ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 12, Issue, 10, pp. 59844-59858, October, 2022
https://doi.org/10.37118/ijdr.25609.10.2022

RESEARCH ARTICLE OPEN ACCESS

DEVELOPMENT OF A COMPARATIVE STUDY BETWEEN LINEAR REGRESSION AND DECISION TREE TECHNIQUES FOR 3rd HARMONIC

Dener Jeferson Horta de Aquino*[1], Jandecy Cabral Leite[2] and Rivanildo Duarte de Almeida[2]

[1,2]Postgraduate Master in Engineering, Process Management, Systems and Environmental (EPMSE/ITEGAM), Manaus – Amazonas, Brazil.
[2]Institute of Technology and Education Galileo of Amazon (ITEGAM), Manaus-Amazonas, Brazil.

ARTICLE INFO

Article History:
Received 22nd September, 2022
Received in revised form
30th September, 2022
Accepted 03rd October, 2022
Published online 30th October, 2022

KeyWords:

Voltage Harmonic Distortion, Current Harmonic Distortion, Electric Power Quality.

****Corresponding author:***
Dener Jeferson Horta de Aquino

ABSTRACT

A comparative study was carried out between the Simple Linear Regression and Decision Tree techniques in the evaluation of the impacts caused by the generation of harmonics, in the distribution network, in a medium voltage substation. Data collection was carried out through a measurement campaign in a 13.8 kV substation that feeds a portion of the companies in the industrial district of the city of Manaus. The objective of this dissertation is to carry out a comparative study of the results of the application of Decision Tree and Simple Linear Regression techniques in the analysis of harmonic impacts in an electrical system. The measurement campaign was carried out for a minimum period of 7 calendar days according to PRODIST. The impact of these harmonics on the electrical system is quite harmful, both for consumers and for utilities. The techniques used in the methodology of this dissertation are the Decision Tree technique, which features the construction of non-parametric models and the Simple Linear Regression technique, which features the construction of parametric models and simple mathematical calculation, easy to interpret and analyze. of data. In this way, this dissertation presented and applied in practice with case studies actions for the analysis of harmonic impacts in electrical energy distribution systems through the construction of mathematical models using Simple Linear Regression analysis and Regression Tree analysis, obtaining important and enlightening results. in the studies carried out, thus validating the techniques applied in the analysis of harmonic impacts in electrical energy distribution systems.

Citation: Dener Jeferson Horta de Aquino, Jandecy Cabral Leite and Rivanildo Duarte de Almeida. 2022. "Development of a comparative study between linear regression and decision tree techniques for 3rd harmonic*", International Journal of Development Research, 12, (10), 59844-59858.*

INTRODUCTION

Electricity is used as a product in all sectors of human life, has a set of specific properties and is directly involved in the development of other types of products, thus requiring optimum power quality, and each electrical/electronic product is designed to operate at certain electrical power parameters: nominal frequency, voltage, current, etc., therefore, for its normal operation, the power quality of electricity must be provided (NAZIROV *et al*, 2021). Power quality is defined based on four important measurements of electrical parameters, i.e. voltage, current, frequency and phase; therefore, both voltage and current should be in sinusoidal form with specified magnitude at a constant frequency without any phase shift (ZOBAA AND ALEEM, 2017). Non-linear loads are defined as sources of harmonic currents that contaminate the electrical system with disturbances such as distortion in the voltage waveform (SOUZA *et al*., 2021).

Power companies produce electrical energy where the generated voltage is maintained sinusoidal using synchronous generators and after some operations, this electrical energy is transferred to consumers at low voltage, except for some industrial consumers and High Voltage (HV) and Medium Voltage (MV) consumers. The use of non-linear loads generates high harmonic currents that end up affecting the phase and line voltages of the electrical system. These harmonic currents cause heating on the conductor surface and cause higher eddy current loss in transformers, and in addition, harmonic currents increase the non-linear voltage drop in transmission lines, causing a distortion in the waveform of the supply voltage that affects consumers in general (SALAM, UDDIN AND MOINUDDIN, 2019). The effects of harmonic currents on power grids are reflected in the energy losses that occur due to the deforming regimes in the voltage and current waveforms of the power grid, leading to increased electricity production costs and oversizing of grid elements, and most electrical equipment is very sensitive to the deformation of voltage and current waveforms as well as the presence of one or more particular harmonics (DOBREF, POPOV AND MOCANU, 2018). According to (ULLAH *et al*., 2019), it is essential the knowledge of the sources of harmonics installed in the electrical system, voltage disturbances and interruptions in the waveform, to then implement improvements for electric power quality. According to (ALMEIDA AND LEITE, 2018), to solve the problem of harmonic distortion, the most common technique is based on the application of passive filters, however an important process that must be performed before implementing the solution, is to identify the main sources of harmonics in an electrical system and know the percentage of contribution of each source, in order to solve more effectively the problem of harmonic distortions of voltage and current. Thus, it is necessary to identify and to measure the harmonic distortions levels in the electric system under study and to identify in which feeders or points there is the biggest contribution in the harmonic distortions, in other words, to identify the non-linear loads that are harming the electric system, to then be able to take the best decision in the solution of this problem. This dissertation uses two techniques of computational intelligence denominated Decision Tree and Linear Regression, to identify and measure the harmonic impacts caused by the injections of harmonic currents generated by non-linear loads. The techniques used seek a correlation between injected harmonic current and the harmonic voltage at the electrical system bar, thus identifying where there is the largest contribution of harmonic sources.

This dissertation presents a comparative study of results in the application of the Decision Tree and Linear Regression techniques in harmonic impact analysis in an electric system. With the purpose of establishing the harmonic simultaneity between the disturbing equipment's, that is, the non - linear loads, and sensitive equipment's, it is necessary the establishment of limits and standards for control of such feats or phenomena. The standards are: Institute of Electrical and Electronics Engineers (IEEE), in the United States, International Electrotechnical commission (IEC), in Europe and Procedures for Distribution of Electrical Energy in the National Electric System (PRODIST), in Brazil. Throughout this dissertation these standards will be addressed in more detail. Currently, electric utilities and consumers are devoting more attention to improve the quality of generated and distributed electric power, and the main objectives are to produce clean electric power and distribute it to end customers with acceptable power quality performance in a cost-effective manner (ZOBAA, ALEEM AND BALCI, 2018). Power quality comprises a wide range of electromagnetic phenomena, among which harmonics stand out, defined as frequency components that are integer multiples of the fundamental frequency of the electrical system, which affect telephone communications and control and protection systems, in addition to increasing network losses, damaging sensitive equipment, causing excessive heating in electrical machinery and being a cause of resonance (MARRERO *et al*, 2021). Power utilities are therefore under pressure to maintain pure sinusoidal line phase voltage waveforms, avoiding harmonics generated by non-linear consumer loads. On the other hand, these harmonics not only affect the voltage waveform, but also produce reactive power affecting the power factor, and in addition, the power quality standard is becoming a sensitive issue due to the additional number of connected nonlinear devices that impose a non-sinusoidal current waveform on the system, and many sensitive loads connected to a power system depend on a pure and stable sinusoidal supply voltage (SALAM, UDDIN AND MOINUDDIN, 2019).

As discussed above, the problem of harmonic distortions is a fact. Therefore, this research is relevant due to the need to measure and identify the harmonic distortion levels and the impacts on the electrical grid, in order to propose solutions that can reduce the impacts caused on the electrical grid, among which we can also mention.

a) Overload in the electrical systems due to the increase in the effective current;
b) Overload in the neutral conductors due to the 3rd order harmonic currents;
c) Heating, vibrations and ageing of alternators, transformers and motors;
d) Heating and ageing of the reactive energy compensation capacitors;
e) False performance of sensors and protection equipment's;
f) Burning or inadequate operation of sensitive equipment;
g) Disturbance in communication networks or telephone lines;
h) Measurement errors.

This research has as contribution the practical application of a valid methodology using two computational intelligence techniques applied case studies for harmonic impact analysis considering the recommendations of IEC 61000-3-6 (IEC/TR, 2008-02), IEEE Std. 519-2014 and PRODIST/MODULO 8/ANEEL, revision year 2021, which serves as a base and support material for more advanced research on harmonic distortion and power quality as a whole. This dissertation also contributes, not only to the incentive of the creation of normative documents that impute responsibilities on harmonic distortion limit violations, but also to the stimulus of the inclusion in electric energy tariffs the implications of the harmonic content of the loads that pollute the feeding system.

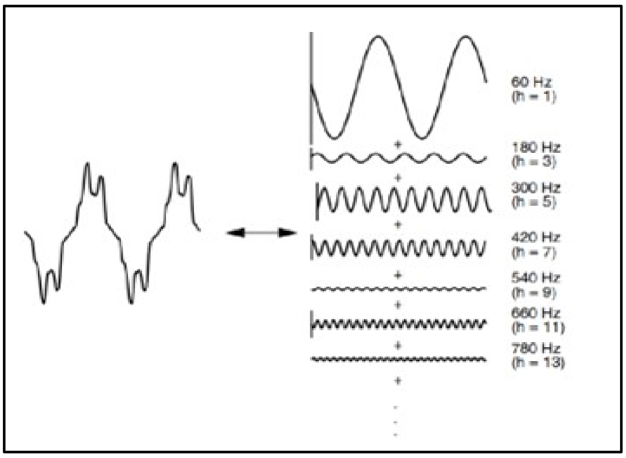
BIBLIOGRAPHIC REVIEW

Fundamentals of harmonic distortions and power quality: According to (IEEE-1159, 2019), the definition of power quality is the wide variety of electromagnetic phenomena occurring in the power grid at a given time and place that characterize voltage and current; these measurements are compared to ideal or acceptable values. In the current era of utility deregulation and competition, the impact of harmonics and interharmonics on equipment and system operations has generated serious concern. Nowadays, it is well known that harmonics have adverse effects on the entire power system, such as malfunctioning of important control and protection equipment, overheating of transformers and overloading of other power appliances, and on the other hand, inter-harmonics cause lighting fluctuations, misfiring of thyristor devices and fluctuation of monitor image or display (AMOO, ALIYU AND BAKARE, 2018). (ZOBAA, ALEEM AND BALCI, 2018) states, all electrical equipment can fail or malfunction when they encounter power quality disturbances, depending on the severity of the disturbance, it is essential that engineers, technicians, manufacturers and power system operators understand well and cope with the various power quality disturbances. Power quality problems include voltage variations (dips, interruptions, swings, etc.), transients (surges, lightning, and switching events), and grounding problems. Figure 1 summarises the common power quality problems. The phenomenon of harmonic power distortion was recognized by utilities in the early 1920s and 1930s when distorted voltage and current waveforms were observed on transmission lines [9, 10]. Harmonics are high frequency steady state energy involving multiple frequencies of the fundamental (60 Hz) flowing along with the fundamental frequency in a power network that can adversely affect the performance of the power system (AMOO, ALIYU AND BAKARE, 2018). (DUGAN *et al*., 2004) state that according to Fourier, any continuous and periodic function over any interval can be represented by a summation of sine components and a constant component. According to (DE OLIVEIRA, 2018), the notion of harmonics in the electrical sense became known in the second half of the twentieth century, when the sine component is of the same frequency as the original signal it is called fundamental and for the other sine components, whose frequencies are integer multiples of the fundamental frequency, they are called harmonic frequencies as shown in Table 1 and Figure 1. For (DE OLIVEIRA, 2018), a distorted waveform can always be written as the superposition of a waveform of the fundamental frequency with other waveforms of different frequencies and harmonic amplitudes. The harmonic spectrum shown in Figure 2 is a good way to identify the decomposition of the waveform in Figure 1.

