

Article

A Method for Responsibility Division of Multi-Harmonic Sources Based on Canonical Correlation Analysis

Yi Zhang ¹ , Youran Wang ², Junyu Guo ^{1,*} and Zhenguo Shao ¹

¹ Department of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350100, China; zhangyi@fzu.edu.cn (Y.Z.); shao.zg@fzu.edu.cn (Z.S.)

² Fuzhou Power Supply Company of State Grid Fujian Electric Power Co., Ltd., Fuzhou 350000, China; wangyouran_86@163.com

* Correspondence: 200127083@fzu.edu.cn

Abstract: The variable background harmonic data and incomplete phasor information make multi-harmonic source responsibility division in three-phase symmetrical power system a significant challenge. In this paper, a background harmonic data selection method based on canonical correlation analysis is proposed to deal with multi-harmonic source responsibility division without phasor information. Firstly, the canonical correlation coefficient between harmonic voltage and harmonic current is used to characterize the fluctuations of background harmonic voltage. Then, the sliding window method is adopted to select the harmonic voltage and harmonic current with small fluctuations. Next, the canonical correlation results for selected data are used to calculate the harmonic responsibility index via the linear regression method. The harmonic responsibility index in the form of percentage represents the harmonic responsibility division. Finally, several experimental results demonstrate that the proposed method has a high accuracy in calculating the harmonic responsibility division, particularly when the user side contains fluctuations of unknown harmonic sources.



Citation: Zhang, Y.; Wang, Y.; Guo, J.; Shao, Z. A Method for Responsibility Division of Multi-Harmonic Sources Based on Canonical Correlation Analysis. *Symmetry* **2021**, *13*, 1451. <https://doi.org/10.3390/sym13081451>

Academic Editor: Christos Volos

Received: 28 May 2021

Accepted: 7 August 2021

Published: 9 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: power quality; harmonic responsibility division; canonical correlation analysis

1. Introduction

In the field of electronic and electrical engineering, nonlinear load is widely used, which makes the harmonic pollution of three-phase symmetrical power system increasingly serious. Nonlinear devices, such as generators, inverters, and rectifiers, inject harmonic currents into their connected power grid, resulting in voltage distortion at the point of common coupling (PCC), which in turn affects the current at this point. Moreover, the interactions among multiple harmonic loads are complicated, making it difficult to locate the sources of distortion. Therefore, to address the issue of harmonic pollution [1], accurately identifying the “problematic” harmonic sources and determining the responsibility of each harmonic source are two extremely important aspects, the latter of which being the focus of this paper.

Existing methods of harmonic responsibility division mainly include: the volatility method [2–4], the covariance method [5–7], the blind source separation method [8–11], and the linear regression method [12–20]. The first three methods have shown a higher accuracy in estimating the harmonic impedance, but they all require accurate harmonic waveform sampling data to obtain the harmonic amplitude and phase angle to realize the separation and calculation of the real and imaginary parts. The linear regression method is vulnerable to the fluctuation of the background harmonic. In addition, the regression formulas derived from following the circuit principles in different works may require different data, and some of them do not require phase angle data [19,20]. The volatility method and the covariance method are both proposed based on the responsibility assessment of a single

harmonic source, while the blind source separation method and linear regression method can be used for a single harmonic source and multi-harmonic sources.

At present, responsibility division for multi-harmonic sources has been widely studied, and the proposed methods are generally based on accurately estimating the harmonic impedance [21]. The blind source separation method has been widely studied in recent years. The independent component method (ICA) as well as its improved variants [8–11] are the most representative. When the harmonic source meets the precondition of being statistical independence and non-Gaussianness, individual harmonic current signals can be separated one by one from their mixed signal with merely measuring the harmonic voltage and without knowing the network harmonic impedance. The least squares method and its improved variants [12–15] are a common method to obtain the harmonic impedance through calculating the coefficient of Norton equivalent circuit of each harmonic source [16].

However, in real situations where the assessment of harmonic impedance is affected by the highly autocorrelated harmonic monitoring data, the problem may become ill-conditioned, and the abovementioned methods could fail. The problem is highly possible to be ill-conditioned when multiple harmonic sources exist, because a large amount of data are required for such systems. Partial least-squares method is an improvement of the least squares method, which overcomes the disadvantage of variable dependence for systems model with multiple harmonic sources [17,19]. M-estimation robust regression method [20] can reduce the influence of singular values to a certain extent and realize the division without phase angle information, but this method is mainly used for responsibility division of a single harmonic source. When used for dividing responsibility of multi-harmonic sources, data segments in which the harmonic resource to be analyzed fluctuates greatly while the fluctuations of other harmonic sources being less than 5~10% have to be selected out. In practical cases, the limited available data segments would make the corresponding result difficult to reflect the overall situation of harmonic responsibility in a long time scale.

The relevant works of harmonic responsibility division in the past five years are summarized in Table 1, from which it can be seen that most methods of responsibility division of multi-harmonic sources require data of harmonic phase angle. However, due to the limitation on communication channels and storage capacity, the harmonic phase angle information is usually not stored in existing regional power quality online monitoring systems. To this end, to apply those methods, harmonic measuring instruments have to be used to conduct special tests on site to obtain short-term measurement data. However, the obtained data is not compatible to the monitoring system, and the corresponding results are hard to reflect the overall situation of long-term harmonic responsibility.

In general, the linear regression method can find highly accurate solutions when the utility harmonic voltage is constant; however, utility harmonic voltage usually varies in real systems [22]. When the utility harmonic voltage is under considerable fluctuations, the use of linear regression methods for harmonic impedance assessment becomes less accurate [23,24]. Data analysis methods can weaken the impacts of fluctuations of the background harmonic on the accuracy but may result in a reduction in the amount of available data. Taking the linear regression method as an example, we discuss three typical data analysis methods.

Table 1. Summary of Harmonic Responsibility division methods in the past five years.

Method	Data Requirement	Data Selection	Method Feature
Covariance method [5–7]	Harmonic voltage amplitude; branch harmonic current amplitude; phase angle difference between harmonic voltage and harmonic current (without harmonic initial phase angle).	Not needed.	Applicable for harmonic responsibility division of a single harmonic source.
Blind source separation method [8–11]	Harmonic voltage amplitude and phase angle; branch harmonic current amplitude and phase angle.	Data selection required; limited available data after selection.	Applicable for harmonic responsibility division of multi-harmonic sources.
Linear regression method [12–20]	Harmonic voltage amplitude and phase angle; branch harmonic current amplitude and phase angle (not necessary for some methods).	Data selection required; limited available data after selection.	Applicable for harmonic responsibility division of one or multi-harmonic sources.
The proposed method	Harmonic voltage amplitude at PCC; branch harmonic current amplitude; (all available in grid online monitoring systems).	Data selection required; sufficient available data after selection.	Accurate even when the background harmonic fluctuates greatly; applicable for harmonic responsibility division of multi-harmonic sources.

