a frequency of 100Hz Fig.33b.
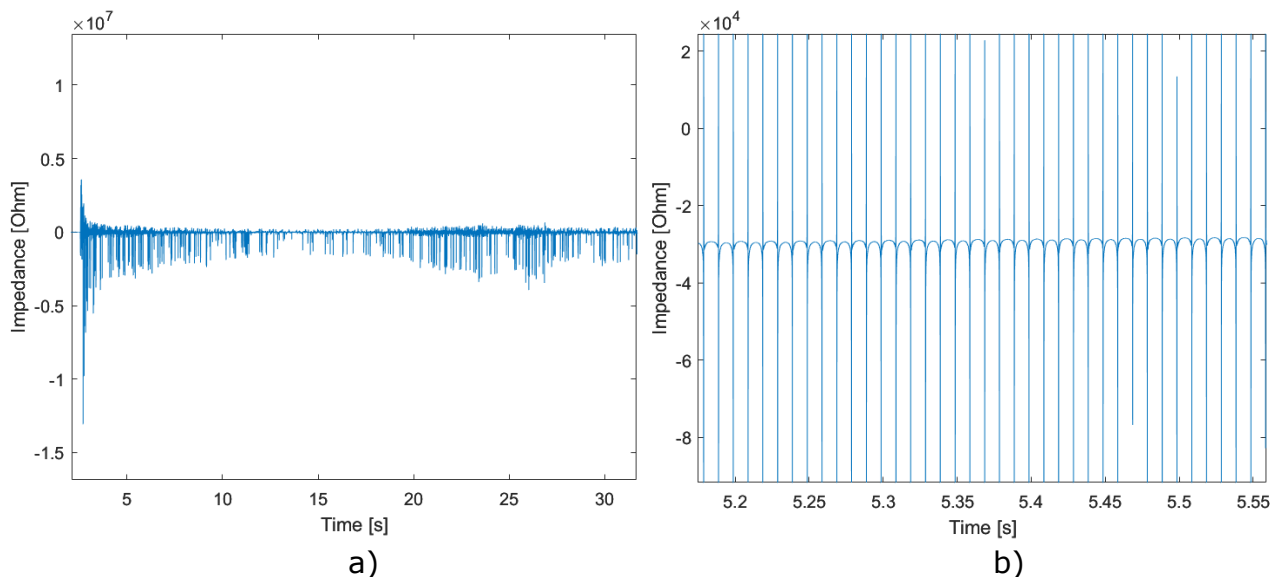
Fig.33 V(t) / I(t) ratio

A low-pass filter with a cutoff frequency of 92 [Hz] was initially used to clean the signal, resulting in the graph in Fig.34, where the expected trend is beginning to be visible. What is desired, however, is a trend without oscillations to be used as a drive to increase the value of R1(t). The output

47

of the first filter was then re-filtered with an additional low-pass at a cut-off frequency of 100Hz to obtain the desired shape.
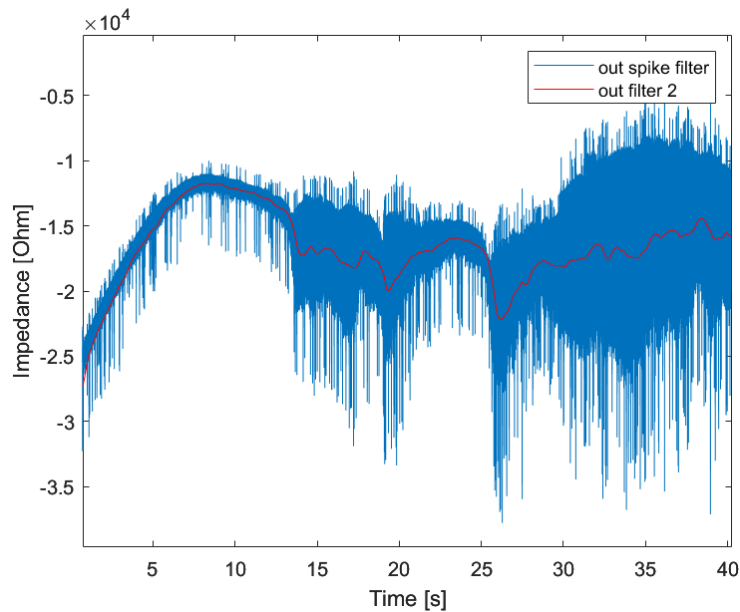
Fig.34 Impedance filtering

It can be seen that for this test, the impedance goes from an initial maximum value of about 25000Ω to a minimum value of 11750Ω reached in about 9s. Subsequently, the impedance rises again following the start of moisture expulsion from the sample, at 14s this phase also ends, and the carbonization phase begins.

5.3 Correlation between impedance and moisture content

In this phase, an analysis of the impedance trend is carried out by comparing various tests belonging to the same species. The aim is to understand which parameters most influence the impedance value and the duration time of the various failure phases.

For this purpose, tests with characteristics summarized in the table below were taken, filtering was carried out for each of them as set out in 4.4, and the impedance trends were compared in a single graph Fig.35.

Table 4: Impedance comparison

Test	Type	I init [A]	Duration [ms]	I limit [A]	Flashover	Species	Moisture [%Wt]	Conductivity [μS/cm]	(1=fire)
230	ph-e	0.449	147211	2	N	Pinus Radiata	56	150	1
238	ph-e	0.081	188477	2	Y	Pinus Radiata	41.8	124	0
260	ph-e	0.06	109311	2	Y	Pinus Radiata	33	123	1
269	ph-e	0.172	144704	2	N	Pinus Radiata	43.2	107	1
270	ph-e	0.208	65706	2	N	Pinus Radiata	50.7	155	0
643	ph-e	0.185	130610	2	N	Pinus Radiata	45.7	85	0
646	ph-e	0.289	64563	2	N	Pinus Radiata	53	96	0
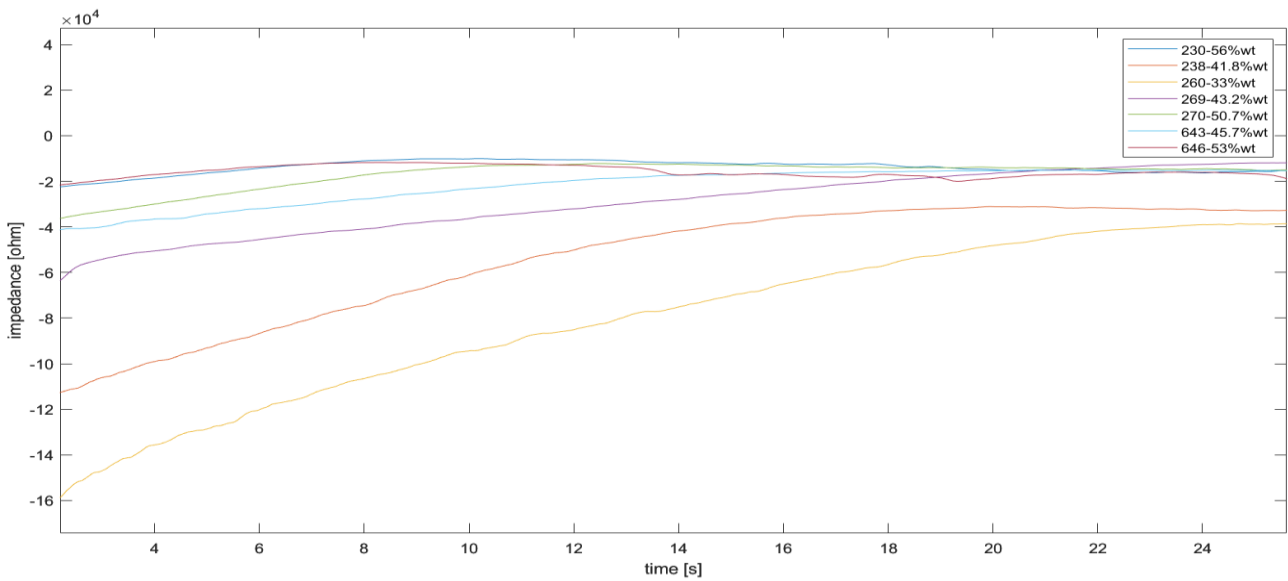
Fig.35 Impedance comparison

From Fig.35, it was realized that the initial impedance value is closely related to the moisture value of the sample. As the moisture value decreases, the initial impedance increases in absolute value.

The moisture content of the sample greatly influences the value of the impedance, take as an example the tests 230 and 260 with moisture 56%wt and 33%wt respectively, we have an initial impedance of approximately 20000Ω for the higher moisture test and 160000Ω for the lower moisture test, thus a considerable difference of 140000Ω.

This relationship between moisture and impedance has been found for all tree species, not shown here. The reason for this behavior is easily understood as the water molecules within the sample provide the current with a lower resistance path [2].

A model describing a high-impedance fault cannot fail to take this type of behavior into account; environmental conditions can lead to considerable

differences in the current value. In particular, during heavily rainy seasons a THIF fault would manifest itself with a higher current value than a fault of the same type during a dry season. This is certainly not a positive factor, as the fire hazard is certainly higher during summer periods that coincide with low humidity values and thus a lower and more hardly detectable fault current value.

The resistance R2(t) of the model, which decreases the value of R1(t), must therefore depend on the humidity value. The objective is to be able to obtain the impedance value for any value of humidity, which is considered to be very useful since, as mentioned before, it will be possible to relate the current value to the environmental conditions.

Once the model of R2(t) has been obtained, it will be possible to verify it with the experimental data. R2(t) will be derived with a moisture value available in the test set (not present in Fig.35) and compared with the actual test data.

