



Article

Using True RMS Current Measurements to Estimate Harmonic Impacts of Multiple Nonlinear Loads in Electric Distribution Grids

Flávia P. Monteiro ^{1,2,*}, Suzane A. Monteiro ^{1,2}, Maria E. Tostes ² and Ubiratan H. Bezerra ²

¹ Campus Oriximiná, Federal University of Western Para, Oriximiná, PA 68270000, Brazil; 1su.aquino@gmail.com

² Electrical Engineering Department, Federal University of Para, Belém, PA 66075-110, Brazil; tostes@ufpa.br (M.E.T.); bira@ufpa.br (U.H.B.)

* Correspondence: flaviamonteiro.comp@gmail.com; Tel.: +55-919-8215-1484

Received: 16 October 2019; Accepted: 25 October 2019; Published: 30 October 2019



Abstract: Currently, for analyzing harmonic impacts on voltage at a point of interest, due to multiple nonlinear loads, the literature recommends carrying out simultaneous and synchronized measurement campaigns in all suspicious points with the use of high cost energy quality analyzers that are usually not available at the customers' facilities and very often also not at the electric utilities. To overcome this drawback this paper proposes a method of assessing the harmonic impact due to multiple nonlinear loads on the total voltage harmonic distortion using only the load current true RMS values which are already available in all customers' installations. The proposed methodology is based on Regression Tree technique using the Permutation Importance indicator which is validated in two case studies using two different electrical systems. The first case study is to ratify the use of Permutation Importance to measure the impact factor of each nonlinear load in a controlled scenario, the IEEE-13 bus test system, using ATP simulation (Alternative Transient Program). The second is to apply the methodology to a real system, an Advanced Measurement Infrastructure System (AMI) implanted on a campus of a Brazilian University, using low cost meters with only true RMS current measurements. The results achieved demonstrated the feasibility of applying the proposed methodology in real electric systems without the need for additional investments in high-cost energy quality analyzers.

Keywords: current true RMS; harmonic distortion contribution; machine learning; power system analysis computing; total voltage harmonic distortion

1. Introduction

With the advance of technology, a problem has become quite frequent in distribution systems: harmonic distortion. This is due to the increasing use of nonlinear loads, mainly in the form of power electronics or consumers' electronics equipment. The first records on harmonic distortion were in the 1920s, when the main effects were on rotating machines [1]. It is noticed a great influence of harmonic distortion in the voltage and current's waveforms, being responsible for harmful effects at various levels in distribution networks.

Heating in distribution transformers, malfunction of protection and overload of neutral conductors are some of the implications on equipment most sensitive to these variations, directly affecting the operation of the system and substantially increasing the likelihood of failure occurrences and interruptions in electricity supply [2]. The electric network itself is a source of harmonic distortions in voltage and current due to its operation characteristics but loads are the main responsible for this type of phenomenon [3,4].

However, the electric energy as a product needs to be delivered at certain quality levels, and the harmonic distortion, as well as other factors impacting the power quality need to be regulated, forcing distribution utilities to continuously improve their services. In Brazil, for example, ANEEL, the Brazilian Electricity Regulatory Agency, establishes the main power quality indicators to be met by utilities and the respective penalties for non-compliance with these limits [5].

Therefore, electric utilities need to continuously monitor their electrical grids to detect any suspicious loads that could be contributing to the voltage harmonic distortion above specified limits as recommended by power quality standards. Therefore, the development of procedures to identify which customers' loads are more impacting to the voltage harmonic distortion observed at specific locations in the grid are relevant and can help to implement a differentiated treatment to these customers aiming at taking remedial actions to mitigate the possible harmonic distortion transgressions.

The point is that monitoring all customers who may be heavily contributing to the voltage harmonic distortion at a specific point in the electric network is not a trivial task and would involve a simultaneous measurement campaign in all these customers, resulting in a high financial investment in power quality analyzers.

In the current scenario of distribution systems, consumers rarely have power quality analyzers in their electrical installations. In this way the implementation of simultaneous measurement campaigns to assess harmonic distortion parameters becomes infeasible. However, all-consuming facilities have power measurement, and consequently, electric voltage and current rms measurements, which is typical in a current Advanced Metering Infrastructure system (AMI) [6].

In an AMI, Smart Meters (SMs) measure the major electrical quantities of the network at regular time intervals and make them available online through a central Meter Data Management System (MDMS). AMI can be used on any power system if coupled to a Power Line Communication (PLC) module, see for example [7–9].

Considering this current scenario, this paper proposes a low-cost methodology capable of identifying and quantifying the impact of multiple nonlinear loads on the voltage total harmonic distortion at a point of interest in the electric grid using only the True RMS current measured at the customers' electric installations.

This new perspective is quite attractive in terms of cost, as it dramatically reduces it by using more affordable meters for utility's reality, but still allows for even greater gain when coupled with an Advanced Metering Infrastructure, enabling real-time power quality monitoring. To achieve this, the Gradient Boosting Regression Tree (GBRT) technique was applied using the permutation importance, an inherent characteristic of regression tree techniques to determine the impact of multiple nonlinear loads in the voltage total harmonic distortion (THD_u) at the point of interest.

The proposed methodology will be tested and validated using the two steps procedure as described in the following:

- (1) Validation using the permutation importance as a metric to measure the impact factor for each nonlinear load by comparison with the result obtained by ATP software in a controlled simulation scheme using the IEEE-13 bus system [10].
- (2) Applying the methodology to the AMI system installed at the Federal University of Pará campus located in the north of Brazil, so that it is possible to indicate the most representative nonlinear loads in impacting the THD_u at the Common Coupling Point (CCP) with the utility electric grid. The result of this test case was also compared to the methodology presented by [10].

2. Related Works

Xu and Liu [11] presented a method for determining the consumer's and the electric utility's harmonic contribution at the common coupling point (CCP) based on a harmonic Norton equivalent circuit capable of separating harmonic current and voltage into two components: one due to consumer and the other due to the power grid.

Following similar methods as the one used in [11], other studies based on harmonic Norton equivalent circuits emerged, as in [12] which has innovated by offering the possibility to evaluate the relative contribution of several customers at the CCP. The main disadvantage of such methodologies is the difficulty of application in systems where there are multiple nonlinear loads constantly changing, making it difficult to establish an adequate and accurate equivalent circuit.

Wang, Nino and Xu [13] presented a measurement method that can determine the source and harmonic impedance for residential and commercial systems supplied by a single-phase transformer, showing the utility's and consumer's contribution at the CCP through a short circuit controlled by a thyristor at the measuring point. The biggest problem with this technique is the need to insert a thyristor in the measurement arrangement which represents an unnecessary additional cost compared to measurement systems already available in the market.

Taking advantage of evolutionary computing, Srinivasan et al. [14] used trained neural networks to extract waveform patterns from the input current to identify harmonic sources and the various types of devices connected to the grid using their uniquely distinct harmonic "signature". The main disadvantage of this method is its inability to provide information about multiple nonlinear loads in the system, such as location, and the application in large power systems.

Mazumdar et al. also conducted studies with recurrent neural networks [15–17], seeking to minimize conflicts between utilities and customers by analyzing the responsibility for harmonic distortions at the CCP. However, the verification is not done in real time, but only when problems related to the responsibility for harmonic contribution in the electric network occur.

Kandev and Chénard [18] present a method capable of locating in real time the consumer responsible for the harmonic disturbance in the grid by vector analysis of harmonic currents measured at different points of the electrical network. Although it has the advantage of real-time analysis, only one harmonic current is analyzed for each experiment.

Manito in [10] presented an artificial neural network (ANN) model to estimate the impacts of nonlinear loads on the voltage harmonic distortion at a point of interest. His method relies on a sensitivity analysis to establish the percentage contribution (impact factor) of load currents on the observed individual voltage harmonic distortion. The study was conducted in an industrial test system with the alternative transient program (ATP). The proposed methodology was able to correctly classify the impact degree of nonlinear load currents on the 5th harmonic voltage distortion at the points of interest, however, only one specific harmonic at a time.

Utilizing a statistical vision, Yin et al. [19] explored statistics tools that provide a solution for identifying sources of harmonic distortions in electric networks with multiple nonlinear loads. Methods with multiple linear regression and partial correlation analysis were proposed. However, this article only determined their contributions to individual harmonic distortions.

Matos et al. [20] presented a method for finding the harmonic impact of multiple nonlinear loads using measurements of individual harmonic voltage and current distortions by linear and nonparametric regression models. Two electrical networks were used for tests and in both the nonparametric models obtained better results.

Following the premise of evaluating harmonic distortion impacts due to multiple loads comparing the performance of evolutionary computation techniques and statistical methods, [21] compares the techniques of multiple linear regression, artificial neural networks and regression tree. This work considered daily and weekly time windows, also analyzing the load curve for different loading levels. However, just like the work previously mentioned, this research evaluates only one harmonic order at a time.

Moradifar et al. [22] presented a new approach to the problem by applying fuzzy logic to determine the optimal allocation of harmonic meters for intelligent detection of multiple nonlinear loads in distribution systems. In addition, the fuzzy algorithm estimates the location and relative level of harmonic sources by investigating the power magnitude and signal considering the harmonic distortions and network reactance.