Table 1. Harmonic Frequencies

Harmonic Order	Network Frequency 50 Hz	Network Frequency 60 Hz
1ª	50	60
2ª	100	120
3ª	150	180
4ª	200	240
5ª	250	300
...	...	...
Nª	50 x n	60 x n

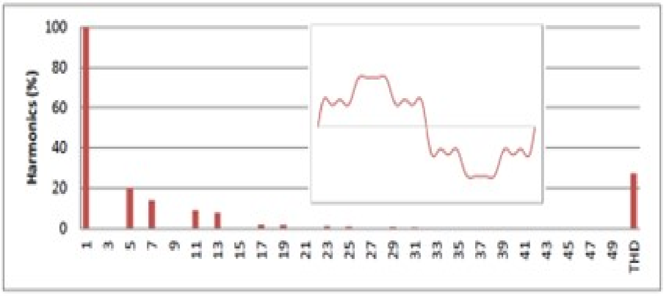
Source: (KAMENKA, 2014).



Source: (DUGAN *et al*, 2004).

Figure 1. Fourier Series Representing a Distorted Waveform

This type of spectrum is also used by almost all power quality measurement devices.



Source: (KAMENKA, 2014).

Figure 2. Distorted Wave and the Distorted Wave Spectrum

For ALBADI, *et al*, (2017) states, total harmonic distortion (THD) is used to define the effect of harmonics on power system voltage. IEEE 519-2014 defines THD as "the ratio of the root mean square of the harmonic content, considering harmonic components up to the 50th order and specifically excluding inter-harmonics, expressed as a percentage of the fundamental". In other words, THD is the contribution of all harmonics to the fundamental. THD is calculated as described by the following formula:

$$THD = \frac{\sqrt{\sum_{h=2}^{h_{max}} M_h^2}}{M_1} \qquad (1)$$

where M_1, is the fundamental component of the rms value of the voltage or current signal. To assess the current harmonic distortion, the total demand distortion (TDD) is commonly used. IEEE 519-2014 defines TDD as "the ratio of the root mean square of the harmonic content, considering harmonic components up to the 50th order and specifically excluding inter-harmonics, expressed as a percentage of the maximum demand current".

$$TDD = \frac{\sqrt{\sum_{h=2}^{h_{max}} I_h^2}}{I_m} = \frac{\sqrt{I^2 - I_1^2}}{I_m} \qquad (2)$$

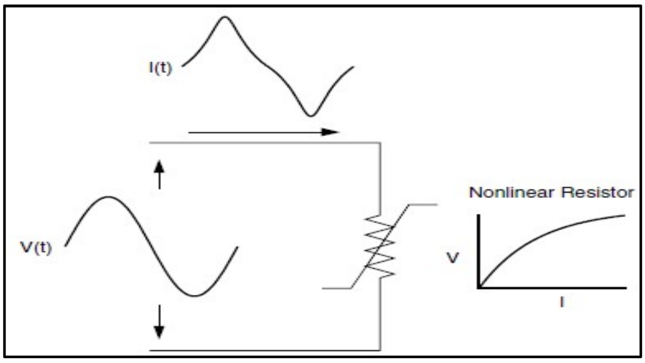
where I_is the maximum demand load current and I_1is the rms value of the fundamental component.

The harmonics sequences can be positive, negative and zero, as described below (PROCOBRE, 2003):

a) Positive sequence: they tend to make the motors rotate in the same direction as that of the fundamental component, thus causing an overcurrent in their windings;
b) Negative sequence: they tend to make motors turn in the opposite direction to the rotation produced by the fundamental, thus braking the motor and also causing undesired heating;
c) Zero sequence: emergence of a neutral current 3 (three) times greater than the phase current, causing excessive heating of the neutral conductor (PROCOBRE, 2003).

Definition of Linear Loads: According to DA SILVEIRA, BERNARDON AND RADUNS, (2018) explains, linear loads are constituted by resistances, inductances and capacitances, where the current and voltage waveforms will always be sinusoidal, except when powered by non-sinusoidal signals, that is, the current waveform of a linear load will be identical to the voltage waveform that feeds it, with only the phase shift. Examples of linear loads: heaters, boiler, incandescent lamps and motors (fans, compressors, pumps, lathe, electric drill, milling machine, etc.) (NASSAR *et al*, 2020).

Definition of Nonlinear Loads: For DE OLIVEIRA, (2018) explains, a load is considered nonlinear if the current projected by the load is not sinusoidal even when it is connected to a sinusoidal voltage. This non-linear current contains frequency components that are multiples of the power system frequency. These harmonic currents interact with the impedance of the power grid to create voltage distortion that can affect the power grid itself and the loads connected to it. Examples of non-linear loads: electric welders, fluorescent lamps, high pressure sodium lamps, instrument and control, control of (lifts, cranes grinders, fans, compressors, pumps, mechanical lathe, electric drill, milling machine, etc . . .) and all such loads controlled or operated by electronic devices (NASSAR *et al*, 2020). Figure 3 shows an example of this concept in the case of a sinusoidal voltage applied to a simple non-linear resistor where the voltage and current vary according to the curve shown. While the waveform of the applied voltage is a pure sinusoid, the resulting current has a distorted waveform. Increasing the voltage by a few percent can cause the current to increase twice as much and take on a different waveform. This is the source of most harmonic distortions in an electrical system (DUGAN *et al*., 2004).



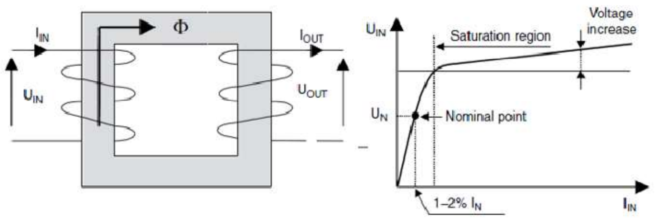
Source: DUGAN *et al*. (2004).

Figure 3. Current Distortion caused by non-linear resistor

Harmonic Sources: Today, as a consequence of the extensive use of power electronics based components in all power system applications, most of today's loads are non-linear. To generalise, three categories can be recognised as primary sources of harmonics in power systems. They are given as follows (ZOBAA, ALEEM AND BALCI, 2018):

(a) equipment based on magnetic cores such as electric motors, power transformers and generators; Arc and induction and arc welders; Equipment based on power electronics.

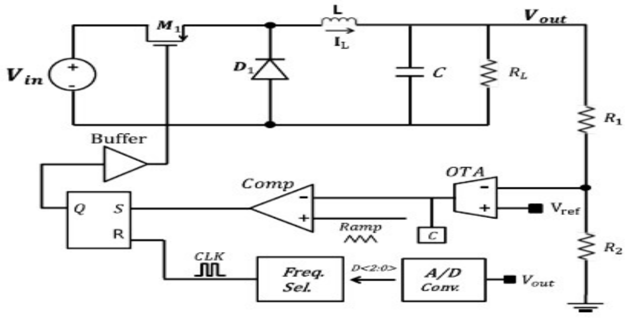
Transformers: Transformers are designed to operate with magnetizing current in the range of 1% to 2% of rated current. The point at which the transformer operates at rated voltage is then located below the "knee" of the magnetization curve, which is the linear region of operation of a transformer, see Figure 4. Even if there are a large number of transformers operating in an electrical power system, they are not considered a significant source of harmonics under normal operating conditions. However if this condition changes due, for example, to a slight voltage increase within the saturation region, even a small voltage increase above the nominal value results in a large increase in the magnetization current and consequently the harmonic content increases significantly (BAGGINI, 2008). The non-linearity of the current transformer core can be assessed using the open circuit test (no load), where results show that the magnetization current of a CT contains harmonic components, and the total harmonic distortion under these conditions can reach a value around 51%, resulting in a THD of 9% in the secondary (MURRAY AND DE KOCK, 2018).



Source: Soares (2012).

Figure 4. Transformer Diagram and Magnetization Curve

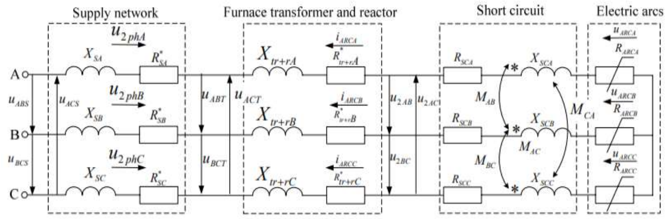
Frequency converters: Linearised modelling of power electronic converters is critical to reveal the causes of harmonic instability in power electronics-based power systems (WANG AND BLAABJERG, 2018). (AHMED *et al*, 2021) states, DC-DC converters are the inevitable components of modern DC distribution systems, and the nonlinear behaviour of these individual converter circuits is thoroughly realized in power electronics circuit studies, and the time-varying behaviour of these converter circuits is due to the low and high frequency oscillations that result in multi-period bifurcations. (WANG AND BLAABJERG, 2018) details, power converters are nonlinear and time-varying dynamic systems, where the nonlinearity is due to the dynamically varying duty cycle (modulator control input) with the closed loop control system, and the time-variation results from the switching modulation process and the time - periodic operating paths of systems In the modern era of technology, most consumer devices are composed of DC-DC converter circuits that require design engineers to focus mainly on eliminating sub-harmonics (AHMED *et al*., 2021).



Source:YU, et al., (2018).

Figure 5. Block Diagram of a Digitally Controlled DC-DC Converter

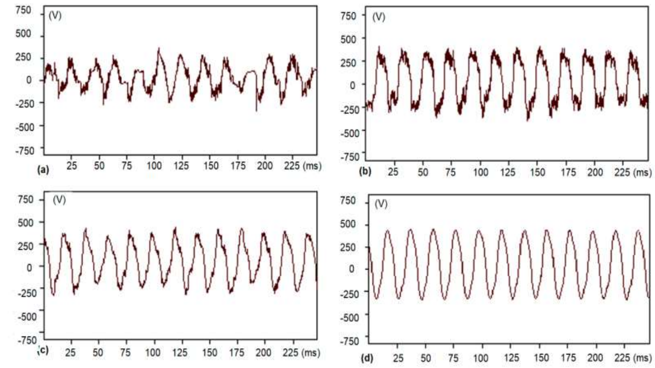
Electric Arc Furnace: According to (JR AND SIMONETTI, 2019), the electric arc furnaces are the main equipment responsible for the transformation of metal scrap into steel in the process of melting and refining of metals, being of essential relevance in the steel industry just to its environmental benefit from the recycling of metallic materials. Therefore, the electric arc furnace in the steelmaking process, becomes of substantial importance in the sustainability of the production chain and human consumption of materials manufactured from steel as a raw material. The analysis of the impact of DC electric arc furnaces (DC EAFs) as a source of harmonic distortions in an electrical network is a complex and relevant objective, and a characteristic of DC furnaces is a very wide frequency range of harmonics and the non-stationary random nature of their appearance during melting, together with certain canonical harmonics of the current that are determined by the rectification circuitry, and the appearance of non-canonical harmonics is linked to the asymmetry of the technological process control pulses due to natural discrepancies, the strongly variable nature of the load and signal distortions in the current loops of the control system. The amplitudes of these harmonics do not exceed the amplitudes of the canonical harmonics and in most cases do not require compensation (ABDULVELEEV et al., 2021). For (NIKOLAEV et al., 2016) explains, the study of the harmonic composition of arc voltage, needs analysis of instantaneous values, and this analysis is made possible by the mathematical model of the circuit as per Figure 6, based on Kirchhoff's laws, differentials instantaneous arc conductivity equation, as well as arc electromotive force equations considering the electrical action similar to a valve.



