

Article

Study on Harmonic Impedance Estimation Based on Gaussian Mixture Regression Using Railway Power Supply Loads

Yankun Xia ^{1,2,*}  and Wenzhang Tang ² ¹ Traction Power State Key Laboratory, Southwest Jiaotong University, Chengdu 610031, China² School of Electrical and Electronic Information Engineering, Xihua University, Chengdu 610039, China

* Correspondence: yankunjtdx@163.com

Abstract: There are a huge number of harmonics in the railway power supply system. Accurately estimating the harmonic impedance of the system is the key to evaluating the harmonic emission level of the power supply system. A harmonic impedance estimation method is proposed in this paper, which takes the Gaussian mixture regression (GMR) as the main idea, and is dedicated to calculating the harmonic impedance when the load changes or the background harmonic changes in the traction power supply system. First, the harmonic voltages and currents are measured at the point of common coupling (PCC); secondly, a Gaussian mixture model (GMM) is established and optimized parameters are obtained through the EM algorithm; finally, a Gaussian mixture regression is performed to obtain the utility side harmonic impedance. In the simulation study, different harmonic impedance estimation models with uniform distribution and Gaussian distribution are established, respectively, and the harmonic impedance changes caused by different system structures in the railway power supply system are simulated. At the same time, the error is compared with the existing method to judge the accuracy and robustness of this method. In the case analysis, the average value, average error, standard deviation and other indicators are used to evaluate this method. Among them, the average error and standard deviation of this method are about one-fifth to one-third of those of the binary linear regression (BLR) method and the independent random vector (IRV) method. At the same time, its index is slightly better than that of the support vector machine (SVM) method.

Keywords: railway traction power supply loads; harmonic impedance; Gaussian mixture regression; uniform distribution; Gaussian distribution; robustness



Citation: Xia, Y.; Tang, W. Study on Harmonic Impedance Estimation Based on Gaussian Mixture Regression Using Railway Power Supply Loads. *Energies* **2022**, *15*, 6952. <https://doi.org/10.3390/en15196952>

Academic Editors: Xiaofeng Yang, Fujun Ma, Lei Wang, Wei Liu and Ke Wang

Received: 28 July 2022

Accepted: 20 September 2022

Published: 22 September 2022

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



Copyright: © 2022 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/>).

1. Introduction

Electrified railways contain a large number of impactful, single-phase and nonlinear loads. Harmonics have become common power quality problems in the traction power supply system [1,2]. The problem of harmonic pollution has become an important problem in the research of the traction power supply system. Serious harmonic pollution will affect the stable operation of electrified railways, and even cause equipment damage [3,4].

Internationally, harmonic pollution is often reduced through “reward and punishment programs”, that is, relevant policies are introduced to impose economic penalties on harmonic emission sources, and at the same time, harmonic victims are financially compensated [5,6]. In heavy load, this kind of scheme has some room for “negotiation”. Therefore, the division of responsibilities for harmonics is very important, which affects the implementation of the “reward and punishment programs”. At the same time, the correct acquisition of harmonic impedance is the main difficulty in the division of harmonic responsibility [7,8].

Harmonic impedance estimation is a power quality analysis problem, which has been widely concerned by experts and scholars. The “non-intervention methods” are often used in the calculation of harmonic impedance of traction power supply system. Harmonic voltages and currents at the PCC are used to estimate the harmonic impedances, in this

type of method [9–11]. Undoubtedly, such methods have no impact on the operation of the system and are easy to implement and operate. Currently, there are some classic “non-intervention methods” in this area of research, including: fluctuation method [12], linear regression method [13], independent random vector method [14], blind source separation method [15], machine learning method [16], etc.