Looking at the mentioned previous works the main contributions of this article can be summarized as follows:

- (1) The methodology uses as input the true RMS current, an electric variable that can be easily measured by low-cost meters, making viable for the reality of both distribution utilities and customers. The use of true RMS current in identifying and quantifying the impact of multiple nonlinear loads on voltage total harmonic distortion levels at the CCP as proposed in this article is a novelty.
- (2) The methodology can be applied with other input and output variables as long as they are representative to the scope of the problem. True RMS current was chosen because it is the most accessible magnitude with high correlation with voltage total harmonic distortion levels at the point of interest;
- (3) It presents an innovative and insightful methodology for solving the problem of identifying and quantifying the impact of multiple loads on the THDu level per phase at the CCP through a metric which is inherent to the regression tree technique. As it is already a step of the computational technique chosen, the estimation of input variables importance in the output variable does not add computational cost.

The methodology, when applied to an AMI system, offers further contributions:

- (1) Understanding that Advanced Metering Infrastructure is a relatively new definition, this article presents some procedures to ensure data reliability during its implementation;
- (2) Provides real-time power quality monitoring in all phases of the distribution grid, identifying meters with big contribution to total harmonic distortion;
- (3) Allows a wide range of experiments with varying time windows. According to the case to be studied by the grid managers, the combination of the proposed methodology with an AMI is able to investigate the meters contributions at different times per day, week or month, investigating the most impacting loads on the grid power quality seasonally.

3. Proposed Methodology

The problem to be solved is illustrated schematically as shown in Figure 1, representing an electric network having m nonlinear loads for which it is desired to calculate how these nonlinear loads are impacting the voltage total harmonic distortion at bus X , $THDu_x$. To reach this objective a GBRT model will be used having the nonlinear loads electric currents as input variables to calculate the harmonic impacts on the total harmonic distortion at a point of interest. To describe the problem mathematically, consider the input matrix as X and the output vector as Y .

Matrix X contains the time series of measured true RMS currents, $I_{RMS_j}(t)$, where $j = 1, 2, 3, \dots, m$, been m the number of nonlinear loads considered for a given time series t . And vector Y is composed of voltage total harmonic distortion values measured at the common coupling point X , $THDu_x$ simultaneously measured with nonlinear loads currents presented in matrix X . Matrix X and vector Y structures are shown in Equations (1) and (2), respectively.

$$X = \begin{bmatrix} I_{RMS_1}(1) & I_{RMS_2}(1) & I_{RMS_3}(1) & \cdots & I_{RMS_m}(1) \\ I_{RMS_1}(2) & I_{RMS_2}(2) & I_{RMS_3}(2) & \cdots & I_{RMS_m}(2) \\ I_{RMS_1}(3) & I_{RMS_2}(3) & I_{RMS_3}(3) & \cdots & I_{RMS_m}(3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_{RMS_1}(t) & I_{RMS_2}(t) & I_{RMS_3}(t) & \cdots & I_{RMS_m}(t) \end{bmatrix} \quad (1)$$

$$Y = \left[THD_X(1) \quad THD_X(2) \quad THD_X(3) \quad \cdots \quad THD_X(t) \right] \quad (2)$$

The proposed methodology is based on the principle of supervised learning for the estimation of a time series, where the quality of this prediction is reflected in how the inputs to build the model are

representative. This analysis is performed by estimating a THD_x within a period t based on a set of input attributes, $I_{RMS,m}$, measured simultaneously according to a specified sampling rate. However, nothing prevents the use of this methodology with other input and output variables, if the high correlation between them is proven.

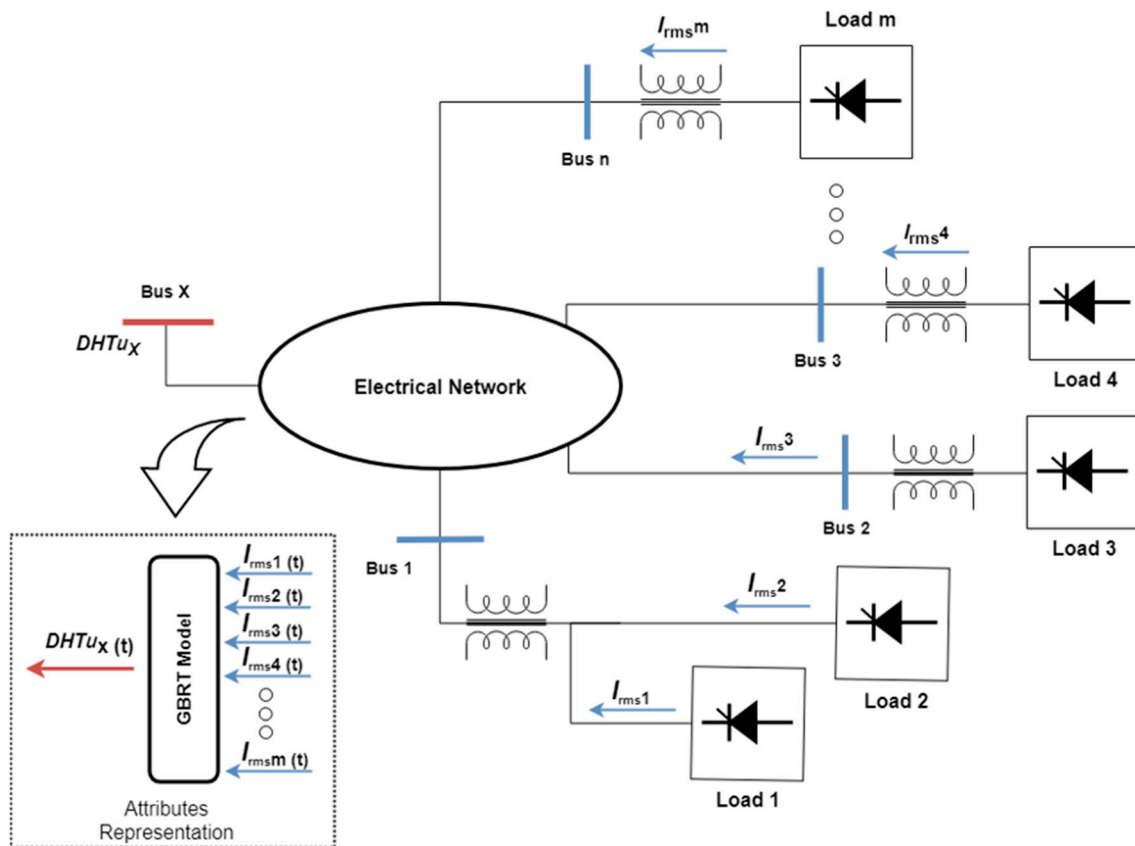


Figure 1. Characterization of the problem of identifying the contribution of multiple harmonic sources in THDu at a point of interest in the power grid. Adapted from [23].

Commonly, machine learning problems apply the attribute selection procedure to combine and identify the most significant variables in a dataset, acting to improve predictive performance and reducing its learning curve. There are techniques that already implement this procedure on its own structure, as is the case of Regression Tree. The regression tree comprises an input X and output Y , being formed by n nodes, subsets of X . The nodes are subjected to binary tests, dividing into new two subsets of X which in turn are labeled with a better value for the output variable, decreasing the degree of impurity with each step, adopting a voracious approach, top-down, recursive, under the divide-to-conquer strategy. Thus, the tree will continue to form nodes until they become pure in terms of Y or when all variables X_j are locally constant.

The importance of an input variable X_m is then evaluated in predicting the output by adding the weighted impurities decreases for all nodes where X_m is used. Thus, using the weighted impurity function to measure the Impact Factor (IF) of each harmonic generating source at a point of interest, is an innovative aspect introduced in this article, not found yet in other references in the technical literature.

Based on this problem formulation, this paper proposes the use of the Gradient Boosting Regression Tree (GBRT) technique [24–26], together with the impurity-based function proposed by [27,28], known as Mean Decrease Accuracy (MDA) or Permutation Importance, to get the impact factor. For the impurity measure $i(n)$ it can be also used the Gini index, the Shannon entropy, the Y variation, among others [27,28]. Generally, (3) can be referred to as the Mean Decrease Impurity (MDI) importance.

And Permutation Importance is an MDI that weighs the importance of a X_m variable in the output when its values are randomly exchanged in out-of-bag samples:

$$Imp(X_m) = \frac{1}{N_{TR}} \sum_{N_{TR}} \sum_{n \in N_{TR}:v(s_n)=X_m} p(n) \Delta i(s_n, n) \quad (3)$$

In (3) N_{TR} represents the number of trees built, $p(n)$ is the proportion of samples reached on n nodes, s_n is the binary test for separating samples in subsets, $v(s_n)$ is the variable used as the separation parameter and $\Delta i(s_n, n)$ is the variation of a measure of impurity.