5.4 Impedance estimation for unknown moisture content – R2(t)

Before proceeding to the impedance estimation step for humidity values other than those provided by the experimental data[16], it was decided to perform a polynomial interpolation for each impedance shown in Fig.34, in order to obtain a function, which is much easier to use in subsequent mathematical processes than a data vector.

Polynomial interpolation is a mathematical method used to approximate an unknown function from a set of known points. This method involves constructing a polynomial of degree 'n', which will be a polynomial equation in the general form:

(5.1) $\qquad P(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0.$

The a_n, a_{n-1}, ..., a_1, a_0 are the coefficients of the polynomial that are determined by solving a system of linear equations. These equations are obtained by imposing that the polynomial passes through each of the known points, i.e. pairs of coordinates (x, y). By solving the system, the coefficients can be calculated uniquely. Once the coefficients have been obtained, the polynomial of degree "n" can be used to approximate the value of the function at intermediate points between the known ones, thus providing an approximate estimate of the general form of the function. For the data subjected to polynomial interpolation, it was observed that a polynomial of degree 6 best approximates the data trend. For this mathematical procedure, MATLAB's "polyfit" function was used, then the curves approximating the impedances in the first 25s of the THIF were drawn using the "polyval" function[22]. The following table shows the polynomials obtained. These polynomials were plotted with respect to the time axis, which was equalized for all tests, from zero to twenty-four seconds with a 0.1s step (Fig.36).

Table 5: Basic impedances for estimation

Test	a0	a1	a2	a3	a4	a5	a6
230	-2,5601E+04	979,2593	698,5181	-130,1068	9,1955	-0,2999	0,0038
238	-1,2340E+05	11487	-1634,9	252,7368	-19,4019	0,6747	-0,088
260	1,8518E+05	28010	-51467	610,3401	-38,8754	1,2501	-0,016
269	-6,1179E+04	7824,7	-1536,9	199,4576	-13,2448	0,4339	-0,0056
270	-3,8979E+04	2024,6	625,3849	-96,8171	4,9559	-0,0993	5,2299E-04
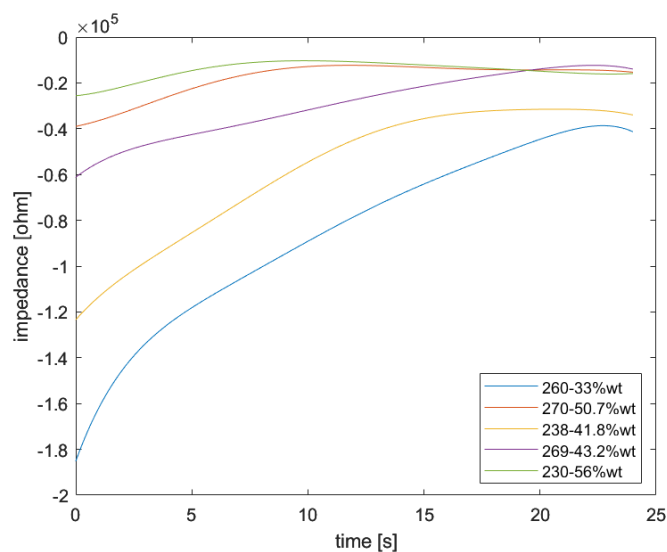
Fig.36 Polynomial interpolation of impedances

To obtain the trend of an impedance with an unknown moisture value, the following procedure was performed: for each value (tx) of the time axis (tx = 0: 0.1:24 [s]), the function passing through the points $(Z_{260}(tx);Wt\%_{260})$, $(Z_{270}(tx);Wt\%_{270})$, $(Z_{238}(tx);Wt\%_{238})$, $(Z_{269}(tx);Wt\%_{269})$, $(Z_{230}(tx);Wt\%_{230})$ was constructed. Through interpolation, the point $(Z_{unknown}(tx);Wt\%_{desired})$ is found. This procedure was iterated until the 1x241 vector containing the values of the unknown impedance was obtained. As the number of tests used for comparison increases, the precision of the estimate increases; only five base curves are given here to simplify the exposition of the procedure.

Some estimated curves for different humidity values are now given as examples Fig.37, they will be used as the basis for R2(t).
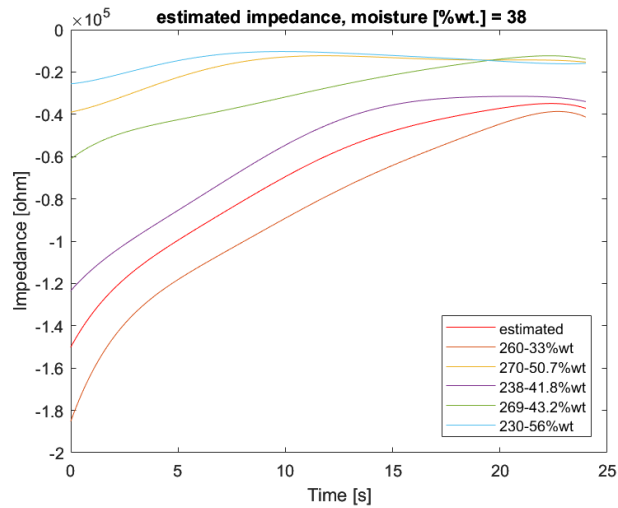
Fig.37a Estimated impedance, %wt.=38
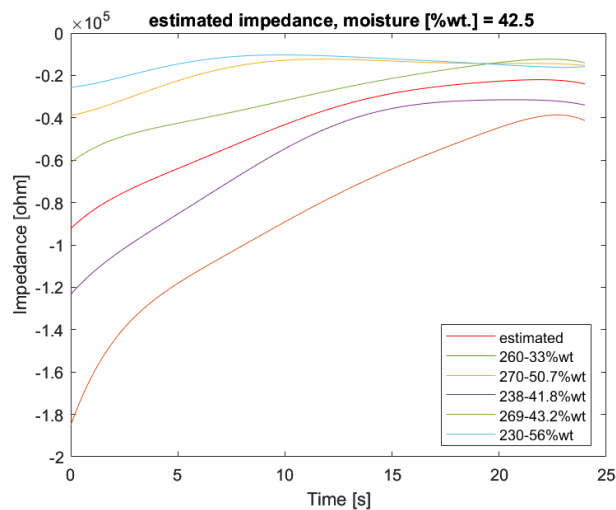
Fig.37b Estimated impedance, %wt.=42.5
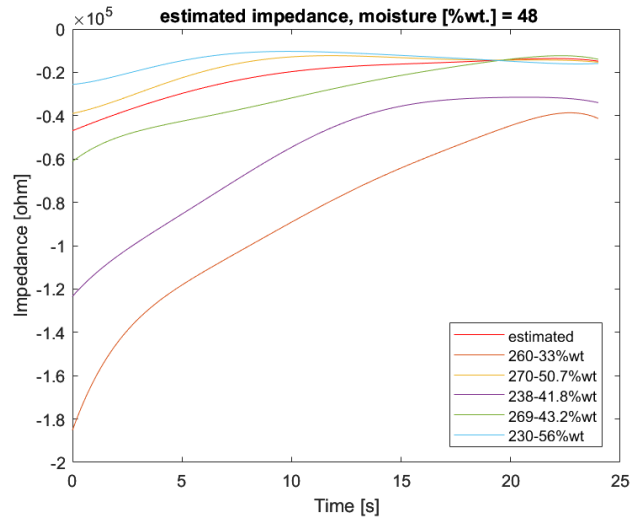
Fig.37c Estimated impedance, %wt.=48

5.5 Current estimation procedure for unknown moisture content

The last step in creating the THIF fault model in the contact formation and moisture expulsion phase is to combine the two resistances R1(t) and R2(t). I the current will be obtained by means of the ratio of voltage to impedance Z(t)=R1(t)+R2(t); initially, the current was obtained for one of the tests used to estimate R2(t), then using one of the sixth-degree polynomials given in Table 5 as the impedance R2(t), this was done in order to compare the current output of the model with the original current and to evaluate the error introduced by the filtering and approximation of R1(t).

As a final step, a current will be estimated for a moisture value not present in Fig.35, but available in the test set[17] to verify the accuracy of the model.

We proceed to the construction of $Z_{final,238}(t)=R1_{238}(t)+R2_{238}(t)$ referring to test No.238, this to evaluate the error introduced by filtering steps and approximation R1(t). In Fig.38a-b the basic R1(t) and the sixth-degree polynomial approximating the impedance decrease are shown. The objective is to repeat the form of Fig.38a so that it follows the trend of Fig.38b. To do this, the mean value of R1(t) base will be used and first this will be set