The first solution is to classify the background harmonic voltage into different categories, and then perform linear regression for each category. For classification of the background harmonic voltage, Zang et al. [25] divided the harmonic data points into segments that take constant utility harmonic voltage, which can be determined from hierarchical K-means clustering. Zhang and coworkers [26] considered the effects of background harmonic fluctuations, and proposed a data selection method based on the improved Hampel weight function. The basic idea of this method is to estimate the utility background harmonic voltage by solving the phasor equation, which avoids the use of iterative methods. However, in real systems where the measurement of phase data is not mandatory, the background harmonic voltage cannot be estimated if the phase information is unknown. The accuracy of the utility background harmonic voltage relies on the accuracy of the utility harmonic impedance; its accuracy is usually low and difficult to calculate when the system harmonic impedance changes [27].

The second solution is to perform linear regressions first and then classifies the background harmonic voltage using iterative methods. As iterative methods are applied in this solution, the computational (both space and time) complexity is significantly increased [28]. The drawback of these two solutions is that both require the harmonic phase data, which are usually not available directly in power quality monitoring systems [29].

To avoid this limitation, the third solution based on data correlation is proposed to assess the harmonic responsibilities by comparing the similarity between the harmonic voltage and the harmonic current and then selecting the strongly relevant data for linear regression [30,31]. However, these methods cannot reflect the degree of correlation for multiple sets of data combinations and are used only for estimating the harmonic emission capability of harmonic sources on both sides of the PCC. For cases where multiple harmonic sources are connected to the PCC, it remains unsolved as to how to analyze the correlations between the harmonic currents and the harmonic voltage.

However, the above-mentioned multi-harmonic source division method requires a large amount of calculations and cumbersome steps, the responsibility division for multi-harmonic sources needs further improvement.

The main contributions of this paper can be highlighted as follows:

- (1) A harmonic responsibility division method based on canonical correlation analysis is proposed. Compared with the existing methods, harmonic phase angle data is not required while still achieving a higher division accuracy even when the background harmonic fluctuates relatively greatly.
- (2) One-to-many correlation analysis between the harmonic voltage at PCC and the data of multi-harmonic sources is realized with canonical correlation analysis method, with which even under great fluctuations of the background harmonic, appropriate data can be selected out.
- (3) Merely using available data from the online power quality monitoring system and without involving additional harmonic measuring instruments, harmonic responsibility division on a longer time scale is realized focusing on long-term stable operation of power grids.

2. Preliminaries on the Division Method of Multi-Harmonic Sources Responsibility

2.1. Projection Coefficient Calculation

Figure 1 shows the Norton equivalent circuit with multi-harmonic sources, where \dot{U}_{PCC} and \dot{I}_{PCC} are harmonic voltage and harmonic current, respectively, at the PCC; $\dot{I}'_u, Z_u, \dot{I}_u$ represent the system side equivalent harmonic current of the harmonic source, equivalent harmonic impedance, and branch harmonic current, respectively; $\dot{I}'_k, Z_k, \dot{I}_k (k = 1, 2, \dots, n)$, similar to $\dot{I}'_u, Z_u, \dot{I}_u$, represent the counterparts on the user side [18].

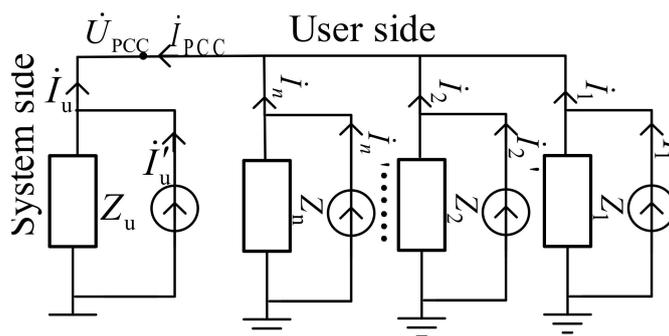


Figure 1. Norton equivalent model.

As proved in [32], it is feasible to use the actual measured current \dot{I} instead of the theoretical current \dot{I}' to carry out the responsibility assessment of centralized multi-harmonic sources, and most current studies default to do so. According to the superposition theorem, \dot{U}_{PCC} can be easily calculated as

$$\dot{U}_{PCC} = \sum_{k=1}^n \dot{U}_k + \dot{U}_0 = \sum_{k=1}^n Z'_k \dot{I}_k + \dot{U}_0 \tag{1}$$

where n is the number of major harmonic sources connected to the PCC; \dot{U}_k is the harmonic voltage at the PCC generated by the harmonic source k ; Z'_k is the equivalent harmonic impedance excluding the branch k ; \dot{U}_0 is the background harmonic voltage for other components at the PCC, including the system harmonic voltage and the user-side non-major harmonic source. Assuming three main harmonic sources ($n = 3$), the relationship between the harmonic source voltage for each user and the harmonic voltage vector at the PCC is shown in Figure 2.

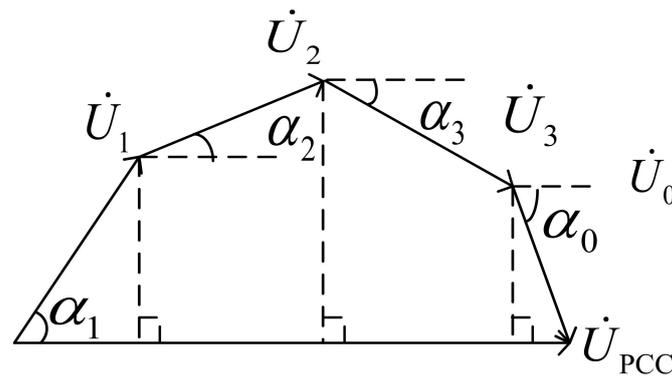


Figure 2. Harmonic voltage vector diagram.

According to the definition, the harmonic responsibility for source k is defined as the ratio of the projection of \dot{U}_k on \dot{U}_{PCC} to the amplitude of \dot{U}_{PCC} , which is expressed as

$$H_k = \frac{|\dot{U}_k| \cos \alpha_k}{|\dot{U}_{PCC}|} \times 100\% = \frac{|\dot{I}_k| Z'_k \cos \alpha_k}{|\dot{U}_{PCC}|} \times 100\% = \frac{|\dot{I}_k| h_k}{|\dot{U}_{PCC}|} \times 100\% \quad (2)$$

where

$$h_k = Z'_k \cos \alpha_k \quad (3)$$

with α_k the phase angle between \dot{U}_{PCC} and \dot{U}_k ($k = 1, 2, \dots, n$). $|\dot{U}_{PCC}|$ and $|\dot{I}_k|$ are the modulus measured at the PCC. h_k ($k = 1, 2, \dots, n$) is the projection coefficient, calculated by the linear regression method.

Figure 2 shows the relationship between \dot{U}_{PCC} and the projection of \dot{U}_k on \dot{U}_{PCC} . Based on Equations (3), Equation (1) can be expressed as

$$\begin{aligned} |\dot{U}_{PCC}| &= \sum_{k=1}^n |\dot{U}_k| \cos \alpha_k + |\dot{U}_0| \cos \alpha_0 \\ &= \sum_{k=1}^n |\dot{I}_k| |Z'_k| \cos \alpha_k + |\dot{U}_0| \cos \alpha_0 \\ &= \sum_{k=1}^n h_k |\dot{I}_k| + |\dot{U}_0| \cos \alpha_0 \end{aligned} \quad (4)$$

Since all the terms in Equation (4) are scalars, hk can be calculated from linear regression, which avoids solving vector equations. Note that the solution of Equation (4) will have high accuracy if $U_0 \cos \alpha$ and hk are both constant. However, $U_0 \cos \alpha$ varies in real systems, which may cause large errors when using the linear regression method.