Finally, the relative contribution of each nonlinear load to the voltage total harmonic distortion at the model output is measured by the impact factor, $IF_{THDu_m}(\%)$, according to (4):

$$IF_{THDu_m}(\%) = \frac{Imp(I_{RMS_m})}{\sum_{j=1}^m Imp(I_{RMS_j})} \quad (4)$$

The impact factor of each j load can be interpreted as a relative percentage value that is calculated for each true RMS current individually for the set of nonlinear loads considered in the analysis, leading to (5):

$$IF_{THDu_1} + IF_{THDu_2} + \dots + IF_{THDu_m} = 100\% \quad (5)$$

An overview of the methodology proposed in this article is presented in Figure 2. As a starting point, it is important to identify the data sources. The methodology points out two possible ways: the first corresponds to the application in an AMI system following procedures as will be indicated in topic A; the second not applying in an AMI system, that is, using traditional measurement campaigns.

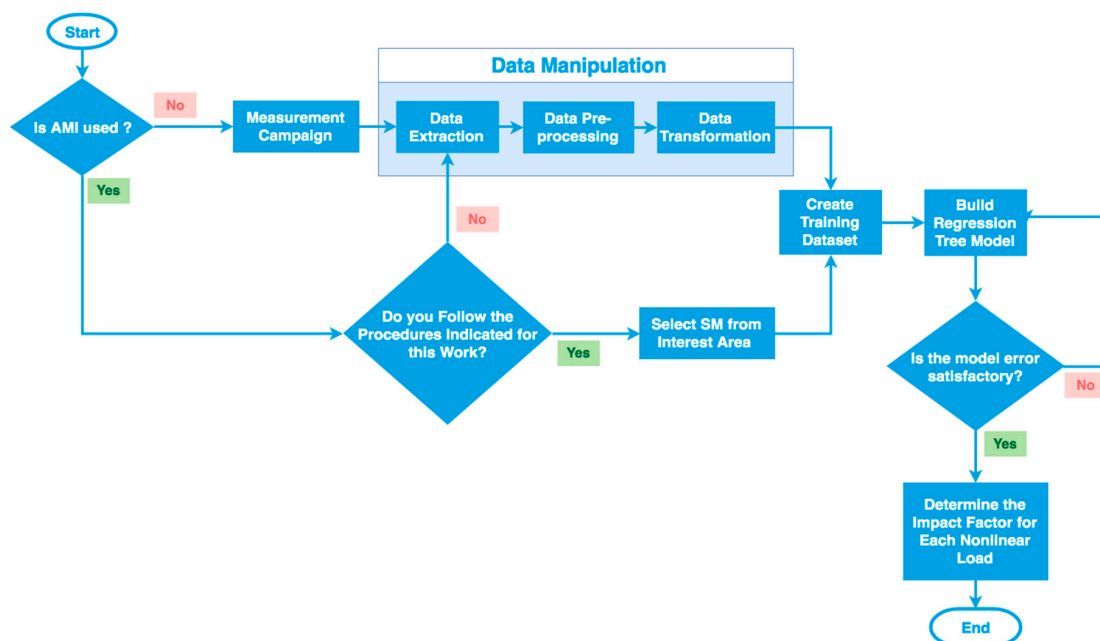


Figure 2. Overview of the proposed methodology.

If measurement campaigns are to be used without the resources of a Data Management System, a time window must be chosen having a sampling interval. For power quality applications a maximum sampling interval of 10 min is recommended by legislation for one-week campaigns, which correspond to 1008 samples [5]. The smaller the sampling rate the greater the possibility of applying data preprocessing techniques, which corresponds to the next step in the methodology.

After the data acquisition comes a data management step, which corresponds basically to the application of process of Knowledge Discovery in the Database (KDD), which is an essential step to obtain a good estimation model.

In the Data Extraction step, it is important to apply a filter to select which measurement points should be considered in the study, depending on how many null samples exist in the time series. Points with more than 30% of null samples must be discarded from the diagnosis. These null samples usually correspond to periods of meter shutdowns or failures. This same procedure is applied in the SM selection step for the interest area.

According to the proposed methodology the next step is the dataset creation and separation that will be used in the training step and later in the validation step of the configured and constructed estimation model. It is advisable to use the cross-validation technique to mitigate effects caused by seasonality which are inherent in electric loads modeling.

If the modelling error is not satisfactory, new parameters configuration must be selected to raise the desired performance. Otherwise, it is recommended to restart the methodology with a greater focus on the Data Manipulation step. When a satisfactory estimation model is obtained the impact factor for each nonlinear load is calculated accordingly.

Procedures for Implementing the Methodology in an Advanced Metering Infrastructure

As previously presented, this work also contributes to the application of the proposed methodology in a real scenario with an already implemented and functioning AMI system. In this section it will be outlined the main steps taken to ensure that data acquisition from the meters be reliable for usage in the methodology.

To achieve that, it begins by highlighting important technical factors for the basic elements that make up the infrastructure required for this application:

- End user devices-Smart Meters;
- Communication;
- Meters Data Management System (MDMS).

In general, smart meters need to be able to send stored data and receive operational commands. Thus, it is crucial a SM to have at least one network communication interface and an internal real-time clock to allow synchronism checking, and in the case different meters are used, it must be possible to standardize the time measurement.

In [6] several recommendations for the implementation of an AMI are presented, but as one of the main motivations of this work is cost reduction so the methodology can be accessible, the efforts were devoted to the development of an application with the function of mediation between meters, database and end users. In other words, routines and mechanisms for communication and data integrity checking were proposed to implement the infrastructure Data Management System.

The application was divided into three steps: data acquisition; data processing; and data persistence. Each step must be associated with an error and exception release. The first is related to the communication with the meters, being responsible for verifying possible changes in the physical network and data acquisition. The essential procedures for this implementation are:

- Flag communication failures;
- Scan network to check if meter is connected or disconnected to system;
- Identify and flag meter configuration changes;
- Request data and apply checksums CRC (Cyclic Redundancy Check).

Cyclic Redundancy Check is an error detection methodology that aims to identify changes in the data chain during information transmission [29]. It basically consists on the polynomial calculation of bytes check, usually called checksum, which are added to the original message for later transmission.

The data acquired during the Data Acquisition step correspond to measurements provided almost instantaneously by the SMs, with requests in every 30 s. Since the purpose of this application is to measure steady state quantities and the meters preferentially work with RMS current values, the interval between data acquisition doesn't need to be smaller than one cycle of the input sinusoid. Thus, the request period can happen in seconds, allowing time for repetition of the data acquisition process in case of errors or inconsistencies.

Such errors need to be detected and corrected so that persisted data in the database are reliable for the estimation model construction step. It is also very important to store the history of faults and errors, as they will be used to filter meters that will participate in the diagnosis. Table 1 shows how the codes associated with each error detected by the AMI system of the test scenario are stored, which it will be presented in the next section.

Table 1. Error table that can be identified in the Data Acquisition step.

Error	Code
Without Error	0
Communication Error	1
CRC Error	2
NaN Error	3
Date Error	4

The second step of the proposed application is Data Processing. The measurement treatment process consists of obtaining a representative value of this set of samples for each electrical quantity, which are I_{RMS} and THD_u in this specific case. Thus, the first step is to remove measurements with an error, in an independent way for each variable. Then the remaining samples are filtered using the Chauvenet criterion presented in the computer code Algorithm 1. Finally, the mean value is calculated, and used to represent the respective variable. If in one of the steps all measurements are removed, the value representing the variable is taken as zero, and an equivalent error is defined as the smallest value among the errors obtained for the set of variables.

Algorithm 1: CHAUVENET CRITERIA	
Input:	$x[1..n]$
Output:	$x_{mod}[1..(n-m)]$
1	Begin
2	$x_{media} \leftarrow \text{MEAN}(x)$
3	$x_{std} \leftarrow \text{STD}(x)$
4	for each x_i in x do
5	$\Delta R \leftarrow x_i - x_{mean} / x_{std}$
6	if $\Delta R < \text{CRITICALVALUE}(n)$ then
7	$x_{mod} \leftarrow \text{APPEND}(x_{mod}, x_i)$
8	end
9	end
10	end
11	return x_{mod}

The Chauvenet Criterion is widely accepted as a deviation detection technique. It defines an acceptable deviation around the mean for a sample with no measurements [30]. The base critical value can be obtained from Table 2.

The last step to address is Data Persistence, which is the step responsible for storing the handled data in the database. It is important to meet the acquisition and treatment rates, because it is an application which should be repeated periodically. These cycles start at different intervals and can be easily configured according to the scenario covered. In summary, the procedures presented in this section ensure the use of the proposed methodology in an AMI environment. It is important to note that all samples must respect a sampling rate and be in the same time window, ensuring synchronism.

Table 2. Critical Values ΔR adapted from [31].

N	AR
3	1.38
5	1.65
6	1.73
7	1.8
10	1.96
50	2.57
100	2.87
300	3.14
500	3.29
1000	3.48

4. Results and Discussion

In order to validate the proposed methodology, two tests were executed. The first one, compares the impact factor performance calculated using the GBRT technique with Permutation Importance, and the impact factor calculated using Artificial Neural Network based procedure as presented in [10] for the IEEE-13 bus industrial system. The second test is accomplished using the AMI system installed at the Federal University of Pará campus.