equal to the starting point (t=0) of R2238(t); R1base(t) is then translated
to obtain $R1_{238}(t)$ (Fig.38c).
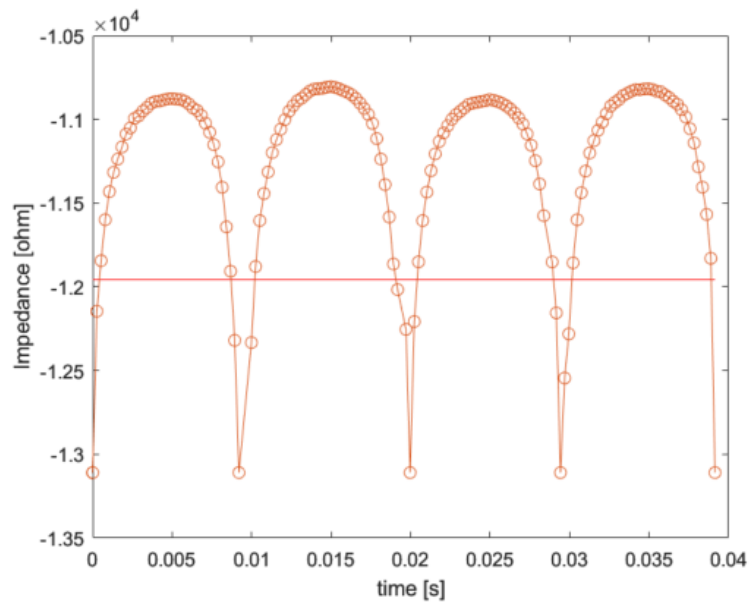
Fig.38a $R1_{base}(t)$
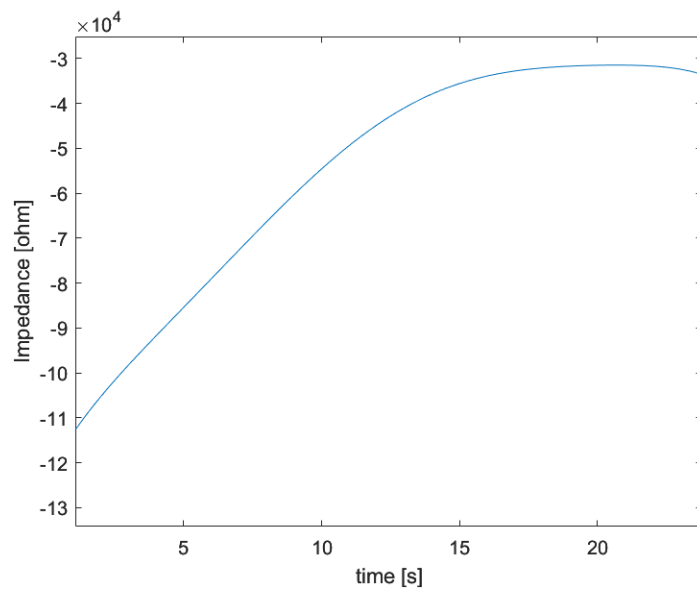
Fig.38b $R2_{238}(t)$
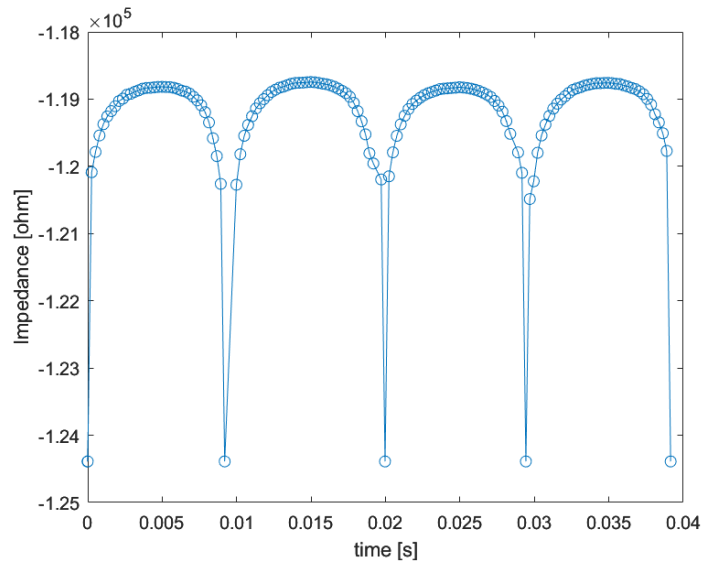
Fig.38c $R1_{238}(t)$

At this point $R1_{238}(t)$ is repeated every 0.4[s] imposing an increment from the mean value and evaluated on $R2_{238}(t)$. The number of increments will be 24[s] / 0.04[s] so a total of 600 increments. The final form of $Z_{238}(t)$ Fig.39 is thus obtained.
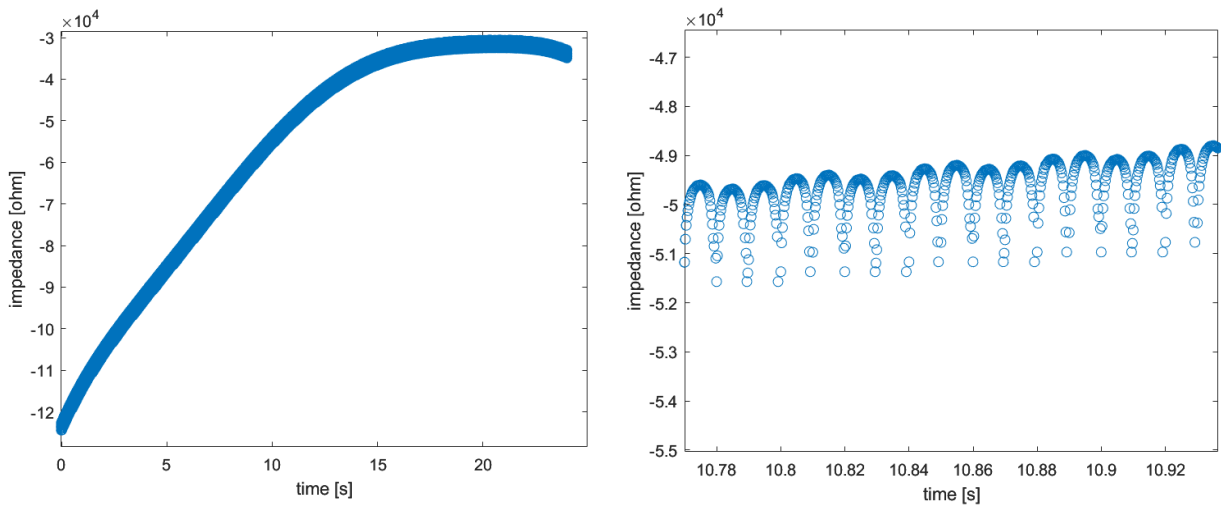
Fig.39 $Z_{238}(t)$

All that remains now is to derive the current by means of the voltage and impedance ratio. A sine wave with an amplitude equal to that used in the tests was created as the nominal voltage in [2].

$$(5.2) \qquad I_{238}(t) = \frac{\frac{22000}{\sqrt{2}} \sin(2\pi \, 50 \, t)}{Z_{238}(t)}$$
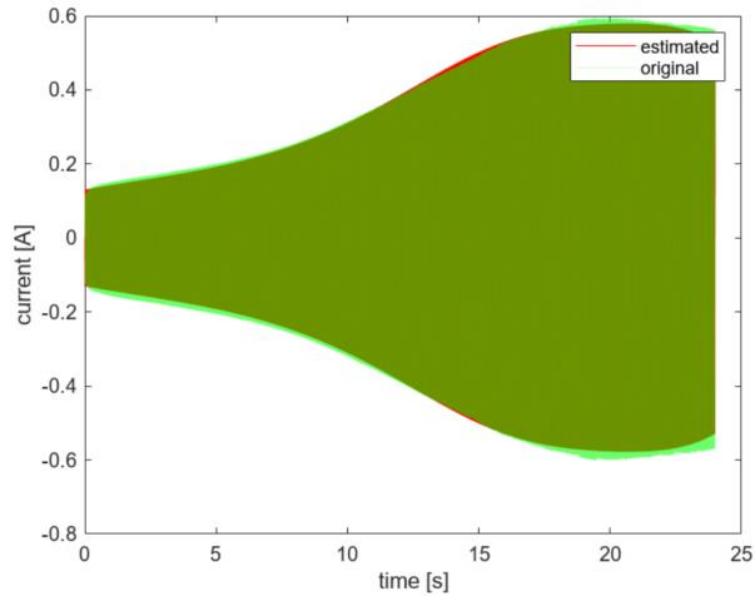
Fig.40 $I_{238}(t)$ original and estimated

Fig.40 compares the current waveform obtained with (5.2) with the original current. It can be seen that the error introduced by the approximations that were performed to derive $Z_{238}(t)$ is extremely small. Is therefore possible to proceed with the validation of the model.

6 Results and Discussion

6.1 Model validation

Once it has been established in 5.4 that the approximations and filters used in the development of the model do not introduce substantial errors, the next step is to validate the model.

We must then proceed to assess the accuracy of the current estimated for unknown moisture values. A moisture value will then be chosen which is available within the dataset but not used as a reference for estimating the full impedance (not present in table 5).

Validation 1

We begin to assess the accuracy of the model by assuming a moisture content of 53%wt. (moisture value not used in the estimation process but present in the data set in Appendix 1 for validation). The current is then constructed following the steps outlined in Chapter 5 and compared with the current available in test 646 (Appendix 1). The characteristics of the test are summarised in Table 6.

Fig.41 shows the trend of the estimated current (blue) and the trend of the measured real current (green). We can see how the model is consistent with the real course, but we notice some differences: In the first phase of contact formation, from time zero until the maximum is reached, it can be seen that the amplitude of the estimated current is greater than that of the real current; the maximum of the estimated current is slightly earlier than the real maximum; finally, in phase two, the phase of moisture expulsion, the amplitude of the real current is greater than that estimated. The graphs of R1(t) and R2(t) used to construct the current are also shown.