Source: NIKOLAEV et al.,(2016).

Figure 6. Electric Oven Equivalent Circuit

The degree of arc voltage distortion has a direct effect on the waveform of the electric current in the furnace, and the distorted current flowing in the supply network causes a distorted voltage drop in the network impedance, which subtracting from the sinusoidal (source) supply voltage causes the steel and steel mill supply voltage to be non-sinusoidal. The measured arc voltage waveforms show changes in the voltage shape during the melting process in the arc furnace, seen in Figure 7 below. In the initial phase, after arc ignition, the voltage shows large stochastic changes that result from the change in the arc length moving along the load (LUKASIK AND OlCZYKOWSKI, 2020).



Source: (LUKASIK AND OlCZYKOWSKI, 2020).

Figure 7. Arc Voltage waveforms in Individual Casting Stages: (a) arc ignition; (b) start of charge melting; (c) main melting phase; (d) final casting stage

Effects of Harmonic Distortions: The impact on harmonics can range from lowering the performance of equipment to its serious failure. The effects of power system harmonics can be grouped into two major groups: as effects on power system networks and equipment and effects on telecommunication systems. According to (ZOBAA, ALEEM AND BALCI, 2018), the most common consequences in different sectors of a power system are summarized below:

a) Excessive power losses due to high non-sinusoidal currents, leading to high power bills.
b) The presence of current in the neutral wire with additional losses. Overheating problem may occur.
c) Equipment failure, motor stalling, conductor overloading, fuse blowouts and lamp blackouts.
d) Errors in the measurement of energy consumption.
e) Interference in telecommunications systems and networks.
f) Loss of data in data transmission systems.
g) Malfunctioning of control and protection system performance.
h) Harmonic resonance in series and parallel, which can cause damage to system components, equipment failure and service interruption.
i) Harmonic instability leading to damage of generator shafts.
j) Audible noise in transformers, rotating machines and motor vibrations.
k) Blockages of computerized programmable logic controllers and correct operation.
l) Malfunctioning of voltage regulators and generators with frequent maintenance problems.
m) Premature aging of the equipment.
n) UPS dimensioning problems.
o) Worsening of the power factor of the loads with its adverse effects and useful penalties.

Standards and Recommendations for Power Quality Monitoring: This section will present the standards that limit harmonic distortion levels in electrical distribution networks.There are international and national standards for the control of harmonic distortion of the voltage and current waveforms. They are the following: IEEE Std. 519-2014, EN 50160 and IEC 61000-3-6 (IEC/TR, 2008-02). In Brazil, the Electricity Distribution Procedures of the National Electricity System - PRODIST - Module 8 (ANEEL-PRODIST, 2021) are in force.

PRODIST - Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (Modulo 8): The objective of PRODIST is to establish the procedures regarding the electric power quality - QEE, addressing the quality of the product and the quality of the service provided (PRODIST MODULE 8, 2021). Regarding the product quality, module 8 defines the terminology, characterizes the phenomena, parameters and reference limit values related to voltage compliance on a permanent basis and voltage waveform anomalies, determining mechanisms that allow ANEEL to set standards for QEE indicators (PRODIST MODULE 8, 2021). Regarding the quality of the services provided, module 8 establishes the methodology for verifying the continuity indicators and the response times for urgent events, defining standards and responsibilities (PRODIST MODULE 8, 2021). This dissertation addresses procedures concerning electric power quality only related to harmonic distortions. This subsection will present reference limit values for total acceptable harmonic distortion for each voltage class in the electric power systems (LV - Low Voltage, MV - Medium Voltage and HV - High Voltage). This dissertation is delimited in harmonic impact analysis in medium voltage electrical networks. Table 2 summarizes the terminology applicable to the harmonic distortions calculation. The expressions for the calculation of magnitudes DIT_h %, DTT %, DTT_p%, DTT_i%, DTT_3% are presented in Equations (3), (4), (5), (6) and (7), respectively:

$$DIT_h\ \% = \frac{V_h}{V_1} \times 100 \tag{3}$$

Where:
h = individual harmonic order.

$$DTT\ \% = \frac{\sqrt{\sum_{h=2}^{hmáx} V_h^2}}{V_1} \times 100 \tag{4}$$

Where:
h = all harmonic orders from 2 up to *hmáx*;
hmáx = according to class A or S.

Table 2. Terminology

Magnitude Identification	Símbolo
Individual harmonic distortion of voltage of order h	DIT_h %
Total voltage harmonic distortion	DTT %
Total harmonic voltage distortion for non-multiple-3 even components	DTT_p%
Total harmonic voltage distortion for non-multiple-3 even components	DTT_i%
Distorção harmônica total de tensão para as componentes múltiplas de 3	DTT_3%
Tensão harmônica de ordem h	V_n
Harmonic order	h
Maximum harmonic order	$hmáx$
Minimum harmonic order	$hmin$
Tensão fundamental medida	V_1
DTT% indicator value exceeded in only 5% of the 1008 valid readings	DTT 95%
DTTP% indicator value exceeded in only 5% of the 1008 valid readings	DTT_p95 %
DTTI% indicator value exceeded in only 5% of the 1008 valid readings	DTT_i95 %
DTT3% indicator value exceeded in only 5% of the 1008 valid readings	$DTT_3$95 %

Source: PRODIST MODULO 8 (2021).

$$DTT_P\ \% = \frac{\sqrt{\sum_{h=2}^{hp} V_h^2}}{V_1} \times 100 \tag{5}$$

Where:
h = all even harmonic orders, not multiple of 3 (h = 2, 4, 8, 10, 14, 16, 20, 22, 26, 28, 32, 34, 38.); *hp* = maximum even harmonic order, not multiple of 3.

$$DTT_I\ \% = \frac{\sqrt{\sum_{h=5}^{hi} V_h^2}}{V_1} \times 100 \tag{6}$$

Where:
h = all odd harmonic orders, not multiple of 3 (h = 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35, 37.); *hi* = maximum odd harmonic order, not multiple of 3.

$$DTT_3\ \% = \frac{\sqrt{\sum_{h=3}^{h3} V_h^2}}{V_1} \times 100 \tag{7}$$

Where:
h = all harmonic orders multiple of 3 (h = 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39...);
h_3 = maximum harmonic order multiple of 3.

According to PRODIST MODULE 8 (2021), the limit values for total harmonic distortions are presented in Table 3.

Table 3. Total Harmonic Distortion limits

Indicator	Nominal voltage		
	$Vn \leq 1{,}0kV$	$1{,}0kV < Vn < 69kV$	$69kV \leq Vn < 230kV$
DTT 95%	10,0%	8,0	5,0
DTT_p95 %	2,5%	2,0	1,0
DTT_i95 %	7,5%	6,0	4,0
$DTT_3$95 %	6,5%	5,0	3,0

Source: PRODIST MODULO 8 (2021).

The limits correspond to the maximum target value to be observed in the distribution system. It is important to observe that PRODIST MODULE 8 does not set limits for total current harmonic distortion, only for voltage harmonic distortion. In measurements made using the secondary of VTs with V-type or open delta connection, the allowed limit values for the DTT395% indicator must correspond to 50% of the respective values indicated in Table 3.2 (PRODIST MODULE 8, 2021). Those responsible for the implementation of any equipment in the basic grid shall follow what is determined in the Network Procedures or in specific regulations.

IEEE STD. - 519-2014: The IEEE STD. 519 of 2014 is conceptualized as "recommended practices and standard requirements for harmonic control in power electrical systems". The characteristic of this standard is the determination of the limit values for both, not only the distortion of the voltage supplied by the distribution utility, but also the consumer load current (IEEE STD. 519, 2014). Harmonic limits from IEEE Standard 519 - 2014 have been established for the PCC between the utility and various consumers. Figure 2.9 shows an example of a general shape configuration of an electrical system viewed from the common coupling point between the utility and the consumer unit. In this figure, the consumer unit is represented by linear and non-linear loads, while the electric power utility's system, is represented in a simplified way through a non-detailed complex that includes its fundamental parts. They are: generation, transmission, distribution, other consumers, etc. Analyzing the voltage distortion level (THD) at the PCC, the utility is responsible for maintaining the quality of the voltage waveform at the PCC, below the maximum limit values in permanent regime presented in Table 3. Total harmonic voltage distortion for non-multiple-3 even components.

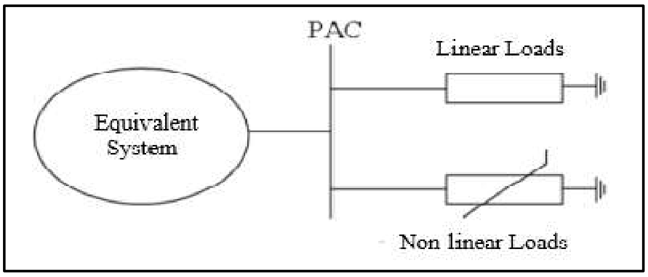
These limits must be met for the worst system operating condition lasting longer than one hour. For short-term periods, during initial or transient conditions, these limits may be exceeded by 50%. The current distortion limits depend on the relationship that exists between the consumer load (IL) and the short-circuit current (Isc) at the PCC. A high Isc/IL means that the system admits higher levels of harmonic distortion (IEEE STD. 519, 2014). Total demand harmonic distortion (TDD) indicator is used to analyse the current harmonic distortion limits. The TDD Indicator can be calculated by the ratio of the square root of the mean square of the harmonic content, weighting the harmonic components up to order 50 and excluding the inter-harmonics, expressed as a percentage of the maximum demand current, expressed by Eq. (8) (IEEE STD. 519, 2014).

$$TDD = \frac{\sqrt{\sum_{h=2}^{\infty} I_h}}{I_1} \times 100 \tag{8}$$

The current harmonic distortion limits are presented in Table 5.

Where:

Isc - is the phase current of the three-phase short circuit at PCC;
IL - is the maximum consumer demand current calculated as the average of those recorded in the last 12 months.



Source: http://200.19.146.153/bitstream/123456789/14588/1/ contribuicaoprocessocompartilhamento.pdf.(2018).

Figure 8. Common Coupling Point (PAC)

Table 4. Voltage Distortion Limit

PAC busbar voltage	Individual harmonics (%)	Total Harmonic Distortion - THD (%)
$V \leq 1kV$	5,0	8,0
$1kV < V \leq 69kV$	3,0	5,0
$69kV < V \leq 161kV$	1,5	2,5
$161kV < V$	1,0	1,5

Source: IEEE STD. 519 (2014).