4.1. Permutation Importance Validation

Although permutation importance is already used successfully to evaluate the impact of input variables on output variables in several applications [32–34], it is necessary to evaluate its performance in the problem addressed in this article. As previously mentioned, in the first test the GBRT technique combined with the permutation importance metric are used to calculate the impact factor in an ATP simulation study using the balanced IEEE-13 bus industrial distribution system, and the obtained results are compared with those obtained in [10] also using the same test system.

Manito et al. (2018) in [10] used an Artificial Neural Network to estimate the impacts of fifth harmonic currents generated by the nonlinear loads on the fifth harmonic voltage at a point of interest, in a scenario simulated in ATP Draw software (Alternative Transient Program [35]).

The same scenario is used with the GBRT modeling technique with permutation importance to calculate the impact factor of fifth harmonic currents on the fifth harmonic voltage at the same point of interest as calculated in [10], comparing the results of the two techniques.

4.1.1. IEEE-13 Bus Industrial Electrical Distribution System

The IEEE-13 bus industrial distribution system [36,37], represented in Figure 3, is a three-phase system used by several authors as a test system for power quality analysis [38–40]. The same scenario of [10] was used in this study, where four nonlinear loads were specified at buses B11, B29, B49 and B51, modeled by the ATP Draw software as shown in Figure 3, configuring nonlinear loads as harmonic current sources HS1, HS2, HS3, and HS4 respectively. Importantly, ATP Draw software is widely used in power quality analysis, simulating the characteristics and the actual behavior of the network satisfactorily [41–45].

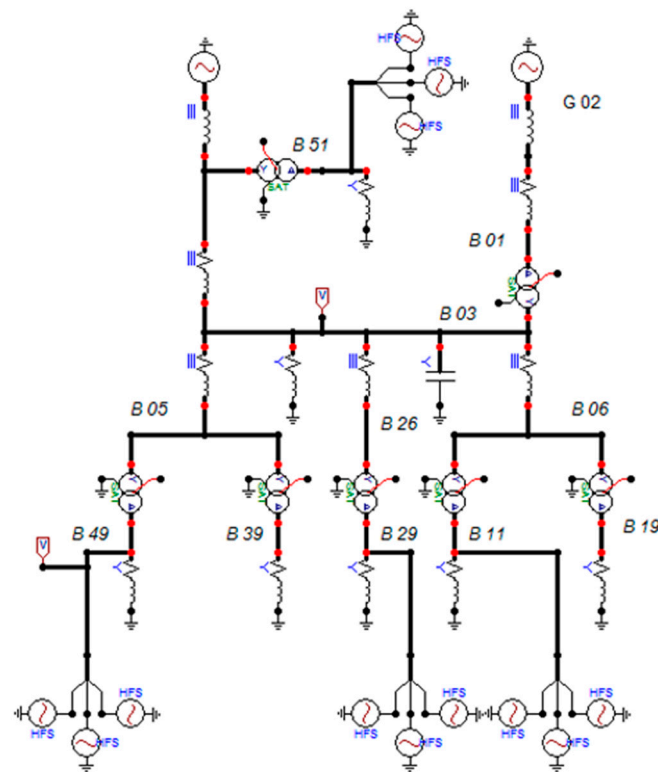


Figure 3. IEEE-13 bus industrial electrical distribution system as used in [10].

4.1.2. Data Manipulation

As described in section III, it is necessary to evaluate the scenario database focusing mainly on data preprocessing. Applying this step to the first test data, outliers were found most frequently in harmonic sources 1 and 4 being identified and removed. At the end of this step the samples were reduced from 10800 to 7427, which is a 31.23% reduction. In the next topic it will be presented the importance of this step for the calculated harmonic impact factor values shown in Table 3.

Table 3. Impact factors calculated for each nonlinear load.

Harmonic Sources	HS1 (%)	HS2 (%)	HS3 (%)	HS4 (%)
ATP Calculated Impact (Reference Values)	11.83	51.82	12.68	23.67
	Fourth	First	Third	Second
GBRT Calculated Impact	11.43	51.63	12.71	24.23
	Fourth	First	Third	Second
GBRT Calculated Impact (without pre-processing)	19.74	42.36	8.82	29.08
	Third	First	Fourth	Second
ANN Calculated Impact	13.83	52.83	12.73	20.61
	Third	First	Fourth	Second

4.1.3. Comparison between Techniques and Metrics

The work developed by Manito et al. (2018) divided the dataset, so the first 80% samples were intended for training and testing the ANN via cross-validation and the 20% remaining for the ANN model validation. To apply the proposed methodology by using GBRT, the above proportion was respected, and the cross-validation technique was also used, but with 10 folds for training the GBRT model.

The methodology was implemented in Python with the aid of the Scikit-learn library [46] and the Gradient Boosted Regression Trees regression technique, as already mentioned. After training and

construction of the GBRT model, it had its performance evaluated by the Absolute Mean Error (MAE) obtained through the test set, as well as in [10].

The best parameter setting found for the GBRT model was: Max_depth = 10, Min_sample_split = 10, Min_sample_leaf = 3, N_estimator = 50, Learning rate = 0.1, Loss functions = huber, alpha = 0.85. With this configuration the GBRT technique obtained a MAE equal to 0.0065 while ANN obtained 0.0107 in this study, so the GBRT improved the estimation performance by 39.25%. In addition, GBRT presented a satisfactory result when compared to impact factors calculated by ATP simulation, which are considered reference values for comparison, as seen in Figure 4.

It is also observed in Figure 4 that the impact factors calculated by the GBRT model are systematically closer to the exact results obtained by ATP, for all harmonic sources HS1, HS2, HS3 and HS4. Other relevant results are summarized in Table 3, for the ANN and GBRT models. While ANN and GBRT without data preprocessing have erroneously ranked third and fourth most impacting harmonic sources as compared to ATP ranking, the GBRT model with data preprocessing as proposed in the article ranked correctly all harmonic sources. This fact makes evident the importance of the data preprocessing step included in the GBRT model.

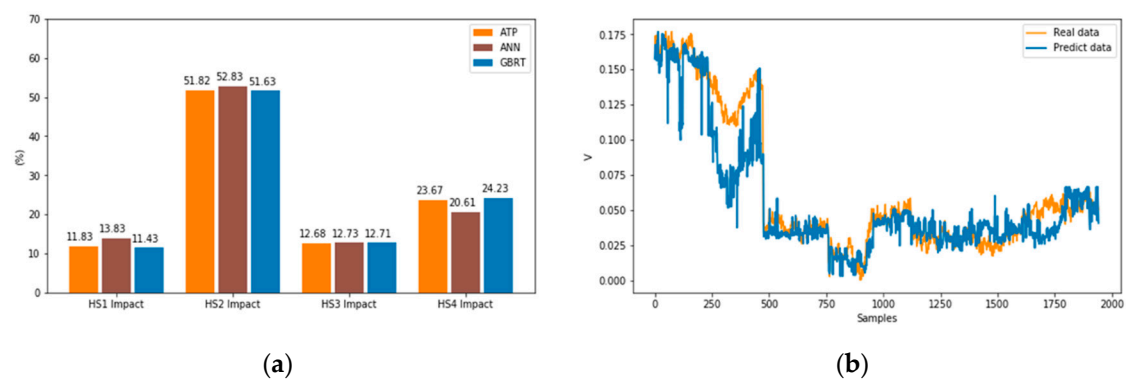


Figure 4. (a) The impact factor of each nonlinear load represented in the IEEE-13 bus system studied by evaluated techniques and (b) Comparing real data (ATP calculated fifth harmonic voltage magnitudes - in orange) at bus B03 with their estimates obtained by the GBRT in blue.

4.2. Test Case in an Advanced Metering Infrastructure Installation

After validating the use of permutation importance as a metric to calculate the impact factor, it was decided to execute the proposed methodology in a real electric distribution grid having an AMI system already implanted. This is located at the Federal University of Pará (UFPA) Campus in Belém City-Brazil, as depicted in Figure 5. The AMI system is named SISGEE that is the short form for Electric Energy Management System.