Table 6: Test 646 characteristics

Test	Type	I init [A]	Duration [ms]	I limit [A]	Flashover	Species	Moisture [%Wt]	Conductivity [μS/cm]	(1=fire)
646	ph-e	0.289	64563	2	N	Pinus Radiata	53	96	0
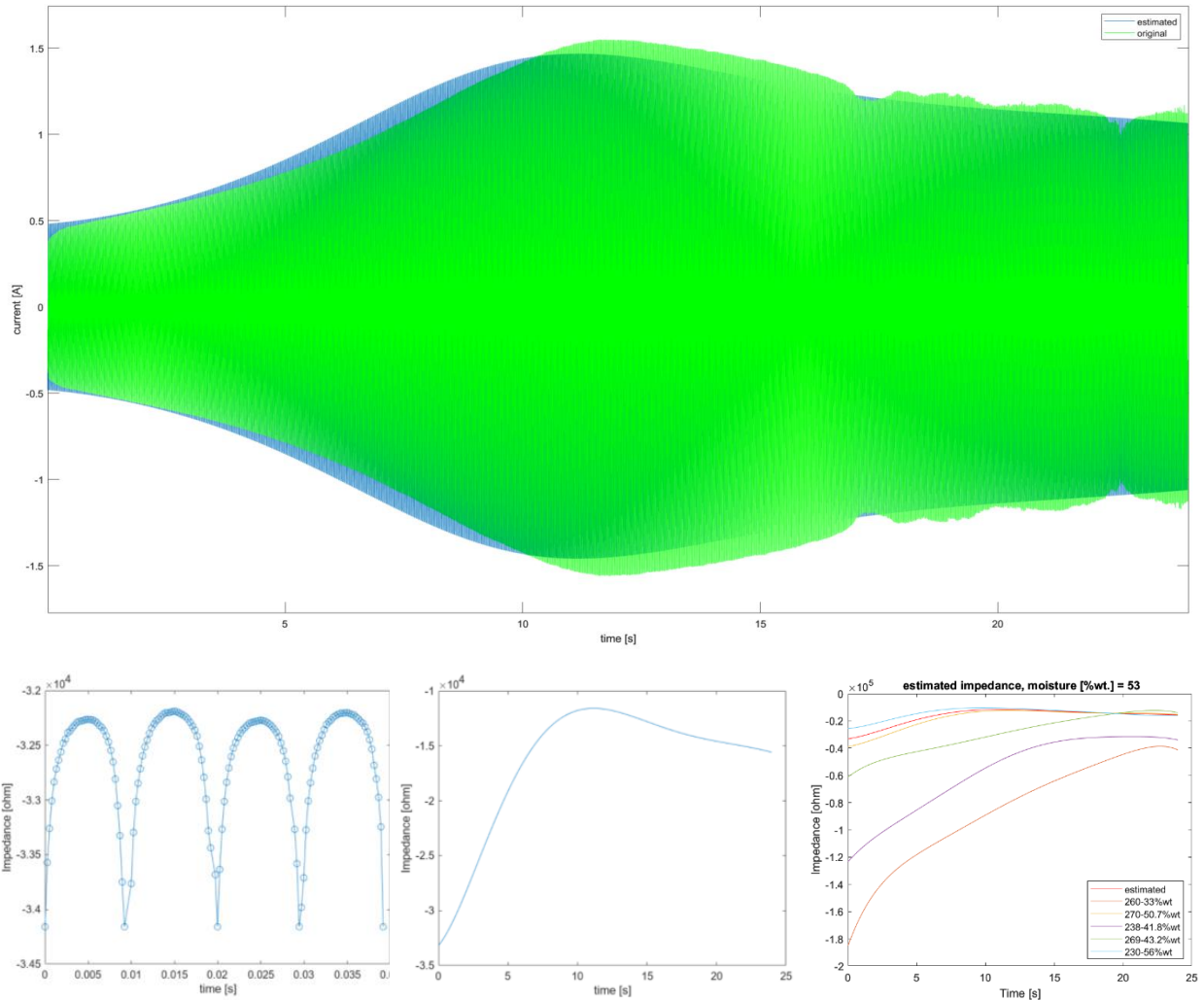
Fig.41 Validation 1

Validation 2

Consider now another validation of the model, let us assume a humidity value of 45.7%wt. (moisture value not used in the estimation process but present in the data set in Appendix 1 for validation). Let us construct the current and compare it with the actual current exactly as we did before. Table 7 summarizes the characteristics of the test.

Fig.42 shows the result, and it can be observed that the matching in this

case is significantly worse than in "validation 1". Estimated current and real current are very similar during the 15s, but the behavior thereafter is different: the estimated current reaches a maximum and then begins to decrease, while the measured current continues to increase and reaches a maximum after the 24s considered in the model.

Table 7: Test 643 characteristics

Test	Type	I init [A]	Duration [ms]	I limit [A]	Flashover	Species	Moisture [%Wt]	Conductivity [μS/cm]	(1=fire)
643	ph-e	0.185	130610	2	N	Pinus Radiata	45.7	85	0
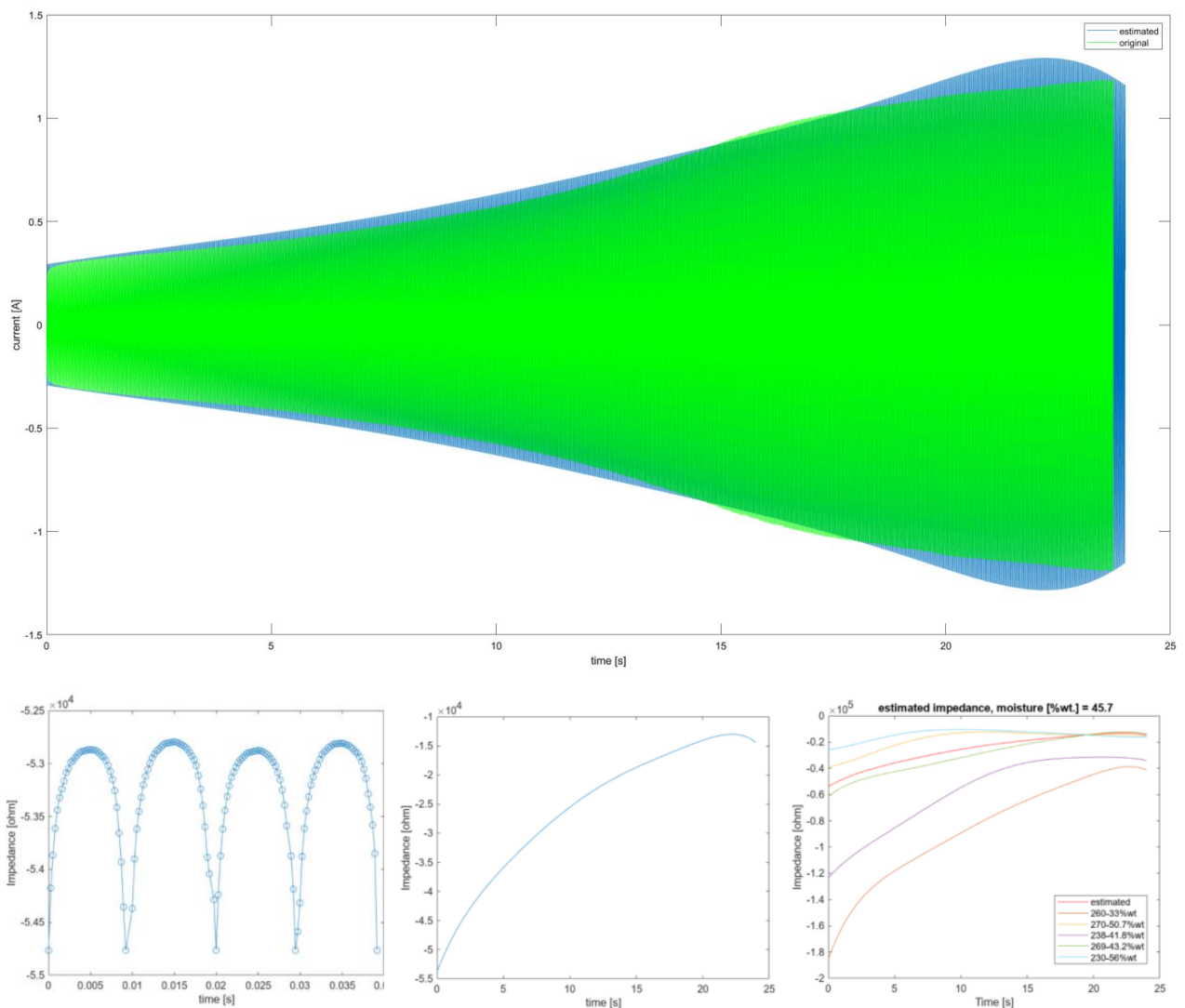
Fig.42 Validation 2

Validation 3

As a final validation, a different tree species to the type used in the estimation process is shown. The aim is to verify if the model can also describe vegetation types different from 'Pinus Radiata' with good approximation.

Assume therefore a moisture content of 51.6%wt. (moisture value not used in the estimation process but present in the data set in Appendix 1 for validation) and derive the current, it will then be compared with the real test of a tree "Eu. Baxteri" (test characteristics summarised in Table 8).

From Fig.43 it can be noted that the current obtained from the model is very similar to the real current but shows differences with similar characteristics to 'validation 1': in the phase of contact formation, the estimated current has a greater amplitude than the measured one; opposite characteristics in the subsequent phase of moisture expulsion.