- (1) The fluctuation method directly uses the ratio of harmonic voltage change rate to harmonic current change rate to replace the harmonic impedance at a specific time [17].
- (2) The linear regression method takes harmonic voltage and current as variables, and the impedance can be obtained by solving the regression parameters [18].
- (3) Independent random vector method, which takes advantage of the property of the probability theory that the covariance of two independent random vectors is 0. The impedance value is solved by canceling the background harmonic variation term [19].
- (4) The blind source separation method uses the statistical principle to separate the harmonic mixed signal, and solves the harmonic impedance by separating the signal [20].
- (5) The machine learning method, the internal connection between the input value and the output value is established. Taking the harmonic voltage and harmonic current collected at the PCC as input samples and the reference harmonic impedance as the output value, the calculation model is obtained and the harmonic impedance is solved [21].

In addition to the above “non-intervention” method, the harmonic impedance of railway power supply system can also be obtained by the “intervention” method [22–24], that is, the harmonic impedance is estimated by controlling the switching of each branch or injecting harmonic current into the traction power supply system [25]. Such methods will affect the stability of the system operation, and the actual use will be limited by the on-site environment.

Obviously, the existing methods all have various deficiencies that have been shown by research. For example, the BLR method mentioned in the literature [26] has high requirements on the measurement accuracy of the harmonic voltage and current, and is greatly affected by the background harmonic fluctuation. Reference [27] proposed the IRV method, which needs to determine whether the customer’s harmonic source is connected with the harmonic source on the utility side. If there is a strong connection, this method is not practical. Reference [28] has proposed the SVM method, which may get trapped in local optima during operation and will affect the stability of the results.

To sum up, the harmonic impedance calculation of the railway power supply system is still restricted by problems such as the large fluctuation of background harmonics and the uncertainty of harmonic source correlation. Therefore, to overcome these problems in harmonic impedance calculation, a method of measuring the harmonic impedance of the railway power supply system via GMR is proposed [29]. GMR has been widely used in time series, automatic control, image denoising and other fields [30]. The proposed method works well in the case of uncertain harmonic source correlation, and has certain progress in timeliness. The accuracy and stability of the proposed method are verified via simulation studies and case analysis.

The work arrangement of each part of this paper is as follows. In Section 2, the basic network structure of the railway power supply system is first reviewed, and the basic model of harmonic impedance estimation is introduced according to the Norton equivalent circuit. Then, GMM and GMR are introduced, and the algorithm description and calculation steps are given. In Section 3, the harmonic impedance estimation models of uniform distribution and Gaussian distribution are built, respectively, and the harmonic impedance of the two models is calculated by the method proposed in this paper and three other existing methods. The conclusions of the three methods were compared and the accuracy and stability of the proposed method were analyzed. In Section 4, the proposed method is validated with a case study of a railway power supply system. Finally, conclusions are drawn in Section 5.

2. Basic Principles

2.1. Basic Structure of Electrified Railway Power Supply System

The power source in the electrified railway power supply system is the secondary transmission bus of the step-up transformer of the power plant. The whole system consists of dedicated high-voltage transmission lines, traction substations, feeders, contact system, return lines and electric locomotives. Its structure is shown in Figure 1. Among them, the traction substation is powered by a two-phase 110 kV bus, and a traction transformer is used to reduce the voltage to a voltage level suitable for the operation of electric locomotives. The circuit is then formed through the contact system, electric locomotives, tracks and return lines.