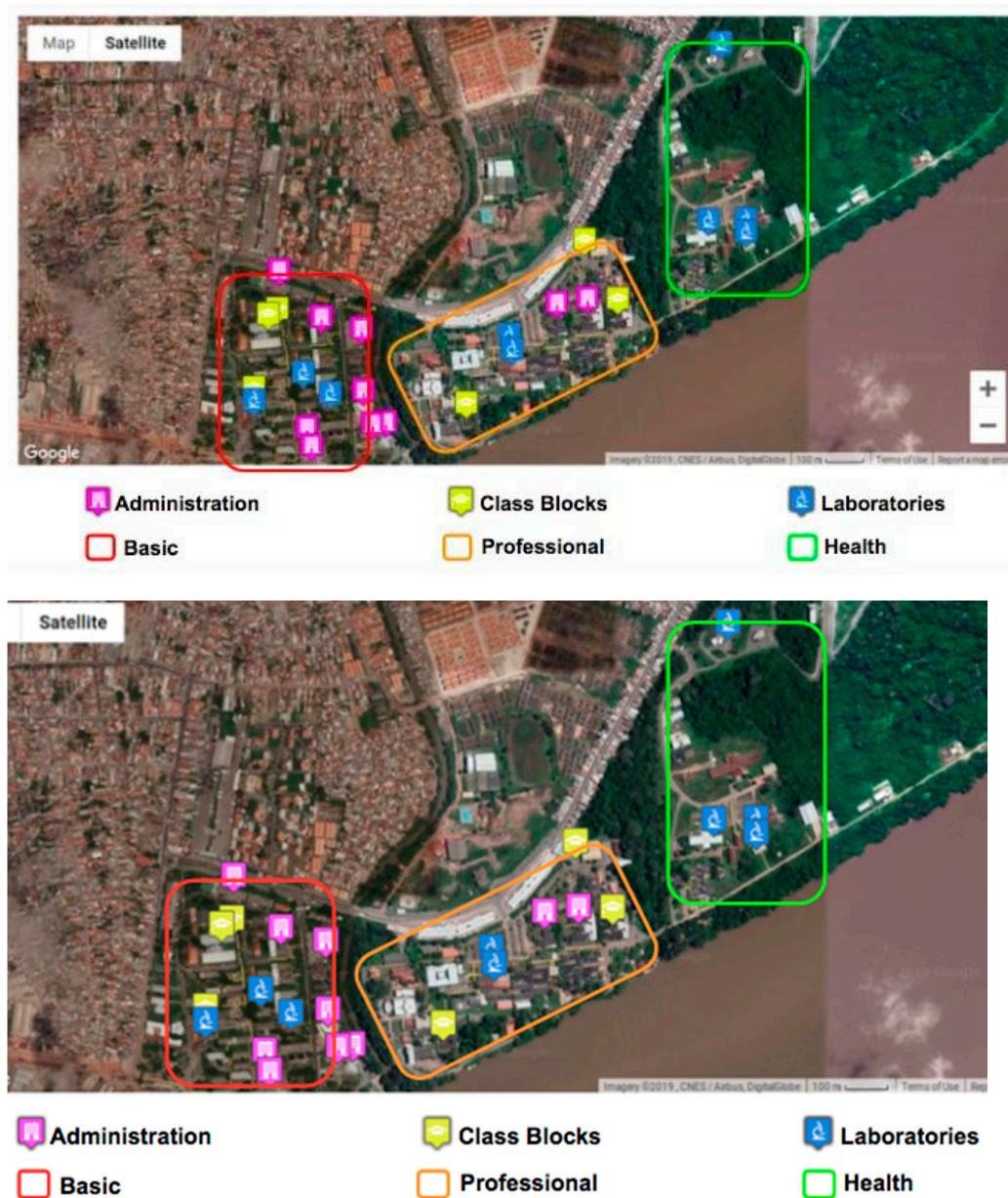


Figure 5. Federal University of Para Campus located in Belém City-Brazil (Adapted from Google Map).

4.2.1. Brief Description of SISGEE (Electric Management System)

SISGEE's main objective is to monitor the steady-state electrical quantities in the University Campus electric grid and to arrange them in an intuitive online environment so that it is easily accessed by computers logged in the Campus available Wide Area Network (WAN). Figure 6 shows schematically the electric grid of the University Campus summarizing the main load blocks.

UFPA's electric loads are distributed in four main sectors, namely: Basic 1, Basic 2, Professional, and Health as detailed in Figure 6. Basic sector is the largest, comprising most graduation courses, administration buildings, banks, convention center, among others, being supplied by two distribution feeders, namely AL1 and AL2. The Professional and Health sectors are supplied by feeders AL3 and AL4 respectively. The common coupling point, where the THDu_x is to be measured is the 13.8 kV substation connecting the internal UFPA electric distribution networks to the local utility distribution grid.

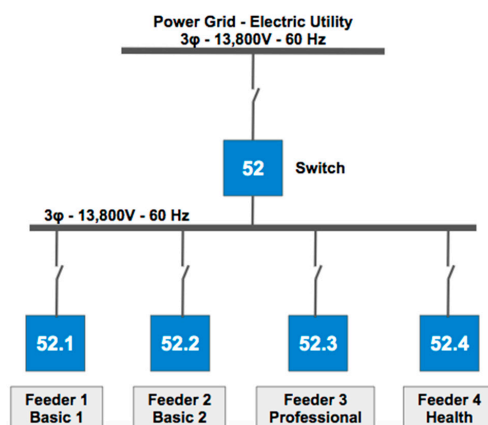


Figure 6. UFFA Belém Campus Electric System Simplified Diagram [21].

SISGEE monitors this power grid through class S Smart Meters (SM) installed at the substations of the largest institutes and buildings. SMs use the MODBUS TCP/IP over Ethernet protocol to send the acquired data to a central server which is responsible for requesting, calculating, storing the acquired data and hosting an internet site for displaying online and real time results. The MS parameters according to factory specification are shown in Table 4.

Table 4. Specifications of Smart Meters Connected to SISGEE.

Electrical Characteristics	
Feed voltage	80 to 300 Vac/Vdc
Consumption	10 VA
Voltage Measurement	30 to 300 Vac (phase-neutral)
Voltage Accuracy	0.5%
Current Measurement	60 A to 3000 A (Model TR4000/TI)
Current Accuracy	0.5%
Phase Angle Accuracy	<5 degree
Power Accuracy	1.0%
Communication	RS-485 MODBUS RTU or Ethernet
Memory Autonomy	Up to 60 days

Table 5 lists the installed SMs by feeder and building highlighting its main activity type and the installed power transformer capacity. In red are the SMs that were inactive in the period of the experiment.

Table 5. SM by feeder and by building.

Feeder	Name	Main Activity	Power (KVA)
1	Language and Communication Institute	Administration	150
	CAPACIT	Administration	225
	Physics Lab (Research)	Laboratories	225
	Communications and IT Center	Laboratories	225
	Department of Material Resources	Administration	225
	Chemistry Lab (Research)	Laboratories	300
	Biological Sciences 1 Institute	Class Blocks	500
	Biological Sciences 2 Institute	Class Blocks	500
	Biological Sciences 3 Institute	Class Blocks	500
2	Central Library 1	Administration	225
	Central Library 2	Administration	225
	Geosciences Institute	Class Blocks	225
	Convention Center Benedito Nunes	Administration	500
	Administration Building 1	Administration	500
	Administration Building 2	Administration	750

Table 5. Cont.

Feeder	Name	Main Activity	Power (KVA)
3	Education Sciences Institute	Administration	225
	Technology Institute	Administration	225
	Architecture Center	Class Blocks	225
	Applied Social Sciences Institute	Class Blocks	225
	Legal Sciences Institute	Class Blocks	300
	Electrical Engineering Lab	Laboratories	300
	Electrical Engineering Lab Annex	Laboratories	500
4	Center of Energy Efficiency in Amazon	Laboratories	225
	Nutrition Center	Laboratories	225
	Physiotherapy College	Laboratories	225
	Odontology	Laboratories	300
—	Main Power Cabin	—	—

Table 6 lists by feeder, the respective nominal power capacity, the effective measured power during the measurement campaign, and the percentage of nominal power capacity used by connected loads. Because, although SISGEE serves the institutes with the largest loads installed, there are some other buildings that are not yet monitored.

Table 6. Percentage of loads by feeder.

Feeder	Location	Nominal Power (kVA)	Measured Power (kVA)	Nominal/Measured (%)
1	Basic 1	5587.5	2850	51.01%
2	Basic 2	3775	2425	64.24%
3	Professional	5950	2000	33.61%
4	Health	3012.5	675	22.41%

4.2.2. Test Scenario Description

This test scenario aims to apply the proposed methodology in an electric network with an implanted AMI system which facilitates the prompt access to the measurements acquired from the real time operation of the electric network. Accessing the UFPA AMI system make available, by phase, real time and historical data involving all electric quantities like voltage, current, power, energy, power factor and also discriminators of power quality like individual and total harmonic distortions indicators.

To validate the proposed methodology of obtaining impact factors in the THDu of a point of interest using only the electric current true rms values, the test was conducted from January 21 to January 27, 2019, because it was the week that had the highest number of THDu violations of the year. Initially, it was selected all meters with status active and that have not extrapolated 30% of null samples. This selection filter is made according to each phase and automatically by MDMS.

Before the final database is formed, it still undergoes a sample removal action that corresponds to the university's general power outages measured at the Main Power Cabin. At the end of this process, the database for constructing and validating the regression model had 897 samples, where the input attributes correspond to true rms current values measured at the points that have active and validated meters, with the THDu measured at the CCP (Main Power Cabin) representing the model output.

It is important to know that of all active meters listed in Table 5, only the meter installed in the Physiotherapy College building was excluded from the Phase B experiment because it did not meet the above prerequisites.

4.2.3. Parameters Definition

Following the steps of the proposed methodology (Figure 2), after selecting the SM that will represent the input and output attributes of the regression model, it is time to create the basis that will be used for training, testing and validation. For this, randomly following the ratio of 80% for the training set and 20% for the validation set. The model was built using the 10-pack cross-validation technique, which forms training and test subsets.