2.2. Principle of Data Selection

As it is shown in Figure 1, where the system side and the user side are considered as two different parts, the harmonic voltage Equation at the PCC is expressed as

$$\dot{U}_{PCC} = Z_u \dot{I}_{PCC} + Z_u \dot{I}'_u \quad (5)$$

Since $Z_u \ll Z_c$, and Z_c is the integrated equivalent harmonic impedance of the user side. Then, we have:

$$\dot{U}'_0 = \frac{Z_u Z_c}{Z_u + Z_c} \dot{I}'_u \approx Z_u \dot{I}'_u \quad (6)$$

Based on Equations (5) and (6), the background harmonic voltage of the system side can be expressed as

$$\dot{U}'_0 = \dot{U}_{PCC} - Z_u \dot{I}_{PCC} \quad (7)$$

The phasor \dot{U}'_0 can be calculated from the vector Equation (7), and then the data for linear regression are selected via the cluster analysis [21].

The cluster analysis methods for data selection have two major drawbacks: (i) vector calculation is required, and (ii) these methods only select the data with small background harmonic voltage fluctuations from the system side. When the user side contains unknown harmonic sources that fluctuate, these methods consider them to be constant, resulting in errors that are unavoidable.

The data with small background harmonic voltage fluctuations can be directly selected using the data correlation methods; the linear equation is then solved according to the selected data, and the projection coefficient is calculated. These methods require no estimation of the system side harmonic voltage, and they also minimize the influence of the unknown harmonic sources on the user side.

Although the correlation between different sequences has been widely studied in the existing data correlation methods, in this paper we propose a new data selecting method based on the CCA. Equation (4) shows that when $U_0 \cos \alpha$ is a constant, U_{PCC} is linearly related to I_k ($k = 1, 2, \dots, n$). This linear relationship between U_{PCC} and I_k ($k = 1, 2, \dots, n$) is affected by the term $U_0 \cos \alpha$ if it fluctuates. In other words, as the fluctuation of $U_0 \cos \alpha$ increases, the linearity between U_{PCC} and the linear combination of I_k ($k = 1, 2, \dots, n$) decreases, and vice versa. The data can thereby be selected according to the degree of linearity between U_{PCC} and the linear combination of I_k ($k = 1, 2, \dots, n$). The higher the degree of linearity is, the smaller the fluctuation of $U_0 \cos \alpha$ is, and vice versa. Since the traditional correlation analysis methods, such as the Pearson correlation coefficient evaluation method, cannot be applied to cases with two groups of multiple random variables, in this paper we use the CCA method to analyze the correlations between the harmonic voltage and the harmonic currents for multiple harmonic sources.

3. Multi-Harmonic Source Responsibility Division Method

3.1. The Method of CCA Data Selection

The analysis in the previous section indicates that the background voltage is considered constant when the harmonic voltage and the harmonic currents are highly relevant, and thus a multivariate correlation analysis is needed to analyze the relationship between the harmonic voltage and the harmonic currents. The CCA is such a method used to study the correlation among multiple sets of data. The two sets of variables studied are $X = (x_1, x_2, \dots, x_n)$ and $Y = (y_1, y_2, \dots, y_n)$, where x_i and y_i are both m -dimensional vectors, and the two groups of variables are standardized. X' and Y' are defined as comprehensive variables where X' and Y' are both m -dimensional vectors; these two comprehensive variables are to replace the original variables according to the optimization method. More details about the optimization method can be found in [33].

The correlation between the input parameters X and Y can be found by calculating the Pearson correlation coefficient. The expression is shown in Equation (8) as

$$r = \frac{\text{cov}(X', Y')}{\sqrt{D(X')} \sqrt{D(Y')}} \quad (8)$$

where comprehensive variable X' and Y' , $X' = X$, Y' is the optimal linear combination of the variables in Y , and the principle is to make the correlation of X' and Y' strongest; $\text{cov}(X', Y')$ represents the covariance between X' and Y' ; $D(X')$ and $D(Y')$ are the variances of X' and Y' , respectively.

If the correlation between X and Y satisfies the criterion in Equation (9), they are considered strongly correlated [34].

$$r \geq 0.7 \quad (9)$$

In order to illustrate the effects of the CCA on data selection, we set two harmonic sources on the user side and set the harmonic source on the system side fluctuating. Figure 3 is a schematic diagram of the data selection results, where U_{PCC} is the harmonic

voltage at the PCC, I_1 and I_2 represent the harmonic current of harmonic sources 1 and 2, respectively. For the plots on left of Figure 3, the data variance is 0.1, and the background harmonic voltage fluctuations are considered small; while the plots on the right show relatively large background harmonic voltage fluctuations with a variance of 0.5. When the background harmonic voltage fluctuations are small, there is no obvious similarity between the harmonic currents of the two harmonic sources and the harmonic voltage at the PCC, and the comprehensive variable have a strong similarity with the harmonic voltage. From the plots on the right of Figure 3, there is no obvious correlation between the comprehensive variable and the harmonic voltage at the PCC, which is in agreement with the theoretical analysis in Section 2 that as the background harmonic voltage fluctuations increase, the correlation between the comprehensive variable and the PCC harmonic voltage decreases. Based on the CCA, the data with small background harmonic voltage fluctuations can be screened for subsequent analysis and calculations.

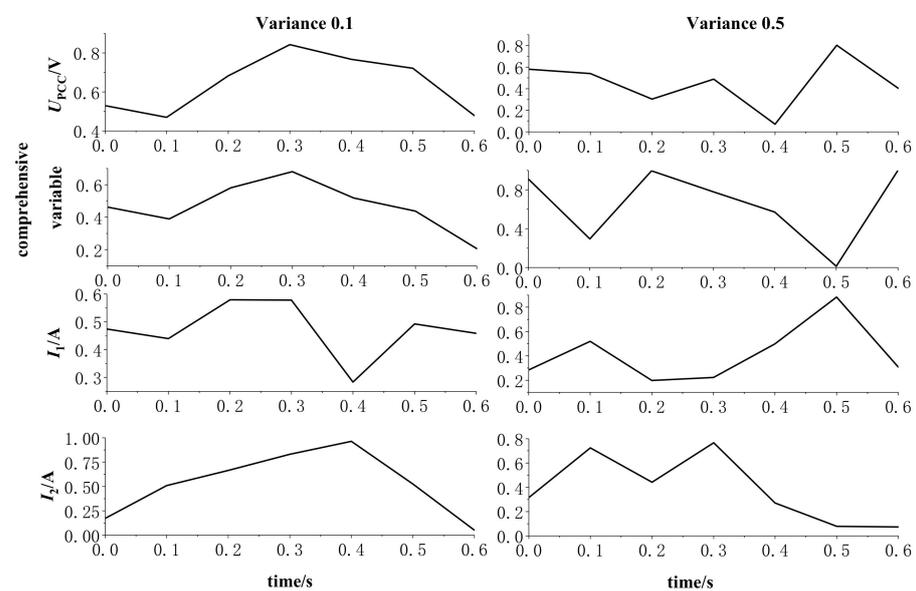


Figure 3. Schematic diagram of variation law of each variable.