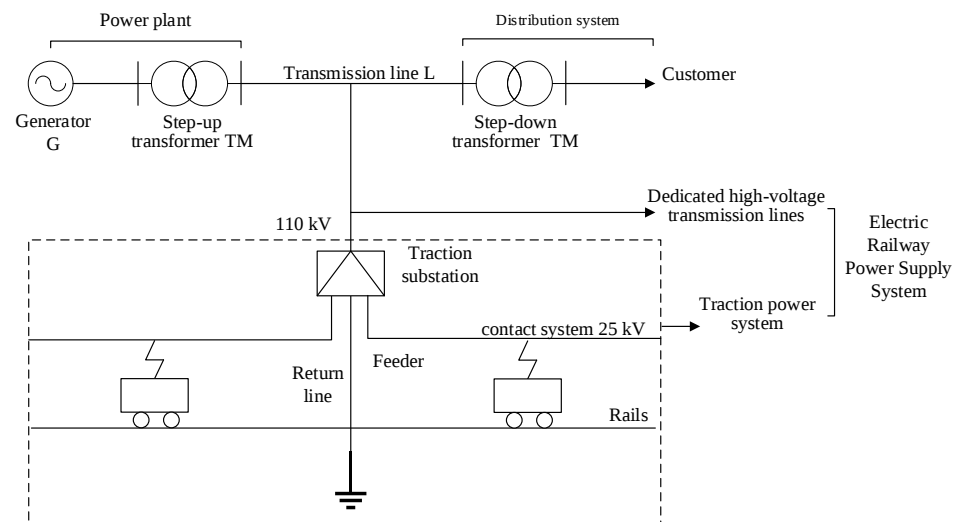


Figure 1. Electric railway power supply system.

2.2. Harmonic Impedance Study

Figure 1 shows the structure diagram of the electrified railway system. Among them, the traction substation and rolling stock on the secondary side can be equivalent to the customer side, and the 110 kV busbar and the upstream system can be equivalent to the utility side. In traction power systems, the Norton equivalent model is one of the most commonly used equivalents, as shown in Figure 2, where $\dot{Z}_s$ and $\dot{Z}_c$ represent the equivalent harmonic impedances on the utility side and customer side, respectively. $\dot{I}_s$ represents the harmonic current value emitted by the utility side. $\dot{I}_c$ represents the harmonic current value emitted by the customer side.

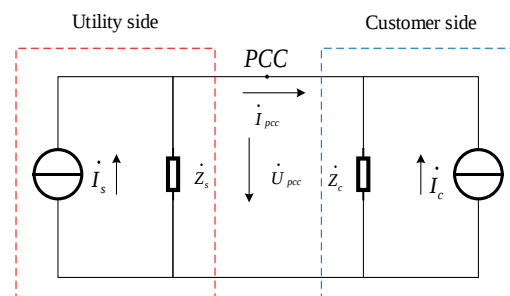


Figure 2. Equivalent circuit in harmonic impedance study.

According to the basic circuit principle, the following equations can be obtained:

$$\dot{U}_{pcc} = \dot{I}_{pcc}\dot{Z}_s + \dot{I}_s\dot{Z}_s \quad (1)$$

$$\dot{I}_{pcc} = \dot{I}_{pcc}\dot{Z}_c - \dot{I}_c\dot{Z}_c \quad (2)$$

where $\dot{U}_{pcc}$ is the harmonic voltage measured at the PCC; $\dot{I}_{pcc}$ is the harmonic current measured at the PCC. Arranging Equations (1) and (2), the following equations can be obtained:

$$\dot{U}_{pcc} = \frac{\dot{Z}_s\dot{Z}_c}{\dot{Z}_s + \dot{Z}_c}\dot{I}_s + \frac{\dot{Z}_s\dot{Z}_c}{\dot{Z}_s + \dot{Z}_c}\dot{I}_c \quad (3)$$

$$\dot{I}_{pcc} = \frac{\dot{Z}_s}{\dot{Z}_s + \dot{Z}_c}\dot{I}_s - \frac{\dot{Z}_c}{\dot{Z}_s + \dot{Z}_c}\dot{I}_c \quad (4)$$

From Equations (3) and (4), it can be seen that there is a nonlinear relationship between the system harmonic impedance $\dot{U}_{pcc}$ and $\dot{I}_{pcc}$ and $\dot{Z}_s$, and GMR has good accuracy and generalization ability for the regression estimation of nonlinear systems. The internal relationship between them can be obtained through GMR, and then the harmonic impedance on the utility side can be estimated.

2.3. Gaussian Mixture Model (GMM)

The main idea of GMM is to classify the parameters of different probability density functions and accurately quantify things. Assuming that each sample of the GMM represents a Gaussian distribution, the GMM model can be regarded as a weighted average of multiple Gaussian distributions [31], that is, multiple single Gaussian models are weighted to approximate an arbitrary continuous distribution.