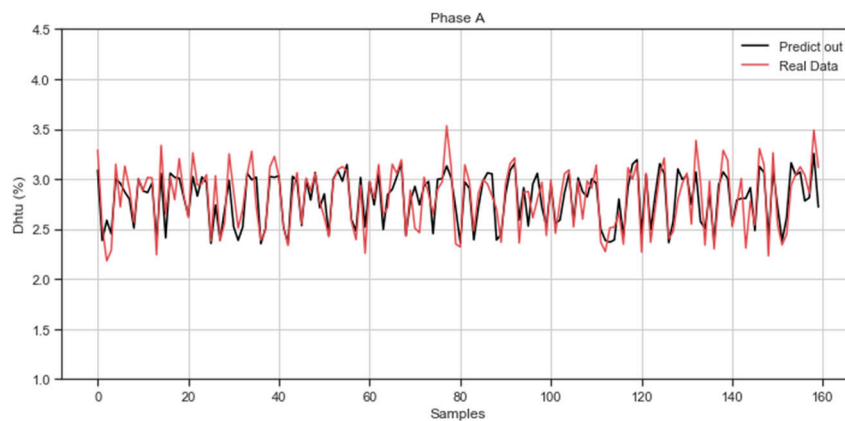
The methodology was implemented using the Python language with the aid of the Scikit-learn library [46] and the GBRT regression technique, through the impurity index obtained from the own GBRT construction process, and the harmonic impact factor is estimated for each SM.

Model performance was evaluated by the Mean Absolute Error (MAE), Mean Square Error (MSE), and Mean Absolute Percent Error (MAPE) metrics obtained by the test set, and the best ratio of configuration parameters for the GBRT model are: Max_depth = 8, Min_sample_split = 9, Min_sample_leaf = 9, N_estimator = 50, Learning rate = 0.09, Loss functions = huber, alpha = 0.85. Table 7 presents the errors obtained for phases A, B, C, by the configured estimator with the above parameters.

Table 7. GBRT model performance using the test set data.

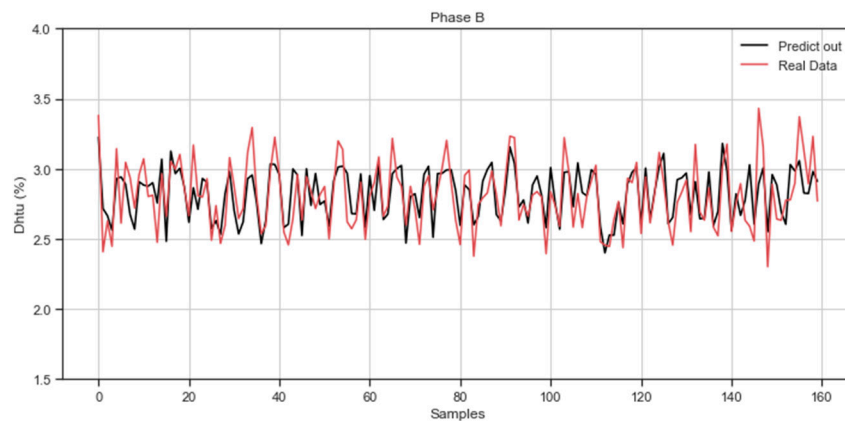
	A	B	C
MAE	0.1379 (+/- 0.000910)	0.1262 (+/- 0.000010)	0.1347 (+/- 0.000047)
MAPE	4.9519 (+/- 0.030815)	4.4932 (+/- 0.000418)	3.1215 (+/- 0.001083)
MSE	0.0304 (+/- 0.000174)	0.0245 (+/- 0.000002)	0.0301 (+/- 0.000019)

Analyzing Table 7, it can be observed that MAPE in all phases remained below 5.0%, which is an acceptable threshold to guarantee a satisfactory performance of this prediction model for the next step, which is the impact factor calculation. Figure 7 compares the THDu values measured at the CCP with the values estimated by the model for phases A-B-C respectively, together with the 95% confidence interval.

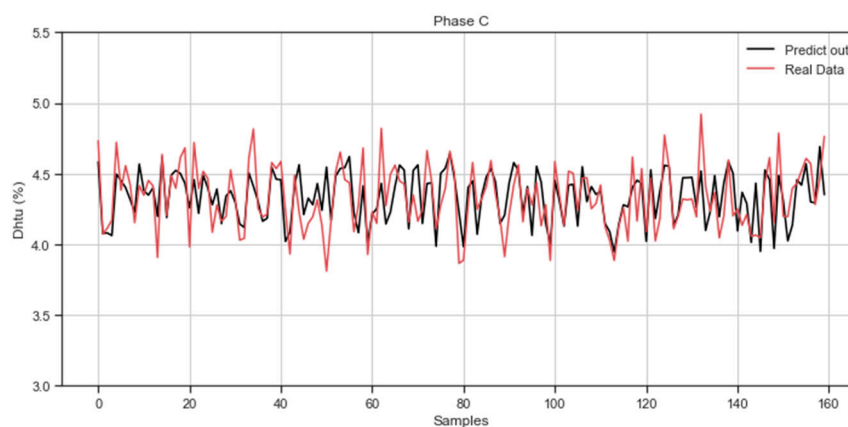


(a)

Figure 7. Cont.



(b)



(c)

Figure 7. THDu measured at the coupling point, Main Power Cabin, during the experiment period in (a) phase A (b) phase B (c) phase C from test set data.

4.2.4. Impact Factor Calculation

In order to calculate the IF by the technique proposed in [10], it is necessary to vary, one by one, each electric current time series in the input vector of the trained model by a percent factor, for example 10%, 20%, etc and then resubmit them to obtain a new estimated output that will be compared with the actual measured values. This procedure was named in [10] as a sensitivity analysis. The difference between both time series is evaluated by the Mean Absolute Percent Error (MAPE), that indicates the IF of the respective load current on the THDu at the point of interest.

As the GBRT regression technique already calculates the importance factor as part of its process, it is evident the gain in processing time offered by the use of GBRT and importance factor. Figure 8 demonstrates how the sensitivity factor varies with varying percent values from 10% to 1000% for phases A-B-C.

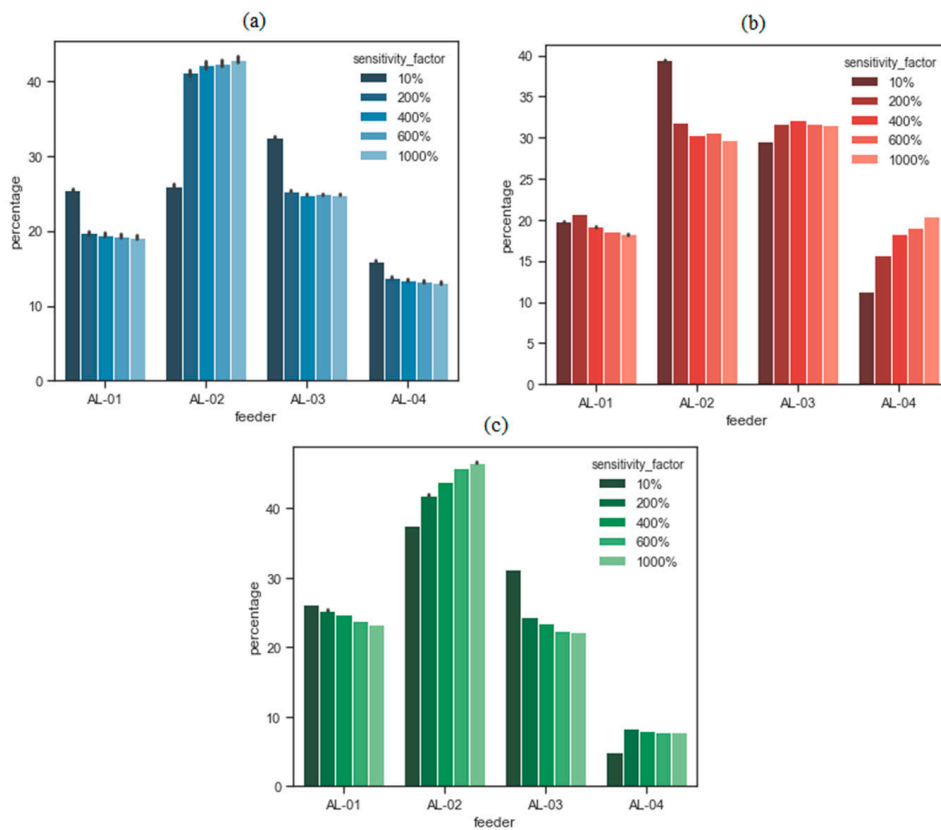


Figure 8. Analysis of the sensitivity factor per feeder in (a) phase A (b) phase B (c) phase C.

To have a good resolution in the sensitivity analysis for the second test system, the input vector was varied 200%, resulting MAPE and Permutation Importance percent values as presented in Figure 9. For the results analysis, the IF calculated for each nonlinear load, represented by the measurements of each SM, were grouped by feeder as can be seen in Figure 9.

This type of representation enables simpler monitoring of the electric network power quality situation and allows the manager to first perform a macro analysis and then focus his actions on specific points.

According to Figure 9, it is highlighted that feeder 2 phases A and B are the most impacting on the THDu at the CCP, by both metrics, that is, MAPE with 41.90% and Permutation Importance with 44.18%. in phase A, and 40.55% and 42.76% in phase B, respectively.