3.2. Harmonic Data Sliding Window Analysis

Since it is almost impossible to directly obtain the correlation of the data at a certain moment, the analysis needs to segment the data first. The CCA method proposed in this paper based on sliding window analysis. Assume that the length of each data set is m , and one sliding window includes t sets of harmonic voltage and current amplitude data.

The CCA is performed for each data segment. r_j is the correlation coefficient of the j th data segment. When $r_j \geq r_{th}$ (r_{th} is the correlation coefficient threshold), the background harmonic voltage fluctuations of the j th segment data are considered small, hence each harmonic current and the harmonic voltage at the PCC in the j th segment can be used to calculate the harmonic responsibility coefficient; when $r_j < r_{th}$, the j th data segment is not available, thus is discarded.

The CCA method based on sliding window analyzes long-term data in stages by intercepting data segments. This method analyzes the correlation between the harmonic voltage and the harmonic currents in each data segment. On this basis, the data segments with smaller background harmonic voltage fluctuations are selected, which can be used to calculate the projection coefficient by linear regression; the harmonic responsibility index can then be calculated from Equation (2).

3.3. Harmonic Responsibility Division Procedures

Figure 4 depicts the procedures of the harmonic responsibility division method for multi-harmonic sources, that are based on the CCA proposed in this paper. A brief description of the procedures is provided as follows.

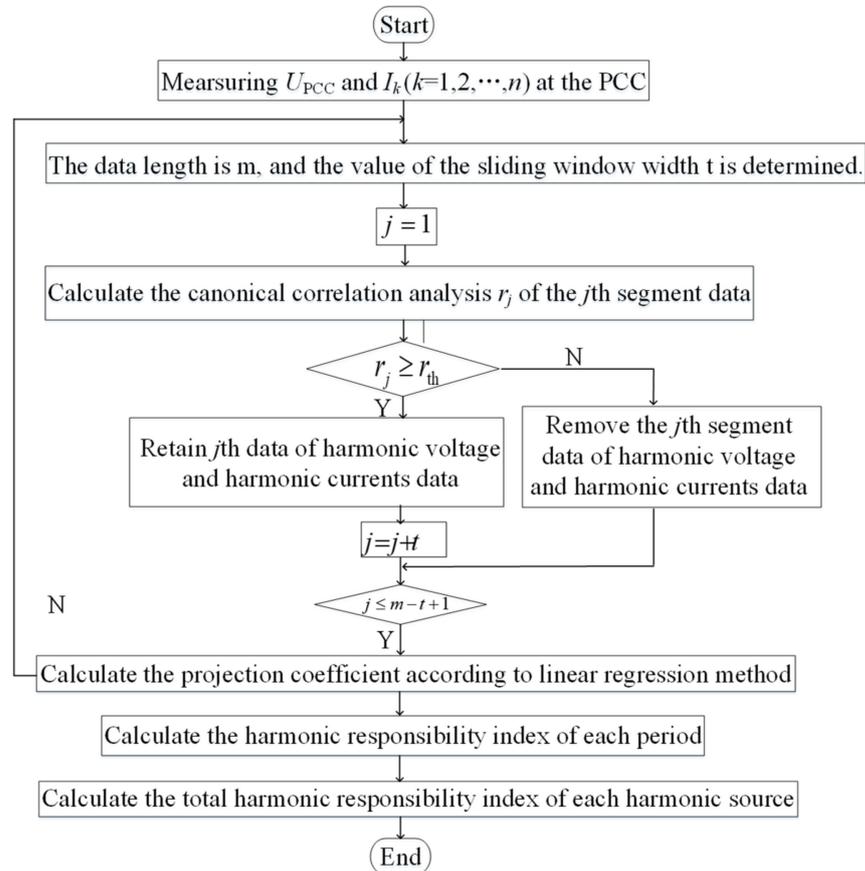


Figure 4. Algorithm flowchart.

Step 1: Measure the harmonic data at the PCC for n major harmonic sources during the time periods of interest. The measurements include the harmonic voltage data at the PCC, U_{PCC} , and the harmonic current data for each source $I_k (k = 1, 2, \dots, n)$.

Step 2: Determine the length of the sliding window t by considering the accuracy and time of the calculation. The smaller sliding window t , the higher calculation accuracy and the longer calculation time.

Step 3: Calculate, using the CCA method, the correlation coefficient, $r_j (1 \leq j \leq n)$, between the harmonic voltage data and the harmonic current data of the j th window. The data are selected if $r_j \geq r_{th}$, and are discarded otherwise.

Step 4: Solve Equation (4) using the linear regression method to calculate the projection coefficient.

Step 5: Calculate the harmonic responsibility index for each harmonic source from Equation (2).

Step 6: Move j one point backwards, and repeat steps (3) to (5) until $j = n$.

Step 7: Calculate the average of the harmonic responsibility index in each period to get the total harmonic responsibility index for each user, the expression is shown in Equation (10) as

$$H'_k = \sum_{j=1}^n H_k / n \quad (10)$$

where, n is the number of total data segments after screening in the analysis period.

4. Case Studies

4.1. Simulation Specifications

The case study demonstrated here is the Norton equivalent circuit with three harmonic sources (see Figure 1 with $n = 3$). The simulation is conducted in Power System CAD (PSCAD), and we take the fifth harmonic as an example. The parameters of the system side and three harmonic sources (user side) are listed in Table 2. The harmonic impedance is a series connection between the resistance R and the reactance L . The fluctuation of each harmonic source is assumed to follow normal distribution with zero mean value, and a variance shown in the last column of Table 2. Note that we used four different variance values to represent varying degrees of fluctuation on the system side.

Table 2. Harmonic Source Parameter Setting.

Harmonic Sources	R/Ω	L/H	I/A	Fluctuation Variance
system side	1	0.001	$1.10\angle-17.5^\circ$	0.15, 0.2, 0.25, 0.3
User 1	5	0.010	$3.03\angle-7.3^\circ$	0.05
User 2	3	0.005	$4.75\angle-14.1^\circ$	0.05
User 3	4	0.008	$2.27\angle-2.7^\circ$	0.05

According to the previous analysis, we set r_{th} to 0.7. Five different methods are cited to divide the harmonic responsibility for each user, which are the partial least squares method [25] (method 1), the mean shift method [35] (method 2), M-estimation robust regression method [20] (method 3), multiple linear regression method [19] (method 4) and the method proposed in this paper. The simulations obtain the harmonic voltage and the harmonic current of each user branch at the PCC.

4.2. Background Harmonic Voltage Fluctuations

The fluctuation of each harmonic source is specified according to Table 2. Here we take the variance of 0.15 as an example. Figure 5 shows the selected data and the removed data of the harmonic voltage and the harmonic currents. The background harmonic source is the system harmonic source, and its degree of fluctuation is larger than any other harmonic source at the PCC. Thus, it can be seen from Figure 5, the data with small system harmonic source fluctuations are selected by the CCA, and the data with large system harmonic fluctuations are removed.