Table 5. Current Distortion Limits for 120V to 69kV Systems

Maximum harmonic distortion as a percentage of the maximum load current demand (I_L)						
Individual harmonic order (odd harmonics)						
I_{sc}/I_L	$3 \leq h < 11$	$11 \leq h < 17$	$17 \leq h < 23$	$23 \leq h < 35$	$35 \leq h < 50$	TDD
< 20	4,0	2,0	1,5	0,6	0,3	5,0
$20 < 50$	7,0	3,5	2,5	1,0	0,5	8,0
$50 < 100$	10	4,5	4	1,5	0,7	12,0
$100 < 1000$	12,0	5,5	5,0	2,0	1,0	15,0
> 1000	15,0	7,0	6,0	2,5	1,4	20,0

Source: IEEE STD. 519 (2014).

Table 6. Values of each Voltage Harmonic

Odd Numbers Harmonics				Harmonic Pairs	
Não múltiplos de 3		Múltiplos de 3			
Ordem	% of nominal voltage	Ordem	% of nominal voltage	Ordem	% of nominal voltage
5	6%	3	5%	2	2%
7	5%	9	1,5%	4	1%
11	3,5%	15	0,5%	6 até 24	0,5%
13	3%	21	0,5%		
17	2%				
19	1,5%				
23	1,5%				
25	1,5%				
DHTt< 8%					

Source: EN 50160 (2004).

EN 50160 - Power Quality Standard: The European standard EM 50160 determines power quality parameters in PCC, indicating acceptable deviations. Regarding voltage harmonics, for a period of one week, 95% of the effective values of each voltage harmonic (average values every 10 minutes), must not exceed the limit values indicated in Table 6 (EM 50160, 2004).

IEC 61000-3-6 Electromagnetic Compatibility Standard

a) The IEC 61000-3-6 standard (IEC/TR, 61000-3-6, 2008) of the IEC (International Electrotechnical Commission) 61000 series determines the harmonic limit values for loads connected to power electrical systems. For theelectricvoltagethis standard defines (LEITE, 2013):

b) (a) the Total Harmonic Voltage Distortion (%THDV). Calculated as the ratio between the rms voltage of the considered higher harmonics and the fundamental voltage (V1), presented in Eq. (9) (IEC/TR 61000-3-6, 2008).

$$\%THDV = \frac{\sqrt{\sum_{h=2}^{\infty} V_h^2}}{V_1} \times 100 \qquad (9)$$

c) The Individual Harmonic Distortion of voltage of order h (%IHDVh). Calculated as the ratio of the voltage of an individual harmonic (Vh) to the fundamental voltage according to Eq. (10) below (IEC/TR 61000-3-6, 2008).

$$\%IHDV_h = \frac{V_h}{V_1} \times 100 \qquad (10)$$

The reference (IEC/TR 61000-3-6, 2008) establishes two types of limits: (1) compatibility levels as shown in Table 2.6, and (2) planning levels, as presented in Table 7. When the harmonic currents generated by non-linear loads do not exceed the established compatibility limit values, good power quality is guaranteed. Otherwise, the planning limit values (more restrictive than the compatibility limit values) are adopted as roadmaps for the system planning and they ensure compliance with the compatibility limit values (IEC/TR 61000-3-6, 2008).

Table 7. Compatibility Levels for Individual Harmonic Voltages in Low Voltage (LV) and Medium Voltage (MV) Networks

Odd-numbered non-multiple of 3		**Odd multiples of 3**		**Pairs**	
Harmonic Order h	Harmonic Voltage %	Harmonic Order h	Harmonic Voltage %	Ordem de Harmonic Order h	Harmonic Voltage %
5	6	3	5	2	2
7	5	9	1.5	4	1
11	3.5	15	0.4	6	0.5
13	3	21	0.3	8	0.5
17≤ h ≤49	2.27(17/h) - 0.27	21≤ h ≤ 45	0.2	10 ≤ h ≤ 50	0.25(10/h) + 0.25

Source: IEC/TR 61000-3-6 (2008).

Table 8. Planning levels for individual harmonic voltages in medium (MV), high (HV) and extra-high voltage (EHT) networks

Odd-numbered non-multiple of 3			**Odd multiples of 3**			**Pairs**		
Harmonic Order h	Harmonic Voltage %		Harmonic Order h	Harmonic Voltage %		Harmonic Order h	Harmonic Voltage h %	
	MT	AT-EHT		MT	AT-EHT		MT	AT-EHT
5	5	2	3	4	2	2	1.8	1.4
7	4	2	9	1.2	1	4	1	0.8
11	3	1.5	15	0.3	0.3	6	0.5	0.4
13	2.5	1.5	21	0.2	0.2	8	0.5	0.4
17 ≤ h ≤ 49	1.9(17/h)-0.2	1.2(17/h)	21≤ h ≤ 45	0.2	0.2	10 ≤ h ≤ 50	0.25(10/h) + 0.22	0.19(10/h) +0.16

Source: IEC/TR 61000-3-6 (2008).

Voltage harmonic limit values are determined to avoid the harmful effects of harmonics on a permanent regime and in a short period of time which is defined as:

(a) permanent regime effects are referred to heating in capacitors, cables, transformers, motors and the other, and are measured on average at 10-minute interval (IEC/TR 61000-3-6, 2008);

b) Short time period effects are those that manifest themselves in electronic equipment sensitive to harmonic limit values, and with a recording interval in the range of 3 seconds or less (IEC/TR 61000-3-6, 2008).

The THDV compatibility level for medium and low voltage is 8% for permanent regime harmonics (10-minute intervals) and 11% for short duration harmonics (intervals shorter than 3 seconds) (IEC/TR 61000-3-6, 2008). According to LEITE (2013), the compatibility limits in Table 6 and planning limits in Table 7 for individual harmonic voltages are valid for stationary harmonics. These limits must be modified by the factor Khvs to be used with short duration harmonics according to Eq. (11).

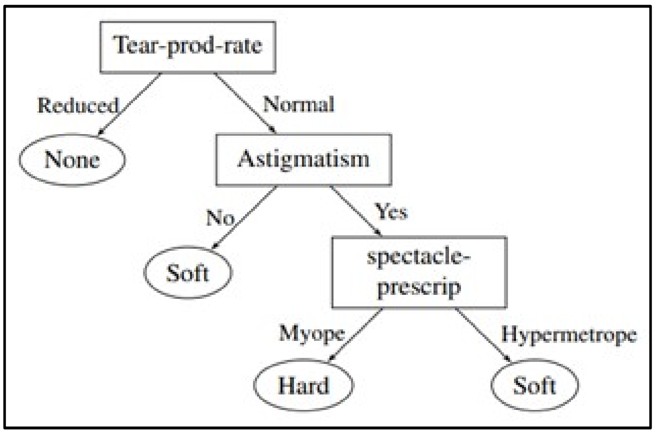
$$\boldsymbol{K_{hvs} = 1.3 + 0.7\frac{h-5}{45}} \qquad (12)$$

According to LEITE (2013), this standard does not expressly determine limit values for harmonic distortion of current in the PCC, however, it considers the chance of converting the voltage harmonic distortion limit values into current harmonic distortion limits when using the impedance at harmonic frequencies of the external power system.

Computational Intelligence Techniques: This section will discuss the computational intelligence techniques used to perform the analyses of the data collected in this work.

Decision Tree: According to (CHEN et al, 2019), Decision Tree technique is a type of supervised learning algorithm that ismainly used in classification problems without any prior assumption placed on the data, and this algorithm has a tree structure similar to a flowchart, where each internal node denotes a test on an attribute, each branch represents a test result and leaf nodes represent classes or class distributions. Compared to other learning algorithms, decision trees have some advantages, such as being robust to noise, low computational cost in building the classifier models, ability to handle repetitive attributes and exclude unnecessary features.

Two different methodologies can be distinguished within the decision tree: classification trees and regression trees. According to (RODRIGUEZ-GALIANO et al, 2015), the learning algorithm of the decision tree technique is executed as follows: (a) To induce DT, recursive partitioning and multiple regressions are performed from the dataset, (b) From the root node, the process of data partitioning at each internal node of a tree rule is repeated until a previously specified stopping condition is reached, (c) Each of the terminal nodes, or leaves, has attached to it a simple regression model that applies only to that node, (d) Once the tree induction process is finished, pruning can be done with the aim of improving the generalization ability of the tree by reducing its structural complexity. The number of cases in nodes can be considered as removal criteria.The Figure below describes an example of decision tree designed for classification, obtained by running the J48 algorithm (C4.5) on the Weka" contact lenses" dataset, considers whether the patient should wear rigid, soft or no contact lenses. The internal nodes (attribute test) are referred to as rectangles, although the leaf nodes are indicated as circles.



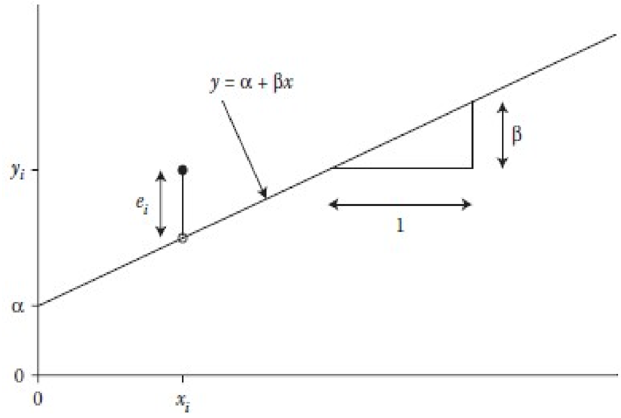
Source: .

Figure 9. An example of Decision Tree: Attributes {Tear production rate, Astigmatism, eyeglass prescription}, Classes {None, Rigid, Soft}

Linear Regression: Linear regression is a data analysis technique used to model the relationship between several variables (independent variables) and an outcome variable (dependent variable). A linear regression model is a probabilistic model that takes into account the randomness that can affect a given outcome. Depending on the previously known input values, a linear regression model predicts a value for the dependent variable (outcome) (SEDKAOUI AND KHELFAOUI, 2020). Regression analysis is used to determine the cause and effect relationship between two or more variables and can vary according to the measurement levels of the dependent variable, and linear regression analysis is used if the type of dependent variable is quantitative and there is a linear relationship between the dependent variable and the independent variable(s). The regression analysis applied is called Simple Linear Regression Analysis if the number of independent variables is one and Multiple Linear Regression Analysis when there are more than one (YASAR et al, 2019) and (LI et al, 2019). According to WELHAM et al. (2015), the simple linear regression model is used to describe the relationship between a response variable (y_i) and a quantitative explanatory variable (x_i) and takes the form of a straight line passing through the scatter of points that arise when the values of the response variable are plotted against the values of the explanatory variable. This model can be represented by Eq. (13).

$$y_i = \alpha + \beta x_i + e_i \qquad (13)$$

Where $y_i$and $x_i$are the response values and the explanatory variables, respectively, for the ith observation. The quantity $e_i$represents a random deviation for the ith observation, sometimes called residual error, and the subindex *i*ranges from *1* to *N,* where *N* is the total number of observations. The model is a straight line defined in terms of the parameters α and *β*, where the parameter α is called the intercept parameter, which corresponds to the point at which the straight line intersects the y-axis when x_i is equal to *0* (zero), and the parameter *β*is called the coefficient of the explanatory variance, it is the slope of the straight line, that is, the change in the response produced by a unit change in the explanatory variable. The simple linear regression model is called simple because it contains a single explanatory variable, and, linear because the response is expressed in a linear form, that is, as a sum of terms that each consists of a coefficient multiplied by an explanatory variable. Figure 10 illustrates in detail the simple linear regression model (WELHAM et al., 2015).