Table 8: Test 392 characteristics

Test	Type	I init [A]	Duration [ms]	I limit [A]	Flashover	Species	Moisture [%Wt]	Conductivity [μS/cm]	(1=fire)
392	ph-e	0.118	85332	2	Y	Eu. Baxteri	51.6	204	1
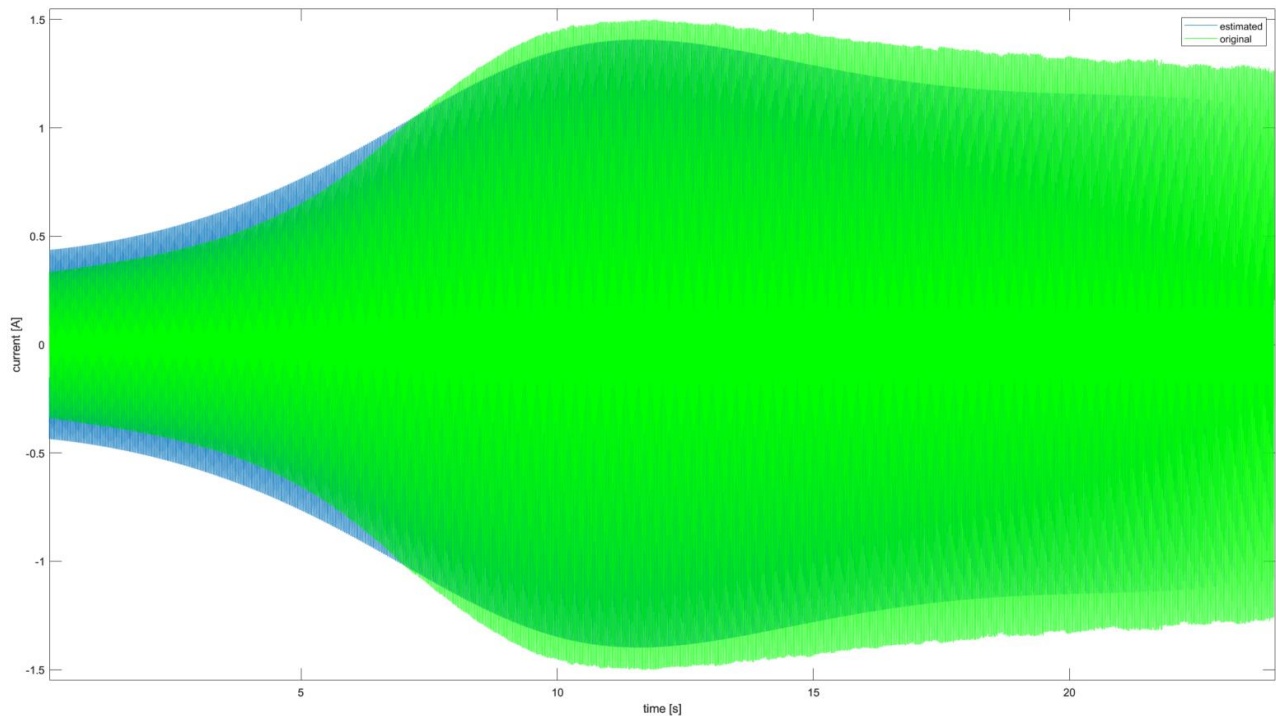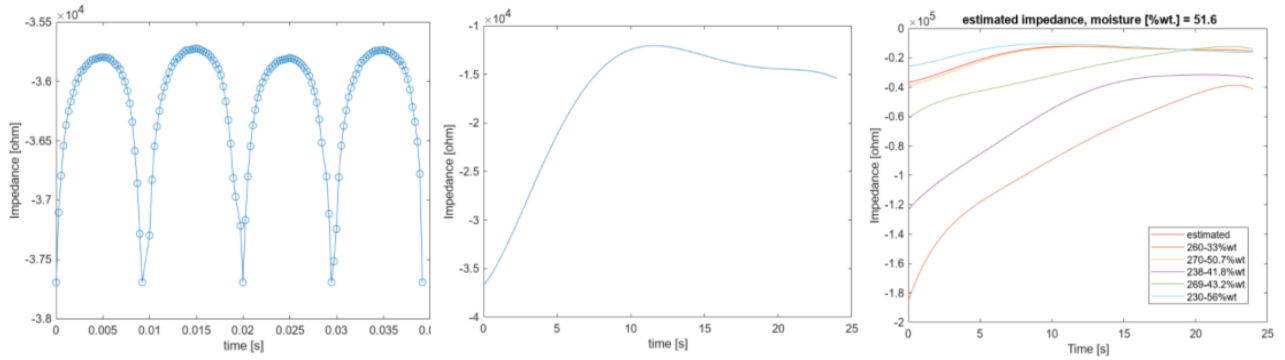
Fig.43 Validation 3

6.2 Discussion

In 6.1, some verifications of the model describing a high-impedance single-phase fault caused by vegetation contact with phase conductor were reported. It is now useful to discuss these results in order to define the limits of this model and show what can be done to increase the accuracy. It was found that the model has a high accuracy in the humidity range 46%wt.-56%wt., whereas for the moisture range 33%wt.-46%wt. the increase in the amplitude of the estimated current differs from that of the real current, in particular the instant of reaching the maximum at the end of phase 1 of contact formation often does not coincide. This poor accuracy can be attributed to the impedance estimation process R2(t).
In order to improve this impedance estimation process, it is necessary to introduce more base impedances by using a larger number of tests. This process is certainly possible given the large data set available (Appendix 1). The model following this process of introducing new base impedances will, however, remain limited in the moisture range of approximately 25%wt.-56%wt., the entirety of the almost 900 tests available, in fact, belong to this range and there is no certainty that an extrapolation process will allow a correct estimate outside this range.

Even using the entirety of the tests for the estimation process, it is considered extremely complex to derive a perfect coincidence between estimated and true current. In this model, in fact, some parameters are not considered which certainly influence the temporal evolution of the current:

the contact point between phase conductor and branch certainly has an influence on the current development [4]; the diameter and length of the branch certainly have an influence on the temporal duration of the contact formation and expulsion phases.

These parameters are extremely difficult to integrate into the estimation process: the contact point is not indicated in [4], the diameter is available but not for all tests, and no information is available on the sample length.

In addition to these limitations, it should be noted that the model is limited to the first 25s of the failure, in this time interval a certain similarity of current evolution can be observed between the various tests. After a time interval ranging from 20s to 30s for the majority of the tests the second phase ends, the current evolution is extremely random in the phases following the expulsion of moisture, and the only similarity to be found between the various datasets concerns the important increase in current due to flashover, which occurs, however, at extremely different times. In addition, it is possible for a THIF to evolve into an arc fault [4], so it might be useful to combine a model such as the one presented with a model of the arc-earth fault.

Despite these limitations, the model even in the elementary form presented in chapters 5 and 6 can be used for the generation of useful datasets for the training phases of artificial intelligence-based fault detection methodologies. Various paths can be followed to generate data: using the Simulink simulation software as proposed in [6] or using the Power Factory electrical network simulation software (methodology used by the company that proposed this thesis topic).

If an AI-based tool is to be applied for fault detection on medium-voltage distribution networks, it should be installed in the HV-MV transformer substation, and a training phase is required for its proper functioning [8] [9]. Datasets on the medium-voltage network that will be monitored are therefore required.

To obtain these datasets, it is necessary to know the network topology and basic parameters (e.g. neutral status, lengths, longitudinal impedances,

etc.). With this information, it will be possible to construct a power factory model of the network; an example of such a model is provided in Fig.44. Once the network model has been generated, it is possible to use the fault analysis functions offered by the software [26]. In particular, it is possible to provide a fault impedance to be applied, in our case a total impedance as shown in Fig.39. It will be possible to change the fault point, generation-consumption scenario, and operating topology to obtain different datasets from a single impedance model.
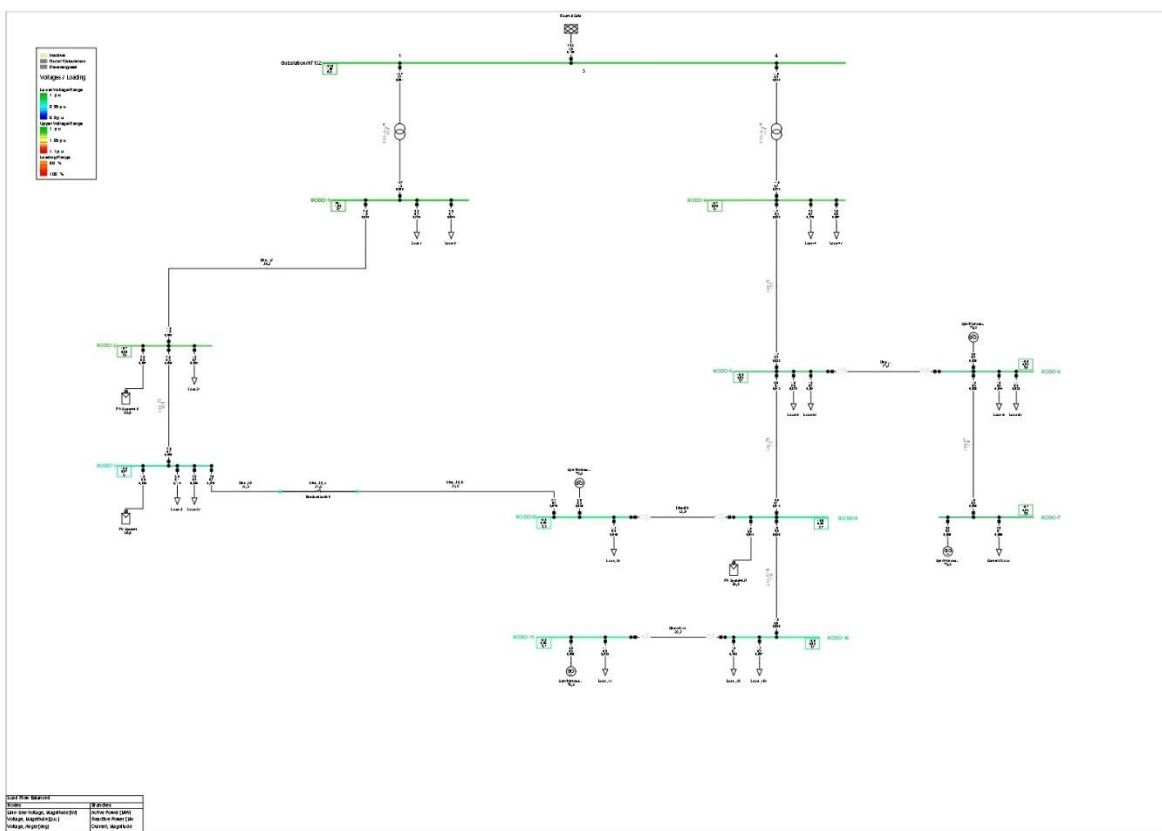
Fig.44 Power Factory model

7 Conclusions and future work

In this thesis work, the focus has been placed on the modernization of electricity grids, and in particular on artificial intelligence-based technologies that could bring important innovation to methods of monitoring, controlling, and securing the power system [2]. As proof of this, if we look at the market and the direction that research is taking in leading companies in the sector, we can see that projects using artificial intelligence-based tools are increasingly numerous. An example of this is the company "Siemens Energy", which is active in numerous projects using AI, such as its implementation for power plant automation [27]; "Schneider Electric", another world leader in the electrical engineering sector, is also very active in the topic, to the extent that it has created an AI-HUB dedicated to research and design [28].
Even this thesis, in its own small way, bears witness to this market focus on the digitization of the electrical system. The subject of analyzing and modelling high impedance faults caused by vegetation contact with phase conductors (THIF) is in fact the result of a very specific request from "Siemens Energy", which is active in the design of AI-based fault detection methodologies in medium voltage distribution networks.