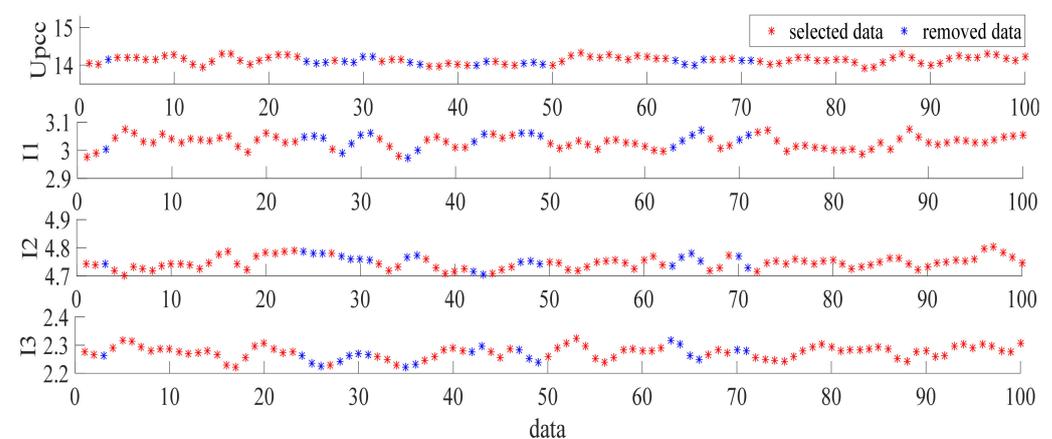


Figure 5. The diagram of selected data and removed data.

Based on the selected data, the projection coefficient is computed from the linear regression method; the harmonic responsibility index then can be calculated from Equation (2). Table 3 shows the harmonic responsibility index calculated from method 1, method

2, and our proposed method. It should be noted that both method 1 and method 2 require phasor information, while our proposed method does not. The actual value in Table 3 represents the ratio of the projection of each user's harmonic voltage on the harmonic voltage at the PCC to the PPC's harmonic voltage when the harmonic source has no fluctuation.

Table 3. Comparison with the Calculation Result of the Harmonic Responsibility Index of the Phase-Required Method.

Harmonic Sources	Fluctuation Variance	Harmonic Responsibility Index/%			
		Actual Value	Method 1	Method 2	The Proposed Method
User 1	0.15	28.41	17.69	26.31	26.48
	0.2		15.45	25.45	29.58
	0.25		13.56	23.88	26.20
	0.3		15.45	22.14	27.42
User 2	0.15	38.90	18.60	36.25	36.03
	0.2		17.88	35.94	37.16
	0.25		15.73	34.75	41.23
	0.3		17.88	34.75	42.87
User 3	0.15	19.89	12.36	19.72	20.32
	0.2		11.07	18.91	22.09
	0.25		10.28	17.72	19.37
	0.3		11.07	9.80	20.41

It can be seen from Table 3 that the proposed method has higher accuracy than method 1, and its error is similar to that of method 2. Since the proposed method requires neither the estimation of the background harmonic voltage nor the phase information, it has higher practical values in engineering applications.

With the background harmonic voltage fluctuation variance set to 0.15, the results from method 3, method 4, and our proposed method are shown in Table 4. Note that neither method 3 nor method 4 requires phasor information.

Table 4. Comparison with the Calculation Results of the Harmonic Responsibility Index of the Phase-Required Method without Phase Method.

Harmonic Sources	Fluctuation Variance	Harmonic Responsibility Index/%			
		Actual Value	Method 3	Method 4	The Proposed Method
User 1	0.15	28.41	27.03	18.61	26.48
User 3		38.90	36.18	20.79	36.03
User 3		19.89	19.27	11.53	20.32

From Table 4, the accuracy of the proposed method is comparable to that of method 3, and is higher than that of method 4. Method 3 used the iterative solution of least squares after weight reduction to weaken the impact of abnormal data on linear regression solutions, but it needs a regression calculation for each harmonic source. Compared with method 3, the method in this paper requires no iteration, only needs one screening and regression to obtain higher accuracy results, and the algorithm is more convenient. The premise of method 4 is that the background harmonic voltage does not fluctuate, so when the background harmonic voltage fluctuates, the method of this paper is more accurate.

Data utilization rate, W , is defined as

$$W = \frac{\text{number of selected data}}{\text{total number of data}} \times 100\% \quad (11)$$

Under different background voltage fluctuations, the data utilization rates of the method proposed in this paper are shown in Table 5:

Table 5. Data Utilization Rates under Different Background Voltage Fluctuations.

Background Harmonic Voltage Fluctuation Variance	W (%)
0.15	66.11
0.2	55.03
0.25	45.30
0.3	12.42

It can be seen from Table 5 that as the background harmonic voltage fluctuations increase, the number of the selected data for linear regression decreases. This can be explained as: with the increase of fluctuation degree, the less available data can be screened out by the CCA method. When the total amount of data is the same, the reduction of filtered data will cause the accuracy of the linear regression method to decrease.

The analysis time scale is another factor affecting the results of the harmonic responsibility index. Table 6 shows the harmonic responsibility index calculated at different analysis time scales.

Table 6. Comparison Table of Harmonic Responsibility Indicators at Different Time Scales.

Analysis Time Scale	Harmonic Responsibility Index/%		
	User 1	User 2	User 3
Actual value	28.41	38.90	19.89
15 h (300 data)	26.48	36.03	20.32
3 days (1440 data)	26.69	36.25	20.13
7 days (3360 data)	26.72	36.30	20.07

Since the harmonic responsibility fluctuates within a small range during the period of analysis, its actual value should be the same as the actual value in Table 3. As is indicated in Table 6, the accuracy of the harmonic responsibility index improves with the increase of the number of data. As the number of data increases, the amount of data selected for linear regression also increases, resulting in more accurate linear regression solutions, and harmonic responsibility indicators. The accuracy of the responsibility division indicator reaches a peak as the number of data is large enough; further increasing the number of data would not significantly increase the accuracy.

4.3. Systems Containing Unknown Harmonic Sources

In the system shown in Figure 1, assume User 3 is an unknown harmonic source, and User 1 and User 2 are subjected to harmonic responsibility division. The fluctuations on the system side follow normal distribution with a mean value of 0 and a variance of 0.15. For User 3, fluctuations have the same mean value of 0 and four different variances, 0.15, 0.2, 0.25, and 0.3. The fluctuations for the harmonic sources in User 1 and User 2 are both normally distributed with the mean value of 0 and a variance of 0.1. The errors of the harmonic responsibility index for each method are shown in Figure 6.

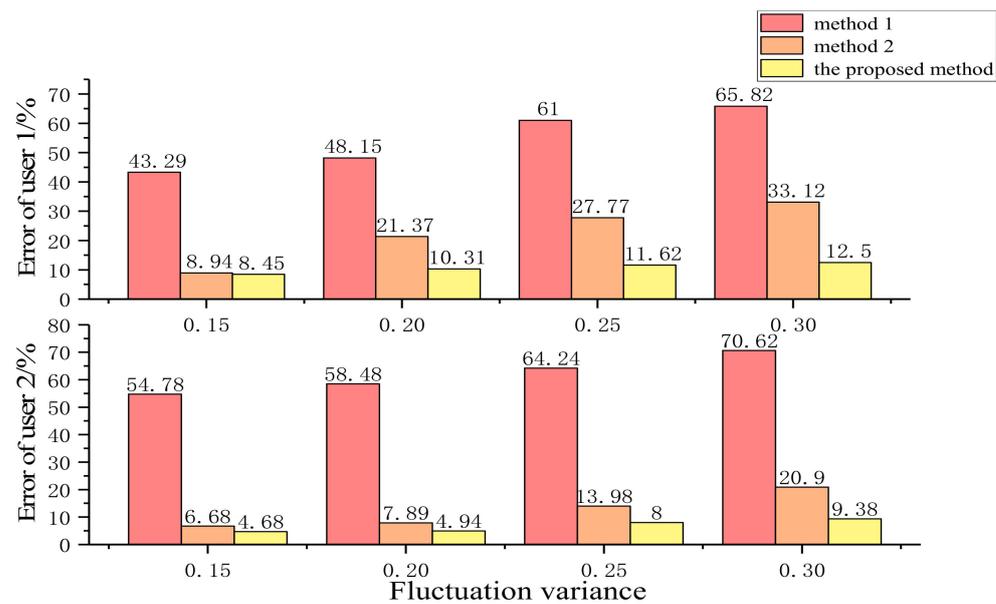


Figure 6. The errors of harmonic responsibility index.