Source: WELHAM et al. (2015).

Figure 10. Simple Linear Regression Model

According to SICHANI and KHALAFINEJAD (2011), the most common method for finding the regression line is the least squares. This method calculates the best-fit straight line for the observed data by minimizing the sum of squares of the vertical distances from each point to the straight line, as illustrated in Figure 10.

The sum of squares of these distances can then be written as:

$$S(\alpha,\beta) = \sum_{i=1}^{n} e_i{}^2 = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \qquad (14)$$

The values of α and β that minimize S (α, β) are given:

$$\hat{\beta} = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \qquad (15)$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \qquad (16)$$

The parameters $\hat{\alpha}$and e $\hat{\beta}$ are called least squares estimates of α and β because they are the solution to the least squares method, the intercept and slope of the straight line that has the smallest possible sum of squares of the vertical distances from each point to the straight line. For this reason, this straight line is called the least squares regression line (LOFTUS, 2022). After building the regression model through the application of simple linear regression, it is essential to evaluate this model through analysis of variance, in such a way as to ensure whether the regression model built is adequate to explain the relationship between the dependent variable and the independent variables. The analysis of variance is a statistical tool that, by means of statistical inference techniques, analyzes two main propositions: the null proposition, which means the possibility of the parameter β (slope of the straight line) being null; and the alternative assumption, which means the possibility of this same parameter not being null (MATLOFF, 2017). According to (LOFTUS, 2022), in the analysis of variance, three essential parameters are obtained: the total sum of squares (SS_{Total}), the error sum of squares (SS_{Res}) and the regression sum of squares (SS_{Reg}). These parameters are calculated using Eqs. (17), (18) and (19), respectively.

$$SS_{Total} = \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad (17)$$

$$SS_{Reg} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 \qquad (18)$$

$$SS_{Res} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (19)$$

Where:
$\hat{y}_i$ - is the estimated value of variable y by the regression model;

$\bar{y}$ is the hope of variable y;
y_i - is the value of variable y used to build the regression model.

De posse desses três parâmetros, SS_{Total}, SS_{Res}, SS_{Reg}, pode-se calcular o valor do indicador de confiabilidade do modelo construído, denominado como coeficiente de determinação R^2. De acordo com RAWLINGS *et al.* (2001), o coeficiente de determinação R^2 é a proporção da soma total dos quadrados da variável dependente pelas variáveis independentes no modelo, conforme a Eq. (20). Com base no valor do coeficiente de determinação R^2, pode-se rejeitar ou não o modelo de regressão.

In possession of these three parameters, SS_{Total}, SS_{Res},SS_{Reg},, one can calculate the value of the reliability indicator of the model built, called the coefficient of determination R^2. According to RAWLINGS et al. (2001), the coefficient of determination R^2 is the proportion of the total sum of squares of the dependent variable by the independent variables in the model, according to Eq. (20). Based on the value of the determination coefficient R^2, one can reject or not the regression model.

$$R^2 = \frac{SS_{Reg}}{SS_{Total}} = \frac{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} = 1 - \frac{SS_{Res}}{SS_{Total}} \qquad (20)$$

The following chapter will present the materials and methods used in the development of this research.

MATERIALS AND METHODS

The methodology proposed in this dissertation is based on the accomplishment of a comparative study of a data correlation analysis, using the Decision Tree and Linear Regression techniques, in such a way to extract a regression tree model and a linear regression model which describes the existent relation between the harmonic current of a non-linear load (harmonic current generator source) and the harmonic voltage at the PCC of the electrical grid which is intended to be analysed, and, therefore, provides a better diagnosis of the non-linear load influence on the voltage harmonic distortion level of the point under analysis. These models are built based on measurements of the rms value of harmonic voltages and currents obtained in the field with power quality analyzers, therefore the voltage and current transducers, and the integralization range of these measurement equipment's exert a great influence on the construction of these regression models. The equipment used to perform the field measurement campaigns was the power quality analyzer HIOKI MODEL PW3198, which will be discussed in more detail in later subtopics. The techniques used in the methodology of this dissertation are the decision tree technique, which has as its characteristic the construction of non-parametric models, and the linear regression technique, which has as its characteristic the construction of parametric models and simple mathematical calculation, of easy interpretation and data analysis. A decision tree is a tree-based method in which each path starting at the root node represents a sequence of splitting data until a Boolean result is reached at the leaf node, where each path in the decision tree is a decision rule that can be easily translated into human languages or programming languages. A regression tree (a type of supervised learning technique) is a prediction model, which is tree structured, for a given numeric objective characterized by a vector of attributes that are also numeric. The regression tree is a category of decision tree, therefore, it is trained and tested using the algorithm called Classification and Regression Tree (CART). The simple linear regression technique has as its characteristic the construction of parametric models and simple mathematical calculation. As previously discussed the simple linear regression technique is used to describe the relationship between a response variable (y_i) and a quantitative explanatory variable (x_i) and assumes the form of a line passing through the dispersion of points that arise when the values of the response variable are plotted against the values of the explanatory variable. The reliability indicator of the constructed model is coefficient of determination R^2. The coefficient of determination R^2 is the ratio of the total sum of squares of the dependent variable by the independent variables in the model, according to Eq. (21). According to value of determination coefficient R^2, one can reject or not the constructed regression model. The correlation intensity associated with the R^2 value is presented in Table 9.

Table 9. Correlation Intensity R^2.

ValueR^2	Correlationintensity
0,00	Nil
(0,00 – 0,09)	Low
0,09 – 0,36)	Moderate
(0,36 – 0,81)	High

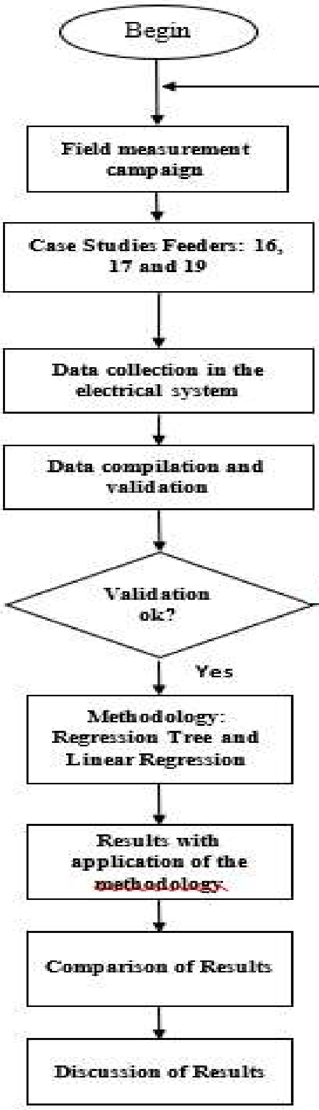
Source: Authors, (2022).

H is justified by the simplicity and easy computational interpretation in the creation of the models, since the simplicity of such models helps to ensure that the regression models built are easy to interpret, thus obtaining a faster response time. The regression models resulting from the application of the regression tree and linear regression techniques on the measured harmonic voltages and currents were built using the AAQEE software, which will be discussed in the next subtopic. The methodology consists in comparing the results of the analysis of the harmonic impacts performed by the two techniques according to the text above, both qualitatively and quantitatively, but only the qualitative analysis is not enough, it is necessary to quantify the harmonic contribution of each customer or feeder of the analyzed electrical system, because probably all customers or feeders of the analyzed electrical system presented contribution for the harmonic voltage distortions in the point or bar of the analyzed electrical system. The techniques used in this dissertation were applied in three case studies in such a way as to show the validity of their applications in solving problems that seek to investigate the dependency relationship between two variables.

AAQEE Software: The AAQEE software is a dedicated data mining tool for electric power quality (EQ) analysis capable of estimating the degree of individual harmonic impact of industrial consumers on electric power distribution networks, using simultaneous measurements of harmonic voltages, currents and powers, as well as computational intelligence techniques. With it, it is possible to perform graphical analysis of the electric quantities data obtained from measurement campaigns by QEE analyzers, calculating the respective QEE indicators established by PRODIST2021, as well as presenting the functionality of calculating the harmonic impacts caused by consumer units on the utility distribution network.

Power Quality Analyzer HIOKI PW 3198: The PW3198 power quality analyzer is an analytical instrument for monitoring and recording anomalies in the electric power supply, allowing to quickly investigate the causes. This equipment can be applied to analyze problems during the electric power supply, such as: voltage drop, fluctuations, harmonics, etc. The features of PW3198 equipment are the following: (a) It has class A complying with IEC 61000-4-30, (b) It performs high frequency transient overvoltage measurements of maximum 6kv peak with sampling rate of minimum 1MHz, (c) Measurement and recording of harmonic and inter-harmonic voltage and current according to IEC 610004-7 and with minimum sampling rate of 256 samples/cycle, (d) Measurement and recording of phase angles of harmonic and inter-harmonic voltage and current; (e) Measurement of harmonic powers and power factor, (f) Measurement of flicker according to IEC 61000-4-15, (g) Measurement of inrush current, (h) Measurement of sags, swells and interruptions, (i) Mass memory of at least 2 GB, (j) Three current sensors in the range (approx.): 1A to 100A, (l) Three flexible current sensors in the range: 200 A to 5000 A.

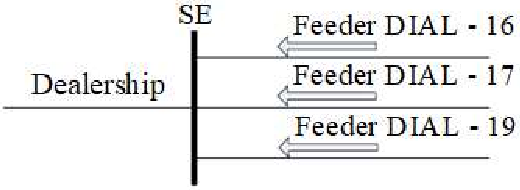
Flowchart of the research: The general structure of the research as shown in Figure 11 defines the steps to be followed according to the methodology adopted for the same, the scope of work and the structure of the literature review. The next chapter will present the results obtained with the application of simple linear regression and decision tree techniques for the evaluation of harmonic impacts on electrical networks.



Source: Authors, (2022).

Figure 11. Flowchart of the research

Application of the linear regression and decision tree techniques: This chapter will present the results of the application techniques Linear Regression and Decision Tree through case studies.The case studies were performed in the electrical system of an electric power utility company, which for ethical reasons its name will not be mentioned. In order to evaluate the methodology previously proposed, a study was performed in three feeders and one bus of the electrical system, in such a way to determine the influence of some non-linear loads in the harmonic distortion of the bus voltage under study, as shown in Figure 12.



Source: Authors, (2022).

Figure 12. Identification of the harmonic impact of each feeder on a common bus in a substation

During the studies in the system, measurements of harmonic voltages and currents were obtained at strategic points of the system, which allowed the construction of regression models that presented the existing relationship between these quantities.The analyses were performed through a field measurement campaign carried out from May 15, 2017 to May 22, 2017 in a 13.8 kV voltage level substation of the Industrial District, in which 4 HIOKI PW 3198 model QEE analyzers were installed to perform simultaneous measurements at the following measurement points: transformer DITF4-04; and feeders DIAL2-16, DIAL2-17 and DIAL2-19. Annex A shows the single-line diagram of the Industrial District substation and the location of the power quality analyser installation points (points circled in blue) for this measurement campaign, totalling 4 simultaneous measurement points. The purpose of the installation of the QEE analyzer in transformer DITF4-04 is to monitor the harmonic voltage at the DIBR2-03 bar (circled in green).These analyses seek to evaluate the existing correlation between order 3a, 5a and 7a harmonic currents of feeders DIAL2-16, DIAL2-17 and DIAL2-19, and the same order harmonic voltages at the DIBR2-03 (13.8 kV) bar of this substation, thus encompassing the three sequences, zero (3a), negative (5a) and positive (7a). These feeders serve companies that have large amounts of non-linear loads installed, such as CNC machines, electric arc furnaces, plastic and aluminum injection machines and others, therefore, large sources of harmonics.