As mentioned previously (1.2), such design phases often come up against the lack of datasets in particular for the algorithm training phase, and the initial aim of this study was precisely to show how this problem can be curbed without resorting to expensive field experiments. The problem was initially approached by analyzing the THIF phenomenon using freely available experimental data, the main characteristics of the fault current were described, and it was shown how the moisture content affects the value of the current.
Following this analysis phase, we proceeded with the description of the model created to derive the value of the current in the first 24[s] as a function of the humidity value; this model was then verified using the same experimental data used previously. Good accuracy was found, but actions

needed to improve accuracy were still reported. To conclude, it was illustrated how such a model can be used to generate datasets useful for the training phase of AI-based fault-detection algorithms.

It is certainly possible to follow up on this thesis work, firstly to improve the accuracy of the model by differentiating it for different vegetation species and then to apply it in network simulation software to derive datasets. Another topic that is considered important to address and analyse concerns the correlation between environmental moisture and vegetation moisture. By providing an AI-based tool with meteorological data, it would then be possible to estimate the THIF fault current instant by instant for better detection accuracy.

Staying with the THIF theme, it is necessary to develop a model for phase-phase failure, which could occur if a branch makes contact with two phase conductors of a power line. Within the experimental data (Appendix 1), phase-phase tests are also available, so it would be possible to follow a similar path to that set out in Chapter 5 for the development of a two-phase THIF model.

In addition to the type of faults dealt with, it would be of great benefit to develop models that would allow the design of AI-based detection tools for faults caused by the presence of salt or ice in the phase conductors of overhead lines.

Special environmental conditions can lead to the formation of ice sleeves on the phase conductors of overhead lines, a phenomenon that is often overlooked and can lead to considerable damage to the electrical system by affecting the mechanical behaviour of the infrastructure [29]. In fact, the ice accumulated on the conductors due to its weight can cause the following problems:

- "sagging", the increase in weight due to the presence of ice in the phase conductors leads to a low-sag, additional tensile forces are introduced which can exceed structural limits and cause ruptures and thus short circuits [30].

- "galloping", is a lateral or vertical oscillation of phase conductors often ignited by wind. With the presence of ice, the additional weight can lead to structural strength limits being exceeded and cause short circuits [31].

Contamination of phase conductors with salt, also caused by environmental conditions, is extremely dangerous and can lead to line damage and thus short circuits [32]. In order to prevent such phenomena, artificial intelligence can be of great help. Thanks to its data-analysis capabilities, large amounts of data can be analysed using, for example, weather recordings or microwave sensors [33].

In addition to the design phases of AI algorithms for monitoring and controlling electricity grids, great attention will have to be paid to the installation phases of these tools. For a fault-detection methodology such as the one discussed in this thesis, it will certainly be necessary to assess how to integrate it into HV-MV transformer substations, a step that is anything but trivial. It is necessary, for example, to assess which communication network structure to use to ensure security; to modify the protection scheme if necessary; to develop the necessary protocols and standards.

References

[1] A. B. F. M. D. I. Mediha E. Mehmed-Hamza, *"Research and analysis of faults in medium voltage distribution grids",* 2018.

[2] M. C. P. L. Tony Marxsen, *Vegetation Conduction Ignition Test report - Final,* Victoria - Australia, 2015.

[3] European Commission, "REPowerEU Plan," in *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee of the regions*, Brussels, 2022.

[4] T. E. C. The European Parliament, *"Fit for 55" - amending Regulation (EU) 2019/631 as regards strengthening the CO2 emission performance standards for new passenger cars and new light commercial vehicles in line with the Union's increased climate ambition,* 2023.

[5] E.DSO, "The value of the digital transformation - Opportunities for distribution system operators (DSOs)," July 2022.

[6] N. S. B. R. J. D.-G. P. Stefanidou-Voziki, *A review of fault location and classification methods in distribution grids,* Elsevier, 2022.

[7] N. S. I. H. N. Saad Abdul Aleem, *Methodologies in power systems fault detection and diagnosis,* 2014.

[8] J. Z. Z. Z. X. W. C. Z. Z. R. Shan, *A Survey: New Generation Artificial Intelligence and Its application in power system dispatching and operation,* Taiyuan, China: IEEE, October 2021.

[9] F. I. H. T. A. Hasbi Ash Shiddieqy, *Power line trasmission fault modeling and dataset generation for AI based automatic detection,* IEEE, 2019.

[10] F. N. H. R. M. Nooshin Bahador, *Modelling and detection of live tree-related high impedance fault in distribution systems,* IET, 2017 .

[11] L. F. Roberto Benato, *Impianti Elettrici (Cap. 7, Cap. 8),* 2014.

[12] R. Turri, *Stato del neutro nelle reti trifase,* 2010.

[13] K. U. B. N. T. H. H. Christian Andre Andresen, *Fault detection and prediction in smart grid,* 2018.

[14] M. K. Z. C. V. S. R. M. Illia Diahovchenko1, *Progress and Challenges in Smart Grids: Distributed Generation, Smart Metering, Energy Storage and Smart Loads,* 2020.

[15] G. U. R. Italiana, «www.gazzettaufficiale.it/eli/id/2014/07/18/14G00113/sg,» 4 Luglio 2014. [Online].

[16] E.-L. d. o. E. laws, «www.eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32009L0072,» [Online].

[17] H. L. G. H. A. M.] Amin Ghaderi, *High impedance fault detection: A review,* 2016.

[18] M. R. H. T. B. R. S. K. J. Peterson, *Staged-Fault testing for high impedance fault data collection,* IEEE, 2005.

[19] C. W. M. T. C. J. Mark Adamiak, *High Impedance Fault Detection On Distribution Feeders.*

[20] J. S. Naser Zamanan, *The Evolution of High Impedance Fault Modeling.*

[21] V. I. R. V. A. S. N. I. V. Yuri V. Makarov, *Blackouts in North America and Europe: Analysis and Generalization.*

[22] B. A. d. S. N. S. D. B. F. B. C. ·. M. R. C. P. J. Wellinsílvio Costa dos Santos, *High Impedance Faults: From Field Tests to Modeling,* 2013.

[23] J. P. Y. K. T. K. S. R. Nam, *A modeling method of a High Impedance Fault in a Distribution System Using Two Series Time-Varying Resistance in EMTP,* 2001.

[24] M. h. c. "-W. models", « What are Hammerstein-Wiener Models? - MATLAB & Simulink - MathWorks Nordic,» [Online].

[25] C. O. Douglas P. S. Gomes, *VeHIF: An Accessible Vegetation High-Impedance Fault Data Set Format.*

[26] Digsilent, «https://www.digsilent.de/en/short-circuit-analysis.html,» [Online].

[27] S. Energy, «https://www.siemens-energy.com/global/en/news/magazine/2022/unleashing-ai-in-power-plants.html,» 2022. [Online].

[28] S. Electric, «https://www.se.com/ww/en/about-us/newsroom/news/press-releases/schneider-electric-advances-its-artificial-intelligence-ai-strategy-with-appointment-of-chief-ai-officer-and-creation-of-new-ai-hub-617aa598b7359463ac1f3144,» [Online].

[29] M. F. Konstantin Savadjiev, *Modeling of Icing and Ice Shedding on Overhead Power Lines Based on Statistical Analysis of Meteorological Data,* 2004.

[30] Q. H. Arsalan Habib Khawaja, *Estimating Sag and Wind-Induced Motion of Overhead Power Lines With Current and Magnetic-Flux Density Measurements,* 2017.

[31] B. Y. L. Z. S. Z. Linshu Zhou, *Study on galloping behavior of iced eight bundle conductor transmission lines,* 2015.

[32] A. J. J. J. W. S. R.K. Aggarwal, *An overview of the condition monitoring of overhead lines,* 2000.