The probability density function of the GMM model of the random variable x is:

$$p(x) = \sum_{r=1}^R \varphi_r N(x|\mu_r, \Sigma_r) \quad (5)$$

where r represents the number of mixture Gaussian components, $r = 1, 2, \dots, R$; φ_r represents the mixture weight of each Gaussian component, and $0 \leq \varphi_r \leq 1$, $\sum_{r=1}^R \varphi_r = 1$; $N(x|\mu_r, \Sigma_r)$ represents the distribution of the r -th Gaussian component; μ_r represents the mean value of the r -th Gaussian component; the covariance of the r -th Gaussian component is denoted by Σ_r . The dimension of the random variable x determines the dimension of $N(x)$, and the proportion of each Gaussian component in the mixture of Gaussians is determined by the mixture weight φ_r .

The hyperparametric solution of GMM usually uses the expectation maximization (EM) algorithm. The core idea is to approach the objective function by continuously maximizing the lower bound function. First, the values of the model parameters are estimated from the known data samples. Then, estimate the value of missing data according to the estimated parameter value. Finally, the parameter values are re-estimated based on the estimated missing data values and the known sample data values. The above steps are repeated continuously, and the iterations are repeated until the final convergence ends.

The core of the EM algorithm is mainly divided into two steps:

E-step: Guess the value of estimated $z^{(i)}$, and predict that the sample parameters may come from a Gaussian component.

$$w_j^i = p(z^{(i)} = j|x^{(i)}; \varphi, \mu, \Sigma) = \frac{p(x^{(i)}|z^{(i)} = j; \mu, \Sigma)p(z^{(i)} = j; \varphi)}{\sum_{l=1}^k p(x^{(i)}|z^{(i)} = l; \mu, \Sigma)p(z^{(i)} = l; \varphi)} \quad (6)$$

M – step: Based on the guess of *E – step*, the calculation is performed, and the model parameters are updated.

$$\varphi_j = \frac{1}{m} \sum_{i=1}^m w_j^{(i)} \quad (7)$$

$$\mu_j = \frac{\sum_{i=1}^m w_j^{(i)} x^{(i)}}{\sum_{i=1}^m w_j^{(i)}} \quad (8)$$

$$\Sigma_j = \frac{\sum_{i=1}^m w_j^{(i)} (x^{(i)} - \mu_j)(x^{(i)} - \mu_j)^T}{\sum_{i=1}^m w_j^{(i)}} \quad (9)$$

where φ_j is the weight parameter; μ_j is the mathematical expectation; Σ_j is the covariance matrix. The hyperparameters of the GMM model can be obtained by iteratively calculating *E – step* and *M – step*, which is expressed as: $Q = [\{\varphi_1, \mu_1, \Sigma_1\}, \{\varphi_2, \mu_2, \Sigma_2\}, \dots, \{\varphi_R, \mu_R, \Sigma_R\}]$.

2.4. Gaussian Mixture Regression

The EM algorithm is used to calculate the established GMM model, and the hyperparameter Q of each group of Gaussian distribution is obtained. At the same time, GMM is performed according to the linear combination of Gaussian distribution and Gaussian condition. The data vector consists of input data and output data. The input data consists of a two-dimensional vector, and the output data consists of a one-dimensional vector. Then, the Gaussian mixture joint probability density can be expressed by Equation (10):

$$\begin{cases} f_{XY}(x, y) = \sum_{n=1}^N \varphi_n N(x, y, \mu_n, \Sigma_n) \\ \mu_n = \begin{bmatrix} \mu_{nx} \\ \mu_{ny} \end{bmatrix} \quad \Sigma_n = \begin{bmatrix} \Sigma_{nxx} & \Sigma_{nxy} \\ \Sigma_{nyx} & \Sigma_{nyy} \end{bmatrix} \end{cases} \quad (10)$$

where μ_n represents the two-dimensional mean vector; Σ_n is the covariance vector group. The conditional mean and variance are:

$$\mu_{n(x)} = \mu_{ny} + \Sigma_{nyx} \Sigma_{nxx}^{-1} (x - \mu_{nx}) \quad (11)$$

$$\Sigma_{n(x)} = \Sigma_{nyy} - \Sigma_{nyx} \Sigma_{nxx}^{-1} \Sigma_{nxy} \quad (12)$$