Application of the Linear Regression Technique

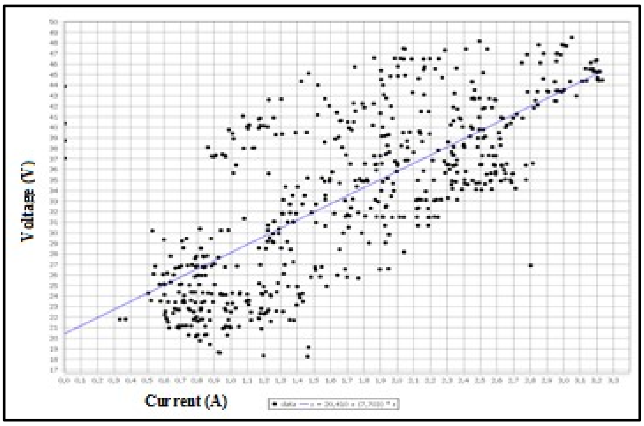
Case Study 1: Impact of 3rd Harmonic: Based on field measurement data, Table 10 presents the percentage impact of the 3rd order harmonic current at DIBR2 - 03 bar of phases A, B and C of feeders DIAL2 - 16, DIAL2 - 17 and DIAL2 - 19 respectively in relation to the background (VBG - unmeasured feeders) using the linear regression technique. Analysing the R^2 metric values of feeder DIAL2 - 16, it is observed that phase A and B show high correlation intensity and low correlation intensity in phase C whose R^2 is 0.043. Table x indicates that phase A is responsible for about 37% of the harmonic voltage distortion impact at DIBR2 - 03 bar caused by the injected harmonic currents at feeder DIAL2 - 16 in relation to VBG, phase B is responsible for about 100% in relation to VBG and phase C is responsible for about 12% in relation to VBG. The simple linear regression models between the 3rd order harmonic currents and voltage of same harmonic order of the DIBR2-03 bar, in phases A, B and C of the feeder DIAL2 - 16 are illustrated in Figures 13, 14 and 15, respectively.

Table 10. Loading plateau, determination coefficient and impact factor of DIAL2 - 16, DIAL2 - 17 and DIAL2 – 19

Load Levelling - Morning and Afternoon			
Start Time		End Time	
00:00:00		14:00:00	
R^2			
Power Feeders	Phase A	Phase B	Phase C
DIAL2 – 16	00,526	00,385	00,043
DIAL2 – 17	00,009	00,165	00,561
DIAL2 – 19	00,755	00,009	00,438
Impact Factor (%)			
BASE	Phase A	Phase B	Phase E C
DIAL2 – 16	37,225	104,527	12,367
VBG-16	62,775	-04,527	87,633
DIAL2 – 17	05,784	41,570	28,812
VBG-17	94,216	58,430	71,188
DIAL2 – 19	59,494	10,680	73,474
VBG-19	40,506	89,320	26,526

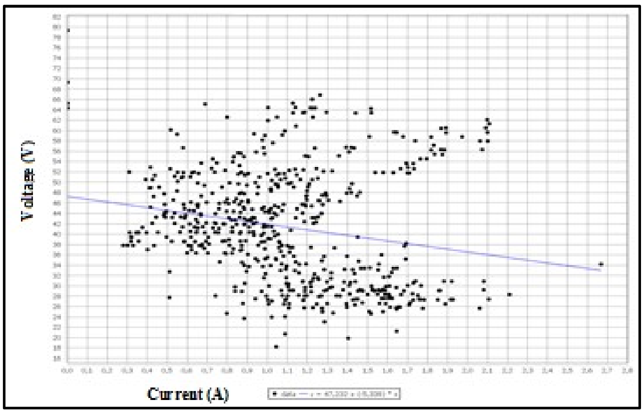
Source: Authors, (2022).

Analyzing the R^2 metric values of feeder DIAL2-17, it is observed that phase A has low correlation intensity, phase B presents moderate correlation intensity and in phase C presents high correlation intensity whose R^2 is 0.561. Table x indicates that phase A is responsible for about 5% of the harmonic voltage distortion impact on bar DIBR2 - 03 caused by the injected harmonic currents at feeder DIAL2-17 in relation to VBG, phase B is responsible for about 41% in relation to



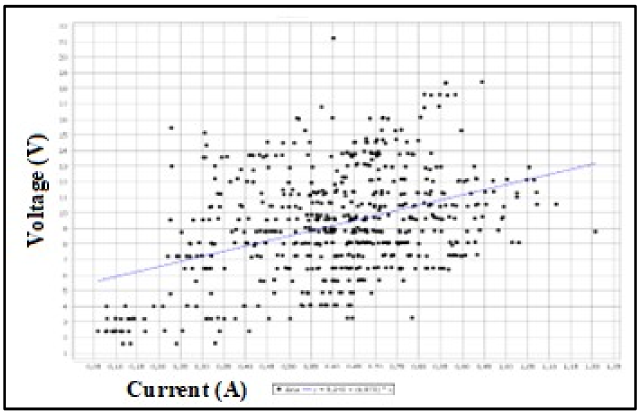
Source: Authors, (2022).

Figure 13. 3rd harmonic of PHASE A, R^2=0.526



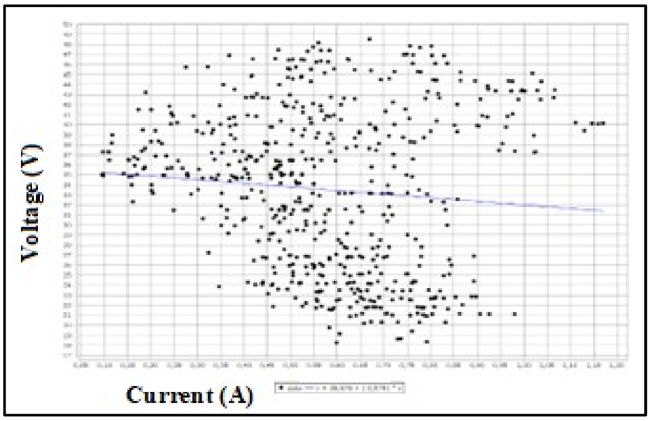
Source: Authors, (2022).

Figure 14. 3rd harmonic of PHASE B, R^2=0,385



Source: Authors, (2022).

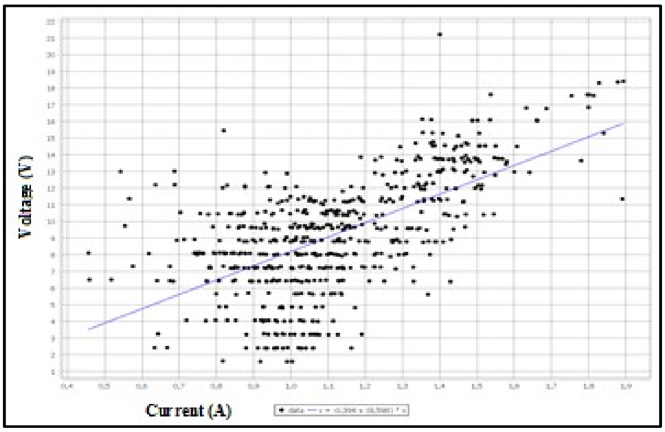
Figure 15. 3rd harmonic of PHASE C, R^2=0,043



Source: Authors, (2022).

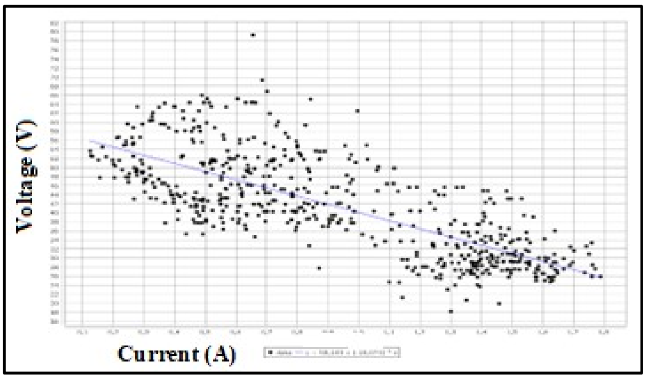
Figure 16. 3rd harmonic PHASE A, R^2=0.009

VBG and phase C is responsible for about 28% in relation to VBG. The simple linear regression models between 3rd order harmonic currents and voltage of same harmonic order of DIBR2-03 bar, in phases A, B and C of feeder DIAL2-17 are illustrated in Figures 16, 17 and 18, respectively.



Source: Authors, (2022).

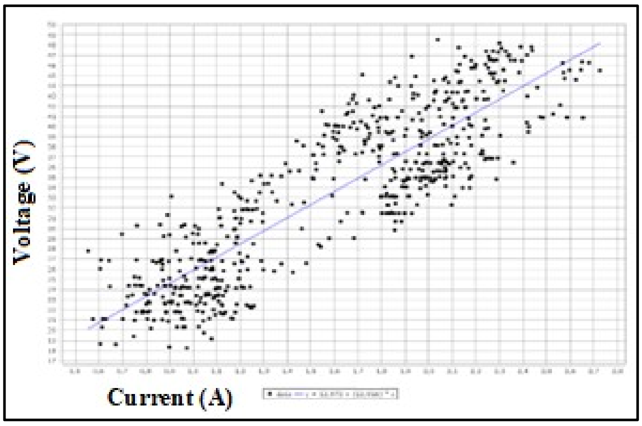
Figure 17. 3rd harmonic PHASE B, R^2=0.165



Source: Authors, (2022).

Figure 18. 3rd harmonic PHASE C, R^2=0.561

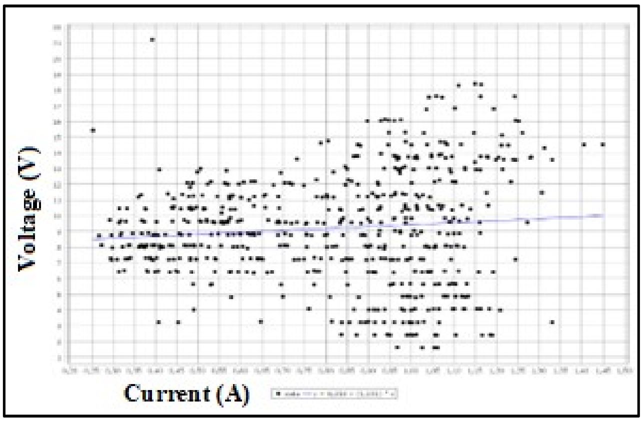
As per Figure 18 the contribution of phase C is negative, which means that the harmonic current injected into phase C will help reduce the voltage distortion levels at DIBR2 - 03 bar. It is important to note, however, that the results are derived from linear regression analysis and there is a margin of error. It is more prudent to conclude that the C-phase of the DIAL2-17 feeder has negligible impact on the voltage distortion levels at the DIBR2 - 03 bar. And analysing Figure 16, the impact of phase A is also insignificant. The R² metric of DIAL2-19 feeder shows that phase A has high correlation intensity with a R² of 0.755, phase B presents low correlation intensity, and phase C presents high correlation intensity with a R² of 0.438.