[33] O. N. M. H. Z. Dima Kilani, *Intelligent Ice Detection based on Artificial Neural Network Using Microwave Sensor.*

Appendix 1

"Powerline Bushfire Safety Program - Vegetation Conduction Ignition Test Report and Data" www. discover.data.vic.gov.au/dataset – Department of Environment, Land, Water Planning – 2016

Appendix 2

Script 1 – MATLAB

This script performs current estimation for unknown impedance values. It begins by performing polynomial interpolation of the impedances of the various tests that will be used within the for loop for estimation, by interpolating the unknown impedance. Finally, the relationship between voltage (a function is used to generate a sinusoid with operator-selected parameters) and estimated impedance is performed to obtain the current.

```
clear all
load('tt')
load('z238')
load('z260')
load('z269')
load('tt269')
load('z270')
load('tt230')
load('z230')

figure;
plot(tt, z38)
hold on
plot(tt,z260)
plot(tt,z270)
hold off

T=(0:0.1:24)

W238 = 41.8;
W260 = 33;
W269 = 43.2;
W270 = 50.7;
W230=56;

coeff=polyfit(tt,z260,6);
Z260=polyval(coeff, T);

coeff1=polyfit(tt,z270,6);
Z270=polyval(coeff1, T);

coeff3=polyfit(tt269,z269,6);
Z269=polyval(coeff3, T);

coeff2=polyfit(tt,z38,6);
Z38=polyval(coeff2, T);

coeff4=polyfit(tt230,z230,6);
Z230=polyval(coeff4, T);

figure;
plot(T,Z260)
hold on
plot(T,Z270)
plot(T,Z38)
plot(T,Z269)
plot(T,Z230)
xlabel('time [s]')
ylabel('impedance [ohm]')
legend(["260-33%wt", "270-50.7%wt", "238-41.8%wt", "269-43.2%wt","230-56%wt"])
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Definisci il vettore dei valori di tempo (X)
X = T; % Sostituisci con il vettore dei valori di tempo

% Definisci il vettore dei valori di umidità per le tre curve (W1, W2, W3)
W = [W260, W238, W270,W269,W230]; % Concatena i vettori di umidità in un unico vettore

% Definisci il vettore dei valori di tempo per i quali desideri stimare Z4(t)
X_desiderati = 0:0.1:24;

% Definisci il valore di umidità desiderato per la curva Z4(t)
W4 = 39.4 ; % Sostituisci con il valore di umidità desiderato

% Inizializza il vettore per memorizzare i valori stimati di Z4(t)
Z4_stimato = zeros(size(X_desiderati));

% Ciclo su tutti i valori di tempo desiderati
for i = 1:length(X_desiderati)
    X_current = X_desiderati(i);
```

d

```matlab
    [~, idx] = min(abs(X - X_current));

    % Estrai i valori di Z per le tre curve corrispondenti all'indice trovato
    Z1_current = Z260(i);
    Z2_current = Z38(i);
    Z3_current = Z270(i);
    Z269_current = Z269(i);
    Z230_current = Z230(i);

    % Calcola la curva Z4(t) stimata per il valore di umidità W4
    Z4_stimato(i) = interp1(W, [Z1_current, Z2_current, Z3_current,Z269_current, Z230_current] - W + W4, W4);
end

coeffZ4=polyfit(T,Z4_stimato,6);
ZSTIMATO=polyval(coeffZ4, T);
figure;
plot(T,ZSTIMATO)
hold on
plot(T,Z4_stimato)
title(['estimated impedance'])
xlabel('time [s]')
ylabel('Impedance [ohm]')
% Grafico della curva stimata Z4(t)
figure;
plot(X_desiderati, Z4_stimato, 'r');
xlabel('Time [s]');
ylabel('Impedance [ohm]');
title([' estimated impedance, moisture [%wt.] = ' num2str(W4)]);
hold on
plot(T,Z260)
plot(T,Z270)
plot(T,Z38)
plot(T,Z269)
plot(T,Z230)
legend(["estimated", "260-33%wt", "270-50.7%wt", "238-41.8%wt", "269-43.2%wt","230-56%wt"])
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%costruzione Z
load('res1_tempo.mat');
tempoassoluto=(Resistenza1.Rmod(2,:)-Resistenza1.Rmod(2,1)).*(10^(-5)); %%%in secondi
%trova medio
min=Resistenza1.Rmod(1,1);
massimo=Resistenza1.Rmod(1,56);
media=(min+massimo)/2;
disp(min);
disp(massimo);
disp(media);

Media_vet=repmat(media,1,146);
figure;
plot(tempoassoluto, Media_vet, 'r');
hold on
plot(tempoassoluto, Resistenza1.Rmod(1,:), '-o' )
hold off
xlabel('time [s]')
ylabel('Impedance [ohm]')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
n_inrementi=24/0.04; disp('numero periodi metà n_incrementi, n_incrementi:'); disp(n_inrementi);
% Inizializzazione del vettore
timee= zeros(1, 600);

% Costruzione del vettore
for i = 1:600
    timee(i) = (i-1) * 0.04;
end

ZSTIMATA2=polyval(coeffZ4, timee,'-o');
figure;
plot(timee, ZSTIMATA2)
xlabel('time [s]')
```

e

```matlab
ylabel('Impedance [ohm]')
Dz=ZSTIMATA2-ZSTIMATA2(1);
dz=ZSTIMATA2(1)-media; disp(dz);
Zeta_zero=media+dz; disp('valore media iniziale su cui valutare incremento'); disp(Zeta_zero);
Z_start= Resistenza1.Rmod(1,:)+dz;
% load('Z_start238_3.mat')
figure;
plot(tempoassoluto, Z_start, '-o' )
% hold on
% plot(tempoassoluto, Z_start238_3, '-o' )
xlabel('time [s]')
ylabel('Impedance [ohm]')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Z_finale=[];
t_finale=[];
for k=1:length(Dz)
    Z_finale=[Z_finale, Z_start+Dz(k)];
    t_finale=[t_finale, tempoassoluto+timee(k)];
end
ZZ_finale=[];
tt_finale=[];
% for k=1:length(Dz)
%     ZZ_finale=[ZZ_finale, Z_start238_3+Dz(k)];
%     tt_finale=[tt_finale, tempoassoluto+timee(k)];
% end

figure;
plot(t_finale, Z_finale, 'o')

xlabel('time [s]')
ylabel('impedance [ohm]')
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
tensione=mySinusoid(t_finale,(22000/(sqrt(2))),50,0);
figure;
plot(t_finale, tensione.*(-1),'r')
corrente=tensione./Z_finale;
% cor2=tensione./ZZ_finale;
load('cor447.mat')
load('tempo447.mat')
xxi=tempo447.*(10^-5);

figure;
plot(t_finale, corrente);

hold on

% plot(t_finale, cor2, 'r');
hold on
plot(xxi,cor447, 'g', 'color', [0 1 0 0.2]);

legend(["estimated", "original"])
xlabel('time [s]')
ylabel('current [A]')
hold off
```

f

Script 2 – MATLAB

This code is used for the construction of R1(t).

```matlab
clear all
load('res1_tempo.mat');
```

```matlab
tempoassoluto=(Resistenza1.Rmod(2,:)-Resistenza1.Rmod(2,1)).*(10^(-5)); %%%in secondi

figure;
plot(tempoassoluto, Resistenza1.Rmod(1,:), '-o' )


tensione=mySinusoid(tempoassoluto,(22000/(sqrt(2))),50,0)
plot(tempoassoluto, tensione.*(-1),'-o')

%%%
corrente=((tensione.*(-1))./Resistenza1.Rmod(1,:));
plot(tempoassoluto, corrente,'r')
```

```matlab
%ripeto forma onda R1 fino ad ottenere 890 cicli (cicli assoluti R2)
incremento=tempoassoluto(end);
tempo_ripetuto = cumsum(repmat(incremento, 1, 445));
tempo_assoluto_ripetuto = repmat(tempoassoluto, 1, 445); % replico tempo_assoluto lungo l'asse delle colonne
incrementi_ripetuti = repmat(tempo_ripetuto, 1, 146); % replico gli incrementi lungo l'asse delle colonne
tempo_assoluto_ripetuto = tempo_assoluto_ripetuto + incrementi_ripetuti; % sommo gli incrementi a tempo_assoluto

Rmod_ripetuto = repmat(Resistenza1.Rmod(1,:), 1, 445);
plot(tempo_assoluto_ripetuto(:), Rmod_ripetuto(:),'o');

tensione2=mySinusoid(tempo_assoluto_ripetuto ,(22000/(sqrt(2))),50,0)
plot(tempo_assoluto_ripetuto , tensione2.*(-1))
```

```matlab
%trova medio
min=Resistenza1.Rmod(1,1);
massimo=Resistenza1.Rmod(1,56);
media=(min+massimo)/2;
disp(min);
disp(massimo);
disp(media);

Media_vet=repmat(media,1,146);
plot(tempoassoluto, Media_vet, 'r');
hold on
plot(tempoassoluto, Resistenza1.Rmod(1,:), '-o' )
hold off
```

```matlab
%%%R2 ottenuta come R2=Rtot-R1med
%R2_assoluta=R2_interp-Media_vet;
%figure;
%plot(t_campionamento, R2_assoluta)
```

g

Script 3 - MATLAB

This script interpolates the current provided by the display script, then the zeros are derived to isolate the periods of interest.

```matlab
%%%
clear all
intervallo=296251:2696247; %%%con 860000:4000000 si vede più non-lin %260000:2000000
phase1I=Current_time(intervallo,:);
phase1V=Voltage_time(intervallo,:);

%%% porzione misurazioni fase 1

figure;

plot(phase1I(:,1), phase1I(:,2))



%%% interpolazione corrente

Colonnainterp=2;

x=linspace(1, size(phase1I,1), size(phase1I,1));
y=phase1I(:,Colonnainterp);

xi=linspace(1, size(phase1I,1), 50000000); %50000000
yi=interp1(x, y, xi, 'linear'); %linear
plot(xi, yi, 'r');

cor447=yi;
tempo447=xi;

%zeri corrente

hold on
xlim = get(gca, 'XLim');
plot(xlim, [0, 0], 'k--');
indici_zeri = find(abs(yi) < 0.00008);

plot(xi(indici_zeri), yi(indici_zeri), 'go', 'MarkerSize', 5, 'MarkerFaceColor', 'g');

xlabel("Time [s]")
ylabel("Current [A]")
legend('Current', 'y=0', 'Zeros');
hold off

coord_zeri = [xi(indici_zeri)', yi(indici_zeri)'];

%%%isolamento perioso corrente

periodo = find(xi >= coord_zeri(2400) & xi <= coord_zeri(2404)); %2400-2404
halfperiodo = find(xi >= coord_zeri(2400) & xi <= coord_zeri(2401));
halfperiodo2 = find(xi >= coord_zeri(2401) & xi <= coord_zeri(2402));
%%%%CORRENTE PERIODO%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
xi_intervallo = xi(periodo);
yi_intervallo = yi(periodo);
xi_half = xi(halfperiodo);
yi_half = yi(halfperiodo);
xi_half2 = xi(halfperiodo2);
yi_half2 = yi(halfperiodo2);
```

h

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure;
plot(xi_intervallo, yi_intervallo, 'r', 'LineWidth', 2);