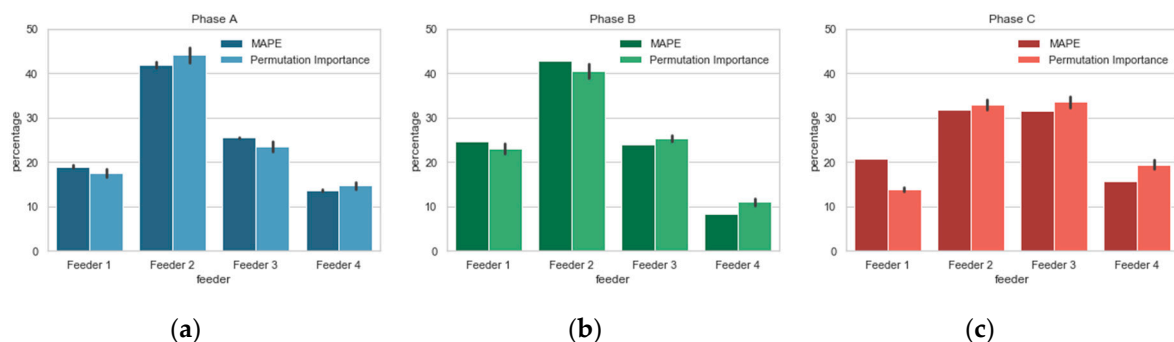


Figure 9. Percent Harmonic impact factor grouped by feeder for phases (a) A, (b) B and (c) C.

It is also important to note that feeder 2 has the largest length and has also the largest installed load. Still looking at feeder 2, and analyzing the detailed data in Table 8, there was a considerable

contribution of the Geoscience Institute with 21.25% of the total IF in phase A at the CCP, the highest percentage of impact factor found in all phases.

In phase C, feeder 2 presented similar results as feeder 3 using both MAPE and Permutation Importance, with approximately 33% of contribution. Summarizing the results in Table 8 it is shown the top three most impacting loads are Geoscience Institute, Administration Building 2 and Legal Sciences Institute.

Table 8. Impact Factor percentage final results of the nonlinear loads at the CCP.

Feeder	Name	Phase A (%)		Phase B (%)		Phase C (%)	
		Permutation	MAPE	Permutation	MAPE	Permutation	MAPE
1	Chemistry Lab (Research)	3.37	3.71	3.45	4.29	2.62	3.65
	Biological Sciences Institute 1	2.28	3.60	4.01	2.58	2.75	7.66
	Biological Sciences Institute 2	5.67	4.45	2.45	2.99	2.32	3.60
	Biological Sciences Institute 3	1.92	1.76	1.26	1.24	1.90	1.32
	Language and Communication Institute	1.56	1.59	5.10	4.31	2.08	1.41
	Physics Lab (Research)	2.82	3.86	6.86	9.32	2.25	3.12
	Total	17.61	18.96	23.14	24.73	13.91	2.76
2	Central Library 2	4.20	2.82	8.76	10.81	2.84	2.58
	Convention Center Benedito Nunes	3.02	4.15	4.60	5.81	6.40	3.92
	Geoscience Institute	21.25	20.85	4.26	4.28	5.42	7.44
	Administration Building 1	7.11	5.82	13.05	12.25	8.69	7.23
	Administration Building 2	8.81	8.27	9.89	9.61	9.63	10.72
	Total	44.38	41.91	40.55	42.76	32.98	31.90
3	Architecture Center	6.72	6.04	2.46	2.54	2.97	3.07
	Electrical Engineering Lab	1.80	2.81	4.80	1.92	3.98	4.28
	Electrical Engineering Lab Annex	4.36	5.72	4.68	3.75	2.89	2.67
	Applied Social Sciences Institute	3.23	2.81	5.61	7.09	3.90	2.76
	Education Sciences Institute	2.89	2.80	1.97	3.92	6.89	4.23
	Legal Sciences Institute	2.22	3.23	3.41	2.05	10.95	12.53
	Technology Institute	2.26	2.10	2.35	2.81	2.08	2.13
Total	23.48	25.51	25.28	24.08	33.65	31.67	
4	CEAMAZON	5.25	4.22	6.48	5.90	3.17	3.94
	Nutrition Faculty	4.47	4.78	4.55	2.53	6.68	3.54
	Physiotherapy College	4.82	4.62	-	-	9.61	8.19
	Total	14.54	13.62	11.03	8.43	19.46	15.67

5. Conclusions

Voltage total harmonic distortion is one of the internationally regulated power quality indicators. The THDu level is a concern for both electric power utilities and larger customers as the responsibility for the negative impact on the grid, may involve legal aspects to be considered. Therefore, it is crucial to find out which loads contribute most to the THD level at the common coupling point, which is the point where the customer and utility are connected.

To contribute to the solution of this problem, this paper presented a new methodology for the identification and quantification of the impacts of multiple nonlinear loads on the electric grid voltage total harmonic distortion at a point of interest by using only true RMS current values employing a concept inherent to the Gradient Boosted Regression Trees technique, which is the impurity index with permutation importance, to find this Percent Impact Factor for each point of interest.

In the first experiment, the use of permutation importance as a metric for the Impact Factor calculation was validated compared to the work developed in [10], finding values very close to the reference scenario obtained from ATP simulations, and reducing the error by 39%. The result obtained by GBRT without preprocessing showed that the technique is very sensitive to bad data, highlighting the need for a data preprocessing step for a good performance of the technique.

Then, the proposed methodology was applied in a real system with an AMI system implanted, obtaining very similar results for both metrics used, that is, MAPE and Permutation Importance. These results make evident that the Permutation Importance can be used as a metric for the calculation of harmonic impacts due to multiple nonlinear loads at a point of interest, which reduces considerably the cost of computer processing in relation to the technique using MAPE as introduced in [10].

Finally, the proposed methodology obtained satisfactory results in both test scenarios through reducing the costs of these evaluations, allowing also the use of this procedure in an integrated way with an AMI system.

Author Contributions: Conceptualization, S.A.M., F.P.M., M.E.T. and U.H.B.; methodology, F.P.M., M.E.T. and U.H.B.; software, F.P.M.; validation, F.P.M. and U.H.B.; formal analysis, F.P.M., S.A.M., M.E.T., U.H.B.; investigation, S.A.M.; data curation, F.P.M.; writing—original draft preparation, S.A.M., F.P.M.; writing—review and editing, F.P.M., S.A.M. and U.H.B.; visualization, M.E.T., U.H.B.; supervision, M.E.T., U.H.B.; project administration, M.E.T., U.H.B.; funding acquisition, M.E.T., U.H.B.

Funding: This work was supported in part by the BRAZILIAN CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO-CNPQ under Universal Project.

Acknowledgments: This research was supported by CEAMAZON—Center of Excellence on Energy Efficiency in the Amazon and by Federal University of Western Para, Campus Oriximiná.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harmonics, P.S. Power System Harmonics: An Overview. *IEEE Trans. Power Appar. Syst.* **1983**, 2455–2460. [[CrossRef](#)]
2. Arrillaga, J.; Watson, N.R. *Power System Harmonics*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2003.
3. IEEE. *IEEE Recommend Practices and Requirements for Harmonic Control in Electrical Power Systems*; IAS—IEEE Industry Applications Society: New York, NY, USA, 2014; pp. 519–1992. [[CrossRef](#)]
4. Variz, A.M. Three-phase Harmonic Load Flow by the Current Injection (Fluxo de carga harmônico trifásico pela injeção de corrente). Master's Thesis, COPPE/UFRJ—Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (Coppe/UFRJ), Rio de Janeiro, Brazil, 2006.
5. Electric Energy Distribution Procedures Guide—PRODIST. *Module 8—Power Quality*; Electric Energy Brazilian Agency (ANEEL): Brasilia, Brazil, 2017.
6. Mohassel, R.R.; Fung, A.; Mohammadi, F.; Raahemifar, K. A survey on Advanced Metering Infrastructure. *Int. J. Electr. Power Energy Syst.* **2014**, *63*, 473–484. [[CrossRef](#)]
7. Sigle, M.; Bauer, M.; Liu, W.; Dostert, K. Transmission Channel Properties of the Low Voltage Grid for Narrowband Power Line Communication. In Proceedings of the IEEE International Symposium on Power Line Communications and Its Applications, Udine, Italy, April 2011.
8. Sigle, M.; Dostert, K.; Liu, W. Channel characterization and system verification for narrowband power line communication in smart grid applications. *IEEE Commun. Mag.* **2011**, *49*, 28–35. [[CrossRef](#)]
9. Bucci, G.; D'Innocenzo, F.; Dolce, S.; Fiorucci, E.; Ciancetta, F. Power Line Communication, Overview of Standards and Applications. In Proceedings of the XXI IMEKO World Congress Measurement in Research and Industry, Prague, Czech Republic, 30 August–4 September 2015.
10. Manito, A.; Bezerra, U.; Tostes, M.; Matos, E.; Carvalho, C.; Soares, T. Evaluating Harmonic Distortions on Grid Voltages Due to Multiple Nonlinear Loads Using Artificial Neural Networks. *Energies* **2018**, *11*, 3303. [[CrossRef](#)]
11. Xu, W.; Liu, Y. A method for determining customer and utility harmonic contributions at the point of common coupling. *IEEE Trans. Power Deliv.* **2000**, *15*, 804–811. [[CrossRef](#)]
12. Pyzalski, T.; Wilkosz, K. Identification of harmonic sources in a power system: A new method. In Proceedings of the 2005 IEEE Russia Power Tech, St. Petersburg, Russia, 27–30 June 2005; pp. 1–6.
13. Wang, W.; Nino, E.E.; Xu, W. Harmonic impedance measurement using a thyristor-controlled short circuit. *IET Gener. Transm. Distrib.* **2007**, *1*, 707–713. [[CrossRef](#)]
14. Srinivasan, D.; Ng, W.S.; Liew, A.C. Neural-network-based signature recognition for harmonic source identification. *IEEE Trans. Power Deliv.* **2006**, *21*, 398–405. [[CrossRef](#)]
15. Mazumdar, J.; Harley, R.G.; Lambert, F.C. System and Method for Determining Harmonic Contributions from Non-Linear Loads. U.S. Patent US 7,013,227, 14 March 2006.
16. Mazumdar, J.; Harley, R.G.; Lambert, F.C.; Venayagamoorthy, G.K. Neural network-based method for predicting nonlinear load harmonics. *IEEE Trans. Power Electron.* **2007**, *22*, 1036–1045. [[CrossRef](#)]
17. Mazumdar, J.; Harley, R.G. Recurrent neural networks trained with backpropagation through time algorithm to estimate nonlinear load harmonic currents. *IEEE Trans. Ind. Electron.* **2008**, *55*, 3484–3491. [[CrossRef](#)]