Source: Authors, (2022).

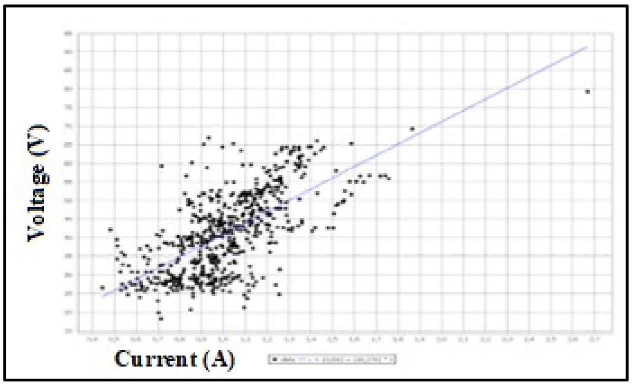
Figure 19. 3rd harmonic of PHASE A, R^2=0.755

Table 11 indicates that phase A is responsible for about 59% of the harmonic voltage distortion impact on bar DIBR2-03 caused by the harmonic currents injected on feeder DIAL2-19 in relation to VBG, phase B is responsible for about 10% in relation to VBG and phase C is responsible for about 73% in relation to VBG. The simple linear regression models between the 3rd order harmonic currents and voltage of same harmonic order of the DIBR2-03 bar, in phases A, B and C of the feeder DIAL2-19 are illustrated in Figures 19, 20 and 21, respectively.



Source: Authors, (2022).

Figure 20. 3rd harmonic of PHASE B, R^2=0.009



Source: Authors, (2022).

Figure 21. 3rd harmonic of PHASE C, R^2=0.438

Application of the Decision Tree Technique

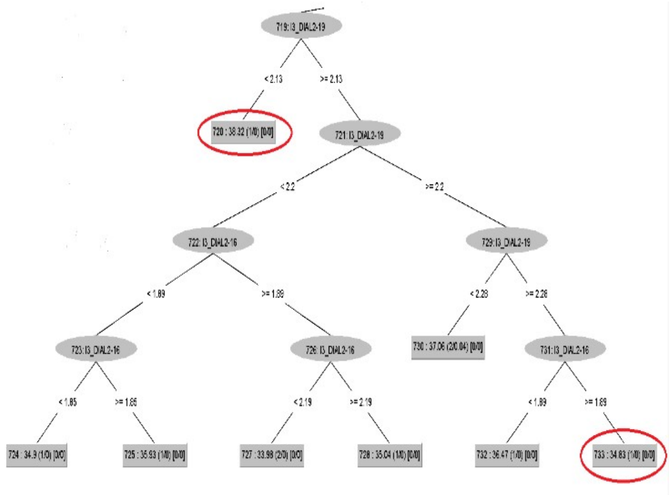
Case Study 1: Impact of the 3rd harmonic: In case study 1 the contribution of feeders DIAL2-16, DIAL2-17 and DIAL2-19 to the increase of harmonic voltage distortion at DIBR2-O3 bar for 3rd harmonic order in phases A, B and C was analysed. During the analysis of the results provided by the application of the decision tree technique it was found that phases A and B of feeder DIAL2-16, phases B and C of feeder DIAL2-17 and phases A and C of feeder DIAL2-19 have higher contribution in harmonic voltage distortion at DIBR2-O3 bar when compared to the results of phases B and C of feeder DIAL2-19 and phases A and C of feeder DIAL2-19.

Table 11. Harmonic Impact Factor for 3rd order (%) MAE

BASE	PHASE A	PHASE B	PHASE C
DIAL2-16	30,366	36,407	23,660
DIAL2-17	16,017	25,852	26,434
DIAL2-19	43,587	15,507	36,422
BACKGROUND	10,030	22,233	13,484
BASE	MAE		
DIAL2-16	5.5054	2.681	10.8398
DIAL2-17	6.7343	3.0231	7.0766
DIAL2-19	4.1127	3.5136	9.2344

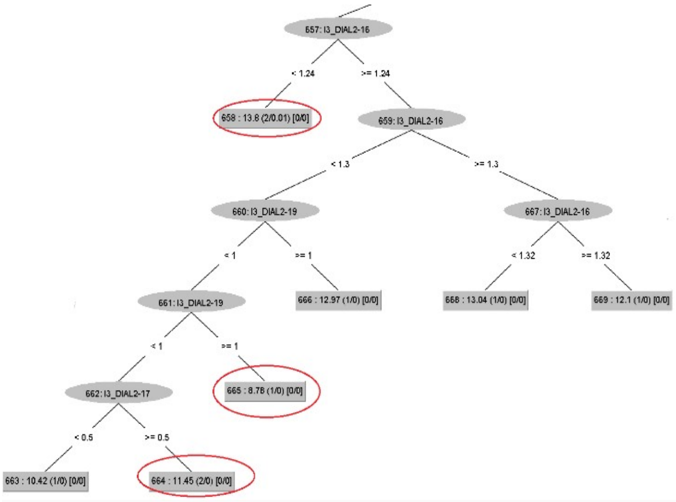
Source: Authors, (2022).

Table 11 presents the impact factor of the 3rd harmonic, through the mean absolute error - MAE (given in percentage) of the three phases for each feeder, including the background which are the other feeders not measured.



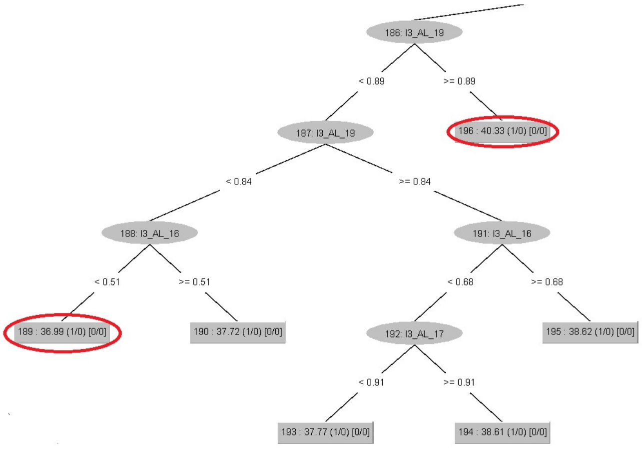
Source: Authors, (2022).

Figure 22. Right branch coming from the root node of the regression tree of case study 1 phase A



Source: Authors, (2022).

Figure 23. Right branch coming from the root node of the regression tree of case study 1 phase B



Source: Authors, (2022).

Figure 24. Right branch coming from the root node of the regression tree of case study 1 phase C

In case study 1, feeder DIAL2-19 contributes with impact factor in the 3rd harmonic for phase A with 43.587%. Feeder DIAL2-16 contributes with impact factor 36.407% for phase B. For phase C feeder DIAL2-17 contributes with an impact factor of 26.434%. Because the decision trees created by the data mining software are quite large, five subtrees were chosen for each phase of each feeder of the substation under study, which are capable of generalizing the knowledge acquired with the use of the technique employed in the evaluation of data for 3rd harmonic order of phases A, B and C of feeders DIAL2-16, DIAL2-17 and DIAL2-19. Figures 22, 23 and 24 show the five subtrees for each phase of each feeder. These were the feeders which most contributed to the increase of harmonic voltage distortion at DIBR2-O3 bar for 3rd harmonic order at phases A, B and C (DIAL2-19, DIAL2-16 and DIAL2-17, respectively). This result enables better decision making in power system management, i.e., directing actions to treat harmonic impacts caused by nonlinear loads firstly to the phases and feeders that cause greater impacts.

Comparison of Results of Linear Regression and Decision Tree Techniques

Comparison of Results of Case Study 1 (3rd harmonic): Analyzing the results of the applications of the Regression Tree and Linear Regression techniques, it can be observed in Tables 12 and Table 13 that the Regression Tree and Linear Regression techniques presented similar results with respect to the harmonic impacts (h=3) in phases A and B of DIAL2-16, phases B and C of DIAL2-17 and phases A and C DIAL2-19. The results of the application of the two techniques converge for feeders DIAL2-16, DIAL2-17 and DIAL2-19. The greatest impacts with the application of the Linear Regression technique were presented in feeders DIAL2-16 (phase A and B) and DIAL2-19 (phase A).

Table 12. Harmonic Impact Factor for 3rd order (%) MAE

BASE	PHASE A	PHASE B	PHASE C
DIAL2-16	30,366	36,407	23,660
DIAL2-17	16,017	25,852	26,434
DIAL2-19	43,587	15,507	36,422
BACKGROUND	10,030	22,233	13,484
BASE	MAE		
DIAL2-16	5.5054	2.681	10.8398
DIAL2-17	6.7343	3.0231	7.0766
DIAL2-19	4.1127	3.5136	9.2344

Fonte: Autores, (2022).

Table 13. Loading plateau, determination coefficient and impact factor of DIAL2 - 16, DIAL2 - 17 and DIAL2 - 19

Load Levelling - Morning and Afternoon		**Load Levelling - Morning and Afternoon**	
START TIME		END TIME	
00:00:00		14:00:00	
R^2			
POWER FEEDERS	PHASE A	PHASE B	PHASE C
DIAL2 – 16	00,526	00,385	00,043
DIAL2 – 17	00,009	00,165	00,561
DIAL2 – 19	00,755	00,009	00,438
Fator de Impacto (%)			
BASE	PHASE A	PHASE B	PHASE C
DIAL2 – 16	37,225	104,527	12,367
VBG-16	62,775	-04,527	87,633
DIAL2 – 17	05,784	41,570	28,812
VBG-17	94,216	58,430	71,188
DIAL2 – 19	59,494	10,680	73,474
VBG-19	40,506	89,320	26,526

Fonte: Authors, (2022).

CONCLUSIONS

Using simple linear regression and decision tree techniques it was accomplished a comparative study in the analysis of the impacts caused by harmonic currents in three feeders in a bus of an electric system. The analyses were performed through field measurement campaigns for a period of 7 calendar days, according to module 8 of PRODIST 2021. With the data collected from this measurement campaign it was possible to make a correlation analysis between the harmonic currents injected into the feeders and the voltage distortion at the busbar of the electrical system under study, through the simple linear regression and decision tree techniques. With these analyses it was possible to identify which of the two techniques used in the analyses obtained better performance in the application, and to create a profile of the feeders DIAL2-16, DIAL2-17 and DIAL2-19, in order to mitigate the harmonic distortion of voltage in the bar DIBR2-03 in study, caused by the impacts of the harmonic currents of these feeders. The technique that obtained the best performance in the analysis of the harmonic impacts on the harmonic voltage distortion in the DIBR2-03 bar under study was the Simple Linear Regression technique. Therefore, this research presented and applied in practice with case studies and actions for analysis of harmonic impacts in electrical power distribution systems through the construction of mathematical models using Simple Linear Regression analysis and Decision Tree analysis, obtaining important results in the studies performed, thus demonstrating the efficiency of the application of techniques in analysis of harmonic impacts in electrical power distribution systems.

Agradecimentos : To the Institute of Technology and Education Galileo from Amazonia (ITEGAM), Postgraduate Master in Engineering, Process Management, Systems and Environmental (EPMSE/ITEGAM). Manaus – Amazonas, Brazil to and Wasion Amazonia Industry of Electronic InstrumentsLtda. Law 6.008/1991 with resources of RD &I Project(SUFRAMA/CAPDA)finance and support the research.