%%% interpolazione tensione

Colonnainterp=2;

xx=linspace(1, size(phase1V,1), size(phase1V,1));
yy=phase1V(:,Colonnainterp);

xxi=linspace(1, size(phase1V,1), 50000000); %50000000
yyi=interp1(xx, yy, xxi, 'linear');
plot(xxi, yyi, 'r');

periodov = find(xi >= coord_zeri(2400) & xi <= coord_zeri(2404)); %2400-2404
halfperiodov = find(xi >= coord_zeri(2400) & xi <= coord_zeri(2401));
halfperiodov2 = find(xi >= coord_zeri(2401) & xi <= coord_zeri(2402));
%%%%TENSIONE PERIODO%%%%%%%%%%%%%%%%%%%%%%%%%%%%
xxi_intervallo = xxi(periodov);
yyi_intervallo = yyi(periodov);
xxi_half = xxi(halfperiodov);
yyi_half = yyi(halfperiodov);
xxi_half2 = xxi(halfperiodov2);
yyi_half2 = yyi(halfperiodov2);
% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

figure;
plot(xxi_intervallo, yyi_intervallo, 'r', 'LineWidth', 2)

figure;

plot(yi_intervallo, yyi_intervallo)

yyisymm = yyi_intervallo.*(-1);


figure;
plot(yi_intervallo,yyisymm)
axis([min(yi_intervallo) max(yi_intervallo) min(yyisymm)-1 max(yyisymm)+1]);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%R1

dati_It=xi_intervallo(1:750:end); %750
dati_I=yi_intervallo(1:750:end); %750

dati_vt=xxi_intervallo(1:750:end); %750
dati_v=yyi_intervallo(1:750:end); %750

dati_vt2=dati_vt.*(10^(-5))
dati_It2=dati_It.*(10^(-5))
figure;
plot(dati_It2, dati_I, '-o');
figure;
plot(dati_vt2, dati_v, '-o');

%%%rapporto v/I
```

i

```matlab
R1=dati_v./dati_I;
R11=R1(2:end);
R111=R11.*(-1);
dati_Itt=dati_It(2:end);
figure;
plot(dati_Itt, R11, '-o');

%%%calcolo deviazione standard

stdev=std(R11);
idx=abs(R11 - mean(R11)) > 2*stdev;

dati_Itt(idx)=[];
R11(idx)=[];

finaldati_Itt=linspace(min(dati_Itt), max(dati_Itt),150); %150
finalR11=interp1(dati_Itt, R11, finaldati_Itt, 'linear');

figure;
plot(finaldati_Itt, finalR11,'-o');
```

Script 4 - MATLAB

This script is used to visualise the data in Appendix 1 (the first 30 lines of
code can also be downloaded from Appendix 1). The downloaded code was
then modified for a better visualisation of the waveforms and to perform the
voltage/current relationships to derive the impedance.

Dataset visualizer

```
clear all
dataset_path='hif_vegetation_dataset.h5'

dataset_path = 'hif_vegetation_dataset.h5'
```

Visualizing groups and attributes

```
h5info(dataset_path).Groups
```

```
Error using h5infoc
Unable to open 'hif_vegetation_dataset.h5'. File or folder not found.

Error in h5info (line 125)
    hinfo = h5infoc(filename,location, useUtf8);
```

```
h5info(dataset_path).Groups(1).Groups
h5info(dataset_path).Groups(1).Groups(1).Attributes
h5info(dataset_path).Groups(2).Groups
h5info(dataset_path).Groups(2).Groups(300).Attributes
```

Visualizing waveforms

```
test_number = 447;

disp(sprintf('Test number - %d',test_number))
disp(sprintf('Fault type - %s',h5readatt(dataset_path,strcat('/test/',num2str(test_number)),'fault_type')))
disp(sprintf('Max set current - %d A',h5readatt(dataset_path,strcat('/test/',num2str(test_number)),'max_current')))

voltage_lf = h5read(dataset_path,strcat('/test/',num2str(test_number),'/voltage_lf'));
current_lf = h5read(dataset_path,strcat('/test/',num2str(test_number),'/current_lf'));
figure;
subplot(2,1,1)
plot(voltage_lf)
title('Low-frequency voltage')
subplot(2,1,2)
plot(current_lf)
title('Low-frequency current')

subplot(2,1,1)
xlabel("ms")
ylabel("V")
subplot(2,1,2)
xlabel("ms")
ylabel("A")
voltage_lf = h5read(dataset_path,strcat('/test/',num2str(test_number),'/voltage_lf'));
current_lf = h5read(dataset_path,strcat('/test/',num2str(test_number),'/current_lf'));

n = length(voltage_lf);
m=1000; %646:1000
num_bloks = ceil (n/m);
Z=zeros(n,1);
t_sample=1/100000; %s
time_vector=zeros(n,1);

for k = 1:n
    if k ==1
        time_vector(k)=0;
    else
        time_vector(k)=time_vector(k-1)+t_sample;
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Current_time = [time_vector, current_lf];
Voltage_time = [time_vector, voltage_lf];
figure;
subplot(2,1,1)
plot(Current_time(:,1), Current_time(:,2))
% subplot(2,1,2)
% plot(Voltage_time(:,1), Voltage_time(:,2))
subplot(2,1,1)
xlabel("Time [s]")
ylabel("Current [A]")
% subplot(2,1,2)
% xlabel("Tme [s]")
% ylabel("Voltage [V]")
for i = 1:num_bloks
    start_index = (i-1)*m+1;
    end_index = min(i*m,n);
    Z(start_index:end_index)=voltage_lf(start_index:end_index)./current_lf(start_index:end_index);
end

figure;
plot(Current_time(:,1), Z);
```

```matlab
time_vector2=time_vector(620471:end); %646:255953 %270:529494 %555:439825  389:612508  706:695791  556:503202  260:597027  269:768546  260:
current_lf2=current_lf(620471:end); %646:255953 %270:529494
Current_time2 = [time_vector2, current_lf2];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
voltage_lf2=voltage_lf(620471:end); %646:255953 %270:529494
Voltage_time2 = [time_vector2, voltage_lf2];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
figure;
subplot(2,1,1)
plot(Voltage_time2);
subplot(2,1,2)
plot(current_lf2);
nn = length(voltage_lf2);
mm=1000;
num_bloks2 = ceil (nn/mm);
for ii = 1:num_bloks2
    start_index2 = (ii-1)*mm+1;
    end_index2 = min(ii*mm,nn);
    ZZ(start_index2:end_index2)=voltage_lf2(start_index2:end_index2)./current_lf2(start_index2:end_index2);
end

figure;
plot((ZZ));
ZZ_tr=ZZ';
time_vectorASS=time_vector2-time_vector2(1);
ZZ_time=[time_vectorASS,ZZ_tr];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fc = 100; % frequenza di taglio %646-270-555:92
fs = 100000; % frequenza di campionamento
[b, a] = butter(6, fc/(fs/2), 'low'); % coefficienti del filtro
filtered_signal = filter(b, a, ZZ_time(:, 2)); % applicazione del filtro
ttt = ZZ_time(:, 1); % tempo
yyy = ZZ_time(:, 2); % segnale originale
y_filtered = filtered_signal; % segnale filtrato
figure; % nuova figura
plot(ttt, yyy, ttt, y_filtered); % plot segnale originale e filtrato
xlabel('Tempo [s]'); % asse x
ylabel('Ampiezza'); % asse y
legend('Segnale originale', 'Segnale filtrato'); % legenda
figure;
plot(ttt, y_filtered);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ttt2=ttt(1:6000000);
y_filtered2=y_filtered(1:6000000);
filtroout=[ttt2,y_filtered2];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
figure;
plot(ttt(1:6000000), y_filtered(1:6000000)); %646-270-555:6000000  556:1:8000:4000000
hold on
plot(ttt(1:12000:6000000), y_filtered(1:12000:6000000),'r');%646: 1:12000:6000000
ttt2=ttt(1:6000000);
y_filtered2=y_filtered(1:6000000);
filtroout=[ttt2,y_filtered2];
ttt3=ttt(1:12000:6000000);
y_filtered3= y_filtered(1:12000:6000000);
filtroout2=[ttt3,y_filtered3];
figure;
plot(ttt(1:12000:6000000), y_filtered(1:12000:6000000),'-o');
```

```matlab
% fcc = 30; % frequenza di taglio %646:92
% fss = 100000; % frequenza di campionamento
% [bb, aa] = butter(4, fcc/(fss/2), 'low'); % coefficienti del filtro
% filtered_signal2 = filter(bb, aa, y_filtered); % applicazione del filtro
% ttt4 = ZZ_time(:, 1); % tempo
% yyy4 = y_filtered; % segnale originale
% y_filtered4= filtered_signal2; % segnale filtrato
% figure; % nuova figura
% plot(ttt(1:12000:6000000), y_filtered(1:12000:6000000), ttt4(1:6000000), y_filtered4(1:6000000)); % plot segnale originale e filtra
% xlabel('Tempo [s]'); % asse x
% ylabel('Ampiezza'); % asse y
% legend('Segnale originale', 'Segnale filtrato'); % legenda
```

I

m