18. Yin, Z.; Sun, Y.; Yu, T. New methods exploration for harmonic source identification technologies. In Proceedings of the 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT), Weihai, Shandong, China, 6–9 July 2011; pp. 399–402.
19. Kandevar, N.P.; Chénard, S. Method for determining customer contribution to harmonic variations in a large power network. In Proceedings of the International Conference Harmonics and Quality of Power (ICHQP), Bergamo, Italy, 26–29 September 2010; pp. 1–7.
20. Matos, O.; Soares, T.M.; Bezerra, U.H.; de Lima Tostes, M.E.; Manito, A.R.A.; Costa, B.C., Jr. Using linear and non-parametric regression models to describe the contribution of non-linear loads on the voltage harmonic distortions in the electrical grid. *IET Gener. Transm. Distrib.* **2016**, *10*, 1825–1832. [[CrossRef](#)]
21. Paixao, U.C., Jr.; Manito, A.R.A.; Rocha, G.V.S.; Monteiro, F.P.; Carvalho, C.C.M.M.; Bezerra, U.H.; Tostes, M.E.L. Evaluation of Harmonic Contribution Impacts in the Electric Grid Through Linear Regression, Artificial Neural Networks and Regression tree. In Proceedings of the IEEE PES Transmission & Distribution Conference and Exhibition—Latin America (T&D-LA), Lima, Peru, 18–21 August 2018. [[CrossRef](#)]
22. Moradifar, A.; Akbari Foroud, A.; Gorgani Firouzjah, K. Intelligent localisation of multiple non-linear loads considering impact of harmonic state estimation accuracy. *IET Gener. Transm. Distrib.* **2017**, *11*, 1943–1953. [[CrossRef](#)]
23. Mazin, H.E.; Xu, W.; Huang, B. Determining the harmonic impacts of multiple harmonic-producing loads. In Proceedings of the IEEE Power Energy Society General Meeting, Detroit, MI, USA, 24–28 July 2011; pp. 1187–1195. [[CrossRef](#)]
24. Ganjisaffar, Y.; Caruana, R.; Lopes, C.V. Bagging gradient boosted trees for high precision, low variance ranking models. In Proceedings of the 34th SIGIR, Beijing, China, 24–28 July 2011; pp. 85–94.
25. Asadi, N.; Lin, J.; de Vries, A.P. Runtime Optimizations for Tree-Based Machine Learning Models. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 2281–2292. [[CrossRef](#)]
26. Zhang, B.; Ren, J.; Cheng, Y.; Wang, B.; Wei, Z. Health Data Driven on Continuous Blood Pressure Prediction Based on Gradient Boosting Decision Tree Algorithm. *IEEE Access* **2019**, *7*, 32423–32433. [[CrossRef](#)]
27. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
28. Breiman, L. *Manual on Setting Up, Using, and Understanding Random Forests*; v3.1; Statistics Department University of California Berkeley: Berkeley, CA, USA, 2002.
29. Warren, H.S. Cyclic Redundancy Check. In *Hacker's Delight*, 2nd ed.; Pearson Education: Upper Saddle River, NJ, USA, 2013; pp. 319–323.
30. Lin, L.; Sherman, P.D. Cleaning data the Chauvenet way. In Proceedings of the South East SAS Users Group—SESUG, Hilton Head, SC, USA, 4–6 November 2007.
31. Holman, J.P. *Experimental Methods for Engineers*, 6th ed.; McGraw-Hill: New York, NY, USA, 1994.
32. Louppe, G.; Wehenkel, L.; Suter, A.; Geurts, P. Understanding variable importances in forests of randomized trees. In Proceedings of the NIPS 2013—Neural Information Processing Systems Conference, Lake Tahoe, NV, USA, 5–10 December 2013.
33. Nandlall, S.D.; Millard, K. Quantifying the Relative Importance of Variables and Groups of Variables in Remote Sensing Classifiers Using Shapley Values and Game Theory. *IEEE Geosci. Remote Sens. Lett.* **2019**, *1*–5. [[CrossRef](#)]
34. Behnamian, A.; Millard, K.; Banks, S.N.; White, L.; Richardson, M.; Pasher, J. A Systematic Approach for Variable Selection with Random Forests: Achieving Stable Variable Importance Values. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1988–1992. [[CrossRef](#)]
35. *Alternative Transiente Program—ATP (ATP Draw)—User's Manual, Version 5.6*; Norwegian University of Technology Trondheim: Trondheim, Norway, 2009.
36. Abu-Hashim, R.; Burch, R.; Chang, G.; Grady, M.; Gunther, E.; Halpin, M.; Harziadonin, C.; Liu, Y.; Marz, M.; Ortmeier, T.; et al. Test systems for harmonics modeling and simulation. *IEEE Trans. Power Deliv.* **1999**, *14*. [[CrossRef](#)]
37. Test System 3: A 13-Bus Balanced Industrial Distribution System, Harmonics Working Group—Power Quality Subcommittee, IEEE Power and Energy Society. Available online: <http://grouper.ieee.org/groups/harmonic/simulate/download.htm> (accessed on 8 October 2019).
38. Jegadeesan, M.; Venkatesh, N. Estimation and Mitigation of Current and Voltage Harmonics in Distribution System Using Distributed Generation. In Proceedings of the International Conference on Innovations in Engineering and Technology (ICIET'14), Madurai, Tamil Nadu, India, 21–22 March 2014.

39. El-Mamlouk, W.M.; Mostafa, H.E.; El-Sharkawy, M.A. Active power filter controller for harmonic suppression in industrial distribution system. *Ain Shams Eng. J.* **2011**, *2*, 161–172. [[CrossRef](#)]
40. Ayunda, S.W. Harmonics propagation and distortion caused by a nonlinear load in balance distribution network. *Afr. J. Phys.* **2016**, *3*, 110–118.
41. Cano-Plata, E.A.; Jimenez-Lozano, G.; Soto-Marín, O.J.; Ustariz-Farfán, A.J. Hybrid Simulation of Power Quality Assessment: An Application for Power Ground Grid in Arc Furnace Systems. *IEEE Trans. Ind. Appl.* **2016**, *52*, 2648–2653. [[CrossRef](#)]
42. Muniz, J.R.S.; Moraes, J.A.A.; Rocha, G.V.S.; Nunes, M.V.A.; Barradas, R.P.S.; Bezerra, U.H.; Brito, A.B.; Monteiro, F.P.; Carvalho, R.L.S. ATP Cards Automatic Generation from an Electrical Network Elements Database Using Python. In Proceedings of the IEEE International Conference on Industry Applications (INDUSCON), São Paulo, Brazil, 11–14 November 2018. [[CrossRef](#)]
43. Tavares, K.A.; Silva, K.M. Evaluation of Power Transformer Differential Protection Using the ATP Software. *IEEE Latin Am. Trans.* **2014**, *12*, 161–168. [[CrossRef](#)]
44. Chen, X. A three-phase multi-legged transformer model in ATP using the directly-formed inverse inductance matrix. *IEEE Trans. Power Deliv.* **1996**, *11*, 1554–1562. [[CrossRef](#)]
45. Apostolopoulos, C.; Tsakiris, D. Design and Performance Evaluation of a High-Impedance REF Scheme for MV/LV Transformers. *IEEE Trans. Ind. Appl.* **2015**, *51*, 5398–5409. [[CrossRef](#)]
46. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).