As can be seen from Figure 6, the proposed method can still achieve high accuracy at the presence of an unknown harmonic source. In addition, as the fluctuations of the unknown harmonic source increase, the accuracy is still higher than either the partial least squares method or the mean shift method.

The proposed method selects the data with strong correlations between the harmonic voltage and the harmonic currents for linear regression. Regardless of where the harmonic voltage fluctuations are from (system side or unknown harmonic source), the proposed method is able to solve the harmonic responsibility index using the linear regression method. The partial least squares method takes all the data into account for linear regression which leads to large errors in the calculation of the responsibility index. The mean shift method estimates the background harmonic voltage vector for the system side and selects the system side background harmonic voltage data with small fluctuations for linear regression. However, if an unknown harmonic source exists, it is impossible to select the data with small fluctuations for the unknown harmonic source. In solving this linear regression problem, the voltage of the unknown harmonic source is considered as a part of the background harmonic voltage. The fluctuations of the unknown harmonic sources reduce the linearity between the harmonic voltage at the PCC and the harmonic currents for all sources, resulting in a linear regression solution with low accuracy. When the fluctuation scale of an unknown harmonic source increases, the accuracy of the harmonic responsibility index decreases.

4.4. Changes in Harmonic Responsibility during the Period of Analysis

The above analysis is based on the assumption that the user's harmonic responsibility is relatively stable around a certain central value. In this section we discuss the scenarios when the harmonic responsibility fluctuates significantly. The parameters of each harmonic source are specified as in Table 2. The harmonic source of User 1 is specified to vary from $3.03\angle-7.3^\circ$ to $0.71\angle-7.3^\circ$ at a certain moment, the background harmonic voltage fluctuation is set to 0.1, and the window width t is 60. The harmonic responsibility index in each window width is shown as a function of time in Figure 7.

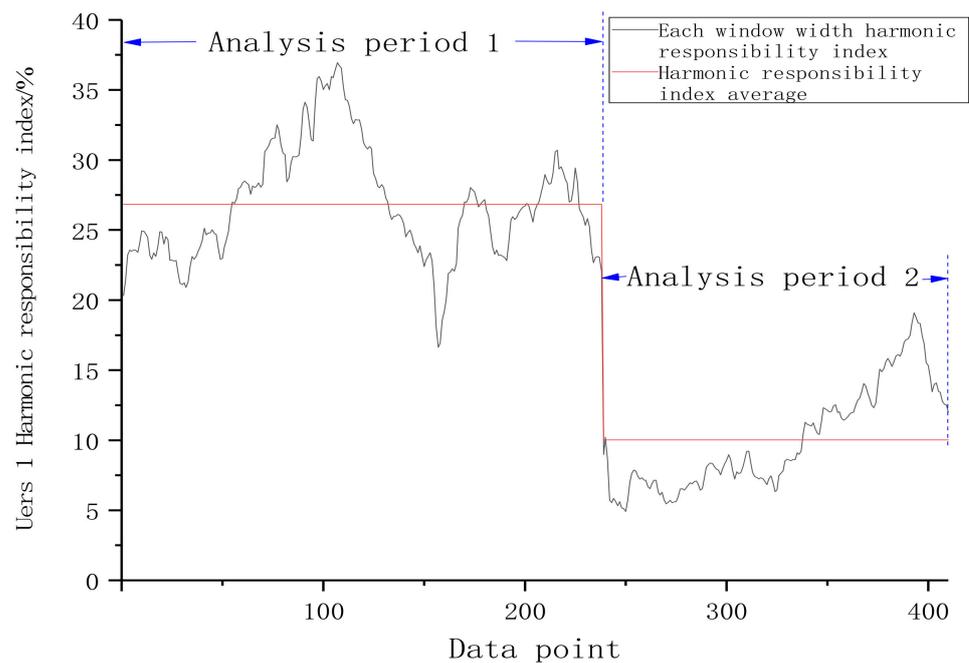


Figure 7. The change chart of harmonic responsibility index.

The black curve in Figure 7 represents the harmonic responsibility index calculated for User 1 within each window width, while the red curve represents the average value of User 1's harmonic responsibility index. It is obvious that the harmonic responsibility has changed significantly, and hence we separate the black curve into two different periods, namely period 1 (left) and period 2 (right). The average value of the harmonic responsibility index for each period is calculated to characterize the harmonic responsibility during that period. The actual values and the average values of the harmonic responsibility indicators are shown in Table 7.

Table 7. Harmonic Responsibility Index when Responsibility Changes Significantly.

User 1 Effective Value of Harmonic Current	Harmonic Responsibility True Value/%	Harmonic Responsibility Average/%
Analysis period 1	28.41	26.84
Analysis period 2	8.03	10.02

According to the data utilization defined in Equation (11), the total data utilization rate is $W = 78.86\%$. The background harmonic voltage fluctuates little in this time period, hence the data utilization rate is relatively high.

As can be seen from Table 7, the method proposed in this paper can still remain a high accuracy under the variance of responsibility. Since filtering data in the proposed method is based on correlation analysis, the selected data have small fluctuations in background harmonic voltage, thus have high accuracy when conducting linear regression. When the harmonic responsibility of the data remains the same in the period of analysis, the harmonic responsibility of the unselected data is approximately the same as that of the selected data; when the harmonic responsibility of the data varies during the period of analysis, the selected data will be the same as in the case when the harmonic responsibility data is unchanged, because the correlation between the harmonic voltage and the harmonic current is unchanged. Since this paper uses the sliding window-based responsibility division, the harmonic responsibility varies significantly with time. To improve the accuracy of the calculation, the average harmonic responsibility of different sections is solved separately.

In the method proposed in this paper, when the background harmonic voltage fluctuates greatly and the harmonic responsibility varies drastically, such data will not be selected, so in this case, a lack of responsibility will occur.

4.5. Practical Case Analysis

In this section, we use a practical case, based on the harmonic data of a 220 kV substation in Zhangzhou, Fujian Province, China, for demonstration. Figure 8 shows the diagram of a substation, and a 110 kV bus with four feeders. The power quality monitoring device measures the 95% probability value of harmonic voltage and the currents at the PCC every 3 min. (In the Chinese national standard, the 95% probability value can be determined by the following method: rearrange the measured data from big to small, then 5% data from the bigger part and select the maximum one from residual part to be 95% probability value [29]). The data used for analysis were collected within 7 days (from 20 September 2019 to 27 September 2019).