The regression function expression is:

$$\mu_{(x)} = \sum_{n=1}^N \varphi_{n(x)} \mu_{n(x)} \quad (13)$$

The conditional variance function is:

$$v_{(x)} = \sum_{n=1}^N \varphi_{n(x)} [\mu_{n(x)}^2 + \Sigma_n^2] - \left[\sum_{n=1}^N \varphi_{n(x)} \mu_{n(x)} \right] \quad (14)$$

According to Formula (13), a regression model can be obtained, and a new output sample can be obtained by substituting the input sample.

2.5. Calculation of Harmonic Impedance by GMR

In this paper, GMR is used for harmonic impedance calculation. The specific process is as follows:

(1) Collect harmonic samples at PCC and take the vector group consisting of $\dot{U}_{pcc}$ and $\dot{I}_{pcc}$ as the input sample. Meanwhile, the output sample consists of a one-dimensional harmonic vector $\dot{Z}_s$. To sum up, the expressions of input variables and output variables are as follows:

$$x_i = \left\{ \dot{U}_{pcc}, \dot{I}_{pcc} \right\}_{i=1}^n; y_i = \left\{ \dot{Z}_s \right\}_{i=1}^n \quad (15)$$

- (2) Establish the combined density GMM model of $\dot{U}_{pcc}$, $\dot{I}_{pcc}$ and $\dot{Z}_s$.
- (3) Use the EM algorithm to obtain the optimal parameters of the GMM model.
- (4) Input the test data to calculate the conditional mean and variance of the test set shown in Equations (11) and (12).
- (5) According to the calculation results of GMR output, evaluate the indicators of this method, and compare the calculation results of other methods.

In summary, the specific process of the proposed method is shown in Figure 3.

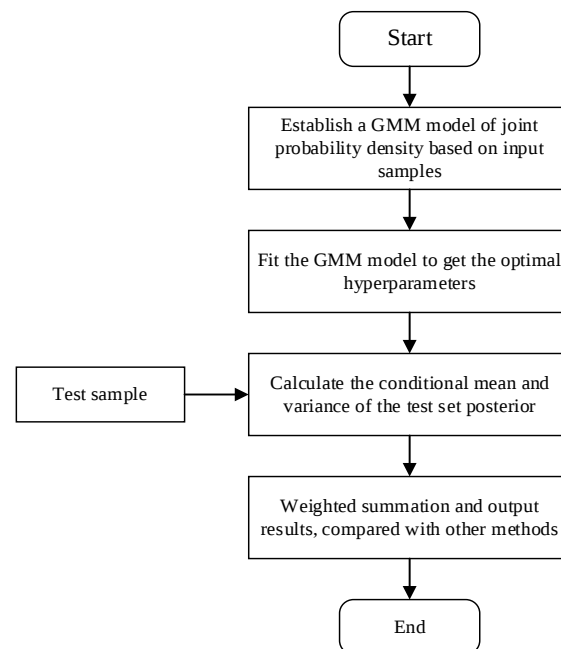


Figure 3. The flowchart of GMR.

3. Simulation Study

In the simulation study, four methods are applied to estimate harmonic impedance, and Matlab software is used to establish a simulation model. Method 1: GMR method, which is the method in the text; Method 2: BLR method; Method 3: IRV method; Method 4: SVM method.

3.1. Simulation I: Harmonic Impedance Estimation with Gaussian Distribution

In the light of the equivalent model in Figure 1, a Gaussian distributed harmonic impedance estimation model is established; meanwhile, the following parameters are set in Simulation I.