REFERENCES

Abdulveleev, I. R.; Khramshin, T. R.; Kornilov, G. P.; Abdulveleeva, R. R. Experimental Study of the Impact of a DC Electric Arc Furnace on a Power Grid. 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), IEEE, Sochi, Russia, 2021.

Agência Nacional De Energia Elétrica–ANEEL. Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional–PRODIST: Módulo 8–Qualidade da Energia Elétrica. 2018.

Ahmed, Sajjad; Kashif, Syed Abdul Rahman; Ain, Noor Ul, Rasool Akhtar; Shahid, Muhammad Sohaib; Padmanaban, Sanjeevikumar; Ozsoy, Emreand Saqib, Muhammad Asghar. Mitigation of Complex Non-Linear Dynamic Effects in Multiple Output Cascaded DC-DC Converters, IEEE, 2021.

Almeida, Rivanildo Duarte; Leite, Jandecy Cabral. Análise de Impactos Harmônicos em Redes Elétricas de Média Tensão. 1. ed., v. 2. 125p, Novas EdiçõesAcadêmicas, Alemanha, 2018.

Amoo, Abdullahi Lanre, Aliyu, Usman O. And Bakare, Ganiyu Ayinde. Compendium of Computational Tools for Power Systems Harmonic Analysis, 2018.

Arrillaga, J. *et al. Power System Harmonic Analysis*. University of Canterbury, Christchurch, New Zealand, 1997.

Baggini, A. *Handbook of Power Quality*. John Wiley & Sons, 2008. 618p.

Bida, Ikram, Aouat, Saliha.Swarm Intelligence-Based Decision Trees Induction For Classification – A Brief Analysis, Department of Computer Science University of Science and Technology Houari Boumediene Bab-Ezzouar - Laboratory for Research in Artificial Intelligence (LRIA), 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being Algiers, Algeria, 2020.

Chen, Guangwu; Zhu, Yuhan; Wiedmann, Thomas; Yao, Lina; Xu, Lixiao; Wang, Yafei. Urban-rural disparities of household energy requirements and influence factors in China: Classification tree models. School of Statistics and Institute of National Accounts, Beijing Normal University, Beijing 100875, China. Sustainability Assessment Program (SAP), School of Civil and Environmental Engineering, UNSW Sydney, New South Wales 2052, Australia

Centre for Integrated Sustainability Analysis, School of Physics, The University of Sydney, New South Wales 2006, Australia. School of Computer Science and Engineering, UNSW Sydney, New South Wales 2052, Australia, 2019.

Da Silveira, Lucas C.; Bernardon, Daniel P.; Raduns, Caroline. Elaboration of an impact study on the medium voltage electrical system caused by nonlinear loads. In: 2018 Simposio Brasileiro de Sistemas Eletricos (SBSE). IEEE, 2018. p. 1-6.

Da Silveira, Lucas C.; Bernardon, Daniel P.; Raduns, Caroline. Elaboration of an impact study on the medium voltage electrical system caused by nonlinear loads. In: 2018 Simposio Brasileiro de SistemasEletricos (SBSE). IEEE, 2018. p. 1-6.

De Oliveira, Edson Farias. Avaliação da distorção harmônica total de tensão no ponto de acoplamento comum industrial usando o processo kdd baseado em medição, Instituto de Tecnologia Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal do Pará – (UFPA), 2018.

Dobref, Vasile, Popov, Petrica, Mocanu, Vlad. The Quality of Electricity Assessment in Marine Power Systems, Marine Engineering Faculty, Naval Academy "Mircea celBtrân" Constana, Romania, 2018.

Dugan, R.; Granaghan, M. et al.Electrical Power Systems Quality, McGraw-Hill. ISBN 0-07-018031-8 1996, 2004.

IEEE Std 1453-2015. IEEE Recommended Practice for the Analysis of Fluctuating Installations on Power Systems, 2015.

IEEE Std 519™-2014. IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems, 2014.

IEEE-1159. IEEE recommended practice for monitoring electric power quality. IEEE Std 1159-2019, v. 99, p. 1–98, 2019.

JR, D. S., E Simonetti, D. S. Análise harmônica e inter-harmônica de um forno elétrico a arco. IEEE/IAS International Conferenceon Industry Applications. Acesso em 21 de janeiro de 2020. Disponível em https://scihub.tw/ https://www.researchgate.net/publication/224226659_Analise_harmonica_e_inter-harmonica_de_um_Forno_Eletrico_a_Arco.

Kamenka, A. Six tough topics about harmonic distortion and Power Quality indices in electric power systems. A White Paper of Schaffner Group, 2014.

Lei, Hong; Huang, Ya; Jiang, Li; Gao, Ge; Zhang, Jie; Wu, Peng.Research on Fourier and Prony analysis algorithm of inter-harmonic in power system. 2021 The International Conference on Power Engineering (ICPE 2021), December 9–11, 2021,Nanning, China, 2022.

Leite, Jandecy Cabral. Projeto multicritério de filtros harmônicos passivos para instalações industriais utilizando técnicas de inteligência computacional. Pará: UFPA, 2013. Tese (Doutorado em Engenharia Elétrica). Universidade Federal do Pará. Instituto de Tecnologia. Programa de pós-graduação em Engenharia Elétrica, Belém- Pará, 2013.

Li, Lianpeng; Dong, Jian; Zuo, Decheng; Wu, Jin. SLA-aware and Energy-efficient VM Consolidation in Cloud Data Centers Using Robust Linear Regression Prediction Model, Department of Computer Science and Technology, Harbin Institute of Technology, No.92, Xidazhi Street, Harbin, China, 2019.

Lukasik, Zbigniew, Olczykowski, Zbigniew. Estimating the Impact of Arc Furnaces on the Quality of Power in Supply Systems. Faculty of Transport, Electrical Engineering and Computer Science, Kazimierz Pulaski, University of Technology and Humanities, Malczewskiego 29, 26-600 Radom, Poland, 2020.

Marrero, L., García-Santander, L., Hernández-Callejo, L., Bañuelos-Sánchez, P., and JARA, V. Harmonic Distortion Characterization in Groups of Distribution Networks Applying the IEEE Standard 519-2014, 2021.

Matloff, N. *Statistical Regression and Classification - From Linear Models to Machine Learning*. University of California, Davis, USA, 2017.

Murray, Ruan ; De Kock, Jan. Instrument transformers influence on harmonic measurements for grid code compliance, 2018 IEEE 4th Global Electromagnetic Compatibility Conference (GEMCCON), 2018.

Nassar, Samira Reda; Eisa, Ayman A.; Saleh, A. A.; Farahat, M.A. And Abdel-Gawad, Amal F. Evaluating the Impact of Connected Non-Linear Loads on Power Quality Nuclear Reactor case study, Journal of Radiation Research and Applied Sciences, (2020). https://doi.org/10.1080/16878507.2020.1828018.

Nazirov, Khurshed B.; Ganiev, Zokirjon S. Experimental Evaluation and Analysis of Electric Power Quality in Electric Networks MunicipalHouseholds, Electric Power Industry Department, Electric Power Industry Department, Branch of the National Research University, 2021.

Nazirov, Khurshed B.; Ganiev, Zokirjon S. Experimental Evaluation and Analysis of Electric Power Quality in Electric Networks MunicipalHouseholds, Electric Power Industry Department, Electric Power Industry Department, Branch of the National Research University

Nikolaev A.A., T. P. The comparative analysis of electrode controle systems of electric arc furnaces and ladle furnaces. International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), pages 1-7, 2016. doi:10.1109/ICIEAM.2016.7910888.

PROCOBRE, S. Qualidade de Energia - Harmonicas. Workshop Instalações. Elétricas de Baixa Tensão. 2003.

Rawlings, J. O., Pantula, S. G., Dickey, D. A. *Applied Regression Analysis: A Research Tool,* Second Edition. Department of Statistics - North Carolina State University, USA, 2001.

Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Global Environmental Change and Earth Observation Research Group, Geography and Environment, University of Southampton, Southampton SO17 1BJ, United Kingdom. Department of Haematology, Wellcome Trust and MRC Cambridge Stem Cell Institute and Cambridge Institute for Medical Research, University of Cambridge, Cambridge CB2 0XY, United Kingdom. Departamento de Geodinámica, Universidad de Granada, 18071 Granada, Spain. Departamento de Análisis Matemático, Universidad de Granada, 18071 Granada, Spain, 2015.

Salam, S M, Uddin, M. Iftekher, MOINUDDIN, M. Rifat Bin. Impact Analysis of Large Number of Non-Linear Lighting Loads on Power Quality in Distribution Network. Dept. of EEE Southern University Bangladesh Chattagram, 2021.

Sedkaoui, Soraya And Khelfaoui, Mounia. Linear Regression, Published by ISTE Ltd and John Wiley & Sons, Inc, IEEE, 2020.

Sichani, M. H., Khalafinejad, S. *Regression, Knowledge-Oriented Applications in Data Mining.*Sharif University of Technology, Iran, 2011.

Souza, Jayne Do Nascimento, Leite, Jandecy Cabral, Almeida, Rivanildo Duarte.Analysis of Harmonic Impacts on an Electric System Bus Using Artificial Neural Networks. Post-Graduation Program in Process Engineering of the Institute of Technology of the Federal University of Pará, Federal University of Pará (UFPA), Belém-Pará, Brazil, Institute of Technology and Education Galileo of the Amazon (ITEGAM), Manaus-Amazonas, Brazil, 2021.

Ullah, Asad; Sheikh, Inam Ul Hasan; Arshad, Shahzad; Saleem, Faisal. Digital Active Power Filter Controller Design for Current Harmonics in Power System, Department of Electrical Engineering, University of Engineering and Technology, Taxila Taxila, Pakistan, 2019.

Wang, Xiongfei; Blaabjerg, Frede. Harmonic Stability in Power Electronic Based Power Systems: Concept, Modeling, and Analysis, IEEE, 2018.

Welham, S. J., Gezan, S. A, Clark, S. J., Mead, A. *Statistical Methods in Biology - Design and Analysis of Experiments and Regression.* CRC Press, Taylor & Francis Group, 2015.

Yasar, Seyma; Arslan, A. Kadir; Çolak, Cemil; Yologlu, Saim. A Developed web-based software can easily fulfill the assumptions of correlation, classification and regression tasks in data processing, Dept. of Biostatistics and Medical Informatics, Faculty of Medicine, Inonu University Malatya/TURKEY, Published in: International Artificial Intelligence and Data Processing Symposium (IDAP), IEEE, 2019.

Yu, Junho; Jeon, Seungki; Choi, Hoyong; Kim, Namsoo. CMOS Integrated PFM DC-DC Converter with Digitally-Controlled Frequency Selector, School of ECE, Chungbuk National University Cheong-ju, chungbuk, 361-763 Korea, 2018.

Zobaa, Ahmed F. And Aleem, Shady H. E. Abdel. Power Quality in Future Electrical Power Systems, Published by The Institution of Engineering and Technology, London, United Kingdom, 2017.

Zobaa, Ahmed F., Aleem, Shady H.E. Abdel And Balci, Murat E. Introductory Chapter: Power System Harmonics—Analysis, Effects, and Mitigation Solutions for Power Quality Improvement, 2018).

****