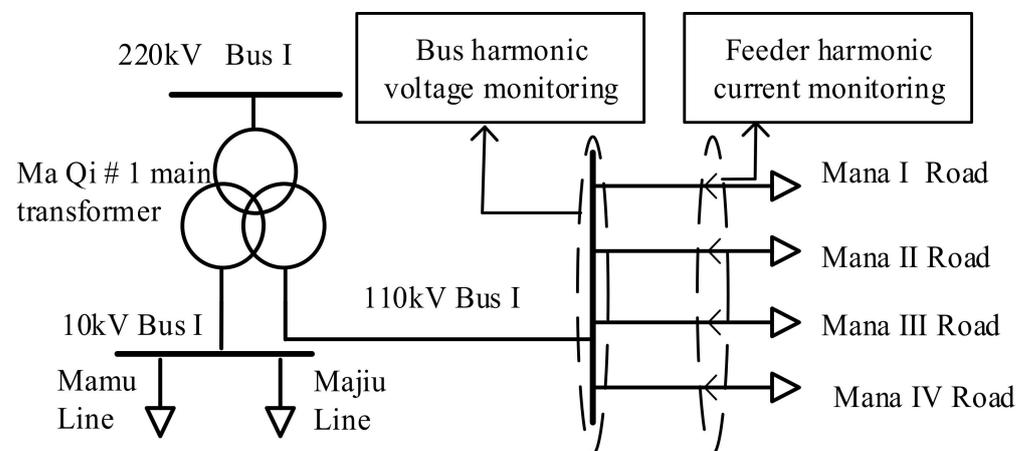


Figure 8. The diagram of the substation.

Figure 9 shows the harmonic voltage and the current data for the fifth harmonic. Since the phase angle data are unavailable in this case, the mean shift method and other methods that require the solution of the phasor equations cannot be applied. In actual systems, the background harmonic voltage is usually under fluctuations, and therefore, the accuracy of the harmonic responsibility index calculated from the linear regression method is low.

The harmonic responsibility division results of the proposed method are shown in Table 8, from which the Manan II Road should take the main harmonic responsibility, as its harmonic responsibility index is the most among the four roads. This point has been proved by the on-site investigation and the actual measurements that the main load of Manan II Road is the electric railway, which results in a large amount of the fifth harmonic emission. The loads for the rest three roads are mainly from the residential electricity or ordinary industries, so their fifth harmonic emissions are small.

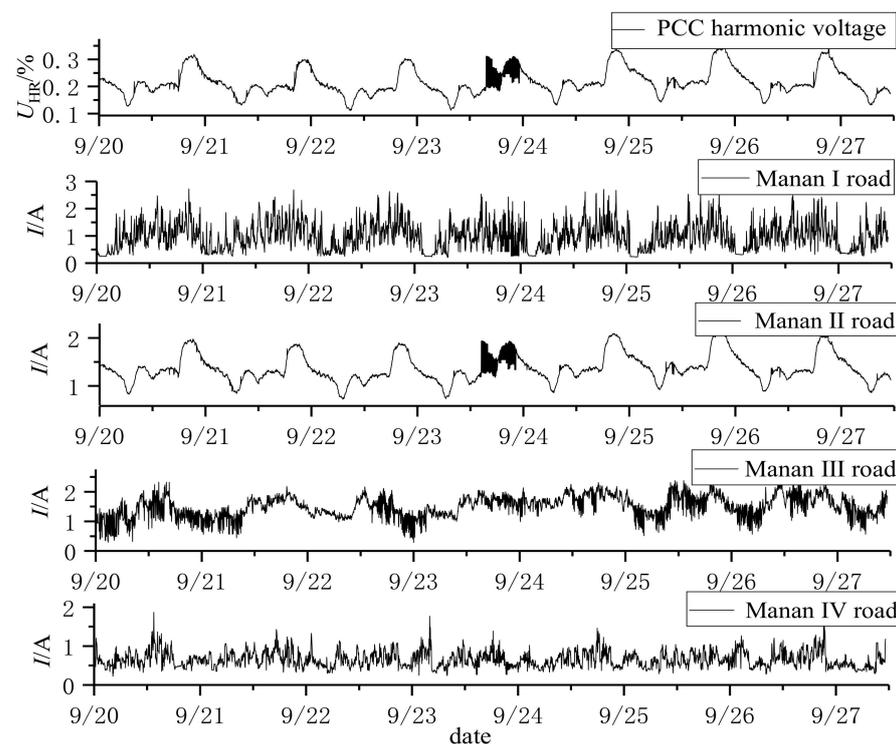


Figure 9. Harmonic voltage and current data diagram for the fifth harmonic.

Table 8. Harmonic Responsibility Index Example.

Feeder	Manan I Road	Manan II Road	Manan III Road	Manan IV Road
Harmonic responsibility index /%	3.66	80.12	7.63	4.67

5. Conclusions

This paper proposes a sliding window based canonical correlation analysis method that can select out appreciated data even when the background harmonics fluctuate greatly, and then divides the multi-harmonic responsibility by partial least square regression method without relying on phase angle information.

In the case study, the following observations are made:

1. Compared with applying traditional partial least squares algorithm and the mean shift method in the harmonic responsibility division, the proposed method shows higher accuracy and restricts the error within 10% in the studied cases.
2. Compared with the M estimation robust regression method, the proposed method shows a similar accuracy that is within the allowable error range of 8%. However, the limited amount of selected data from the M estimation robust regression method in the practical cases will be difficult to reflect the overall situation of harmonic responsibility.
3. Merely using available data from the online power quality monitoring system and focusing on long-term stable operation of power grids, the proposed method realizes harmonic responsibility division on a longer time scale, weakens the impact of short-term measurement and analysis errors and is applicable to practical projects.

As for future work, methodologies for parameter adjustment can be developed to further improve the accuracy of responsibility division for different types of customers, especially users with severe background harmonic fluctuations.

Author Contributions: Conceptualization, Y.Z. and Y.W.; methodology, Y.Z. and Z.S.; software, Y.W. and J.G.; validation, Y.W. and J.G.; writing—original draft preparation, Y.W.; writing—review and editing, J.G., Y.Z. and Y.W.; supervision, Y.Z. and Z.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported in part by the National Natural Science Foundation of China (51777035).