- (1) Harmonic impedance on the utility side is $\dot{Z}_s$, which obeys a Gaussian distribution with mean $3 + j8$ and variance $0.3 + j0.7$, and the unit is Ω .
- (2) Harmonic impedance on the customer side is $\dot{Z}_c$, which obeys a Gaussian distribution with mean $25 + j168$ and variance $0.8 + j11$, and the unit is Ω .
- (3) Harmonic current emitted by the customer side is $\dot{I}_c$, which obeys a Gaussian distribution with mean $7 + j12$ and variance $0.4 + j0.7$, and the unit is A.

- (4) Harmonic current emitted by the utility side is $\dot{I}_s$, which is k times larger than $\dot{I}_c$, and the value interval of k is 0.2~1.0, and the value interval is 0.2.

The number of samples collected in simulation I is 1440. Based on the sample set $\{x_n, y_n\}_{n=1}^{1440}$, four methods are used in Simulation I to calculate harmonic impedance. It is worth noting that methods 2 and 3 require the use of continuous calculation, i.e., continuous calculation of harmonic impedance every 200 sampling points.

According to the calculation results of each method, 1000 calculation results are taken and each 100 points are taken as a data group, and a box plot is drawn, where “ n ” represents the number of data sets. The extreme values, medians, quartiles, and outliers for each group of results are visualized in the image. Since k has different values, this paper only shows the calculation result images of each method when $k = 0.2$, as shown in Figures 4–7.

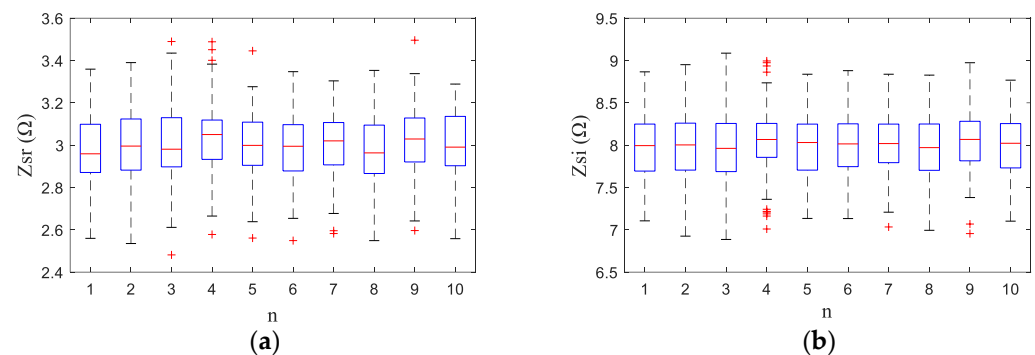


Figure 4. Results of method 1 in simulation I. (a) The result of the real part; (b) The result of the imaginary part.

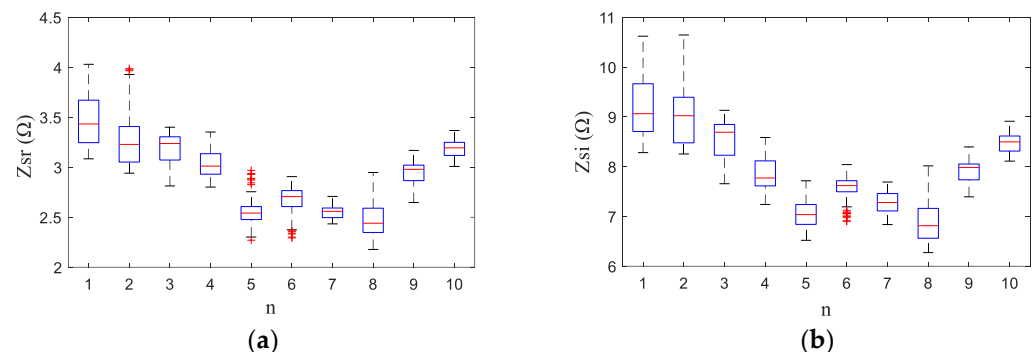


Figure 5. Results of method 2 in simulation I. (a) The result of the real part; (b) The result of the imaginary part.