Data Availability Statement: Some or all data, models, or codes generated or used during the study are available from the corresponding author by request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Fernandez, F.M.; Nair, P.S.C. Method for separation of customer and utility contributions of harmonics at point of common coupling. *IET Gener. Transmiss. Distrib.* **2013**, *7*, 374–381. [[CrossRef](#)]
2. Yang, H.; Pirotte, P.; Robert, A. Harmonic Emission Levels of Industrial Loads Statistical Assessment. In *CIGRE Proceedings; International Council on Large Electric Systems*: Paris, France, 1996; pp. 36–306.
3. Hui, J.; Yang, H.; Ye, M. Assessing harmonic emission level based on the impedance gathering trend discrimination. *Proc. CSEE* **2011**, *31*, 73–80.
4. Gong, H.; Xiao, X.; Liu, Y.; Yang, H.G. A method for assessing customer harmonic emission level based on the dominant fluctuation filtering principle. *Proc. CSEE* **2010**, *30*, 22–27.
5. Wu, Y.; Xu, F.; Zhang, W. System-side harmonic impedance estimation based on modified independent random vectors. *Autom. Electr. Power Syst.* **2019**, *43*, 146–152.
6. Xu, F.; Zheng, H.; Yang, H. Harmonic impedance estimation method on system side based on measurement data without phase angle. *Autom. Electr. Power Syst.* **2019**, *43*, 170–176.
7. Xu, F.; Wang, C.; Guo, K. An Improved Utility Harmonic Impedance Estimation Method Based on Measurement Data without Phase Angle. *Proc. CSEE* **2021**, 3149–3158. Available online: <https://xueshu.baidu.com/usercenter/paper/show?paperid=106t08w0b87n0x80kb4v08j0m6241728> (accessed on 27 May 2021).
8. Yang, X.; Yang, H.; Ma, X. Determination of harmonic contributions based on kernel independent component analysis. *Proc. CSU-EPSSA* **2016**, *28*, 109–114.
9. Lin, Y.; Yang, H.; Wang, Z.; Xiao, C. Assessment method of harmonic emission level based on fast approximate joint diagonalization. *Autom. Electr. Power Syst.* **2017**, *41*, 81–87.
10. Zhao, X.; Yang, H. A new method to calculate the utility harmonic impedance based on FastICA. *IEEE Trans. Power Deliv.* **2016**, *31*, 381–388. [[CrossRef](#)]
11. Xu, F.; Wang, C.; Yang, H. A Responsibility Quantification Method of Distributed Multiple Harmonic Sources without Solving Harmonic. *Autom. Electr. Power Syst.* **2020**, *44*, 178–186.
12. De Andrade, G.V., Jr.; Naidu, S.R.; Neri, M.G.; da Costa, E.G. Estimation of the utility's and consumer's contribution to harmonic distortion. *IEEE Trans. Instrum. Meas. Nov.* **2009**, *58*, 3817–3823. [[CrossRef](#)]
13. Zhang, W.; Yang, H. A method for assessing harmonic emission level based on binary linear regression. *Proc. CSEE* **2004**, *24*, 50–54.
14. Qin, F.; Li, J.; Li, Y.; Sun, W.; Li, Z. Research on the responsibility partition of harmonic pollution and harmonic impedance based on the total least-squares regression method. In *Proceedings of the 2014 International Conference on Power System Technology*, Chengdu, China, 20–22 October 2014; pp. 2122–2127.
15. Li, P.; Tao, S.; Yao, L.; Qian, Y.; Sun, J. An assessment method of power system harmonic impedance based on mixed total least squares. In *Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, Xi'an, China, 25–28 October 2016; pp. 1694–1698.
16. Luo, J.; Zang, T.; Fu, L. A novel method to determine the harmonic responsibility for power distribution network including photovoltaic system. In *Proceedings of the 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, China, 26–28 November 2017; pp. 1–5.
17. Li, X.; Chen, M.; Zheng, Y. Assessing the Harmonic Emission Level Based on Partial Least-Squares Regression with Data Envelopment Analysis. *Adv. Mater. Res.* **2014**, 3326, 3367–3370. [[CrossRef](#)]
18. Homan, E.M.; Huang, B.W.X. Determining the harmonic impacts of multiple harmonic-producing loads. *IEEE Trans. Power Deliv.* **2011**, *26*, 1187–1195.
19. Wang, Y.; Mazin, H.E.; Xu, W.; Huang, B. Estimating harmonic impact of individual loads using multiple linear regression analysis. *Int. Trans. Electr. Energy Syst.* **2016**, *26*, 809–882. [[CrossRef](#)]
20. Sun, Y.; Yin, Z. Quantifying Harmonic Responsibilities of Multiple Harmonic Sources Based on M-estimation Robust Regression. *Proc. CSEE* **2012**, *32*, 166–173.

21. Xu, F.; Wang, L.; Zhao, J.; Chen, Y.; Xu, L.; Hu, C. A new method for harmonic impedance estimation. In Proceedings of the 2018 18th International Conference on Harmonics and Quality of Power (ICHQP), Ljubljana, Slovenia, 13–16 May 2018; pp. 1–4.
22. Yang, X.; Yang, H.; Zeng, Q.; Li, H. Assessment of harmonic current responsibility based on sample data. *Electr. Meas. Instrum.* **2016**, *53*, 13–18.
23. Hua, H.; Liu, Z.; Han, B. Calculation method for the harmonic contribution statistical characteristics. *Proc. CSEE* **2016**, *36*, 2959–2965.
24. Hua, H.; Jia, X.; Zhang, S. Neighborhood multi-point measurement method for harmonic contribution determination. *Power Syst. Technol.* **2014**, *38*, 502–508.
25. Zang, T.; He, L.; Fu, L.; Wang, Y.; Qian, Q. Adaptive method for harmonic contribution assessment based on hierarchical K-means clustering and Bayesian partial least squares regression. *IET Gener. Transm. Distrib.* **2016**, *10*, 3220–3227. [[CrossRef](#)]
26. Zhang, J.; Chen, H.; Zeng, L.; Ding, T. Responsibility division optimization for multiple harmonic sources based on the improved hampel weight function. In Proceedings of the 2018 18th International Conference on Harmonics and Quality of Power (ICHQP), Ljubljana, Slovenia, 13–16 May 2018; pp. 1–6.
27. Karimzadeh, F.; Esmaili, S.; Hosseinian, S.H. A novel method for noninvasive estimation of utility harmonic impedance based on complex independent component analysis. *IEEE Trans. Power Deliv.* **2015**, *30*, 1843–1852. [[CrossRef](#)]
28. Karimzadeh, F.; Esmaili, S.; Hosseinian, S.H. Method for determining utility and consumer harmonic contributions based on complex independent component analysis. *IET Gener. Transmiss. Distrib.* **2016**, *10*, 526–534. [[CrossRef](#)]
29. National Energy Administration. Technical Specification for Power Quality Monitoring and Management System: DL/T 1297—2013[S]. 2013. Available online: <http://www.nssi.org.cn/nssi/front/85161180.html> (accessed on 27 May 2021).
30. Qiu, S.; Yang, H. Assessment on Harmonic Emission Level Considering Volatility Coefficient Screening. In Proceedings of the CSU-EPSCA, Tianjin, China, 29 April 2017; Volume 29, pp. 55–59.
31. Gu, W.; Qiu, H.; Yin, X. Waveform matching based method for harmonic contribution determination. *Autom. Electr. Power Syst.* **2017**, *41*, 129–134.
32. Sun, Y.; Li, J. Quantifying Harmonic Impacts for Concentrated Multiple Harmonic Sources Using Actual Data. *Proc. CSEE* **2014**, *34*, 2164–2171.
33. Nielsen, A.A. Multiset canonical correlations analysis and multispectral, truly multitemporal remote sensing data. *IEEE Trans. Image Process.* **2002**, *11*, 293–305. [[CrossRef](#)]
34. Rumpf, D.L. Statistics for Dummies. 2011. Available online: <http://www.doc88.com/p-19759443825165.html> (accessed on 27 May 2021).
35. Wang, Y.; Zang, T.; Fu, L. Harmonic Contribution Partition of Multiple Harmonic Sources Considering Background Harmonic Voltage Fluctuation. *Autom. Electr. Power Syst.* **2015**, *39*, 55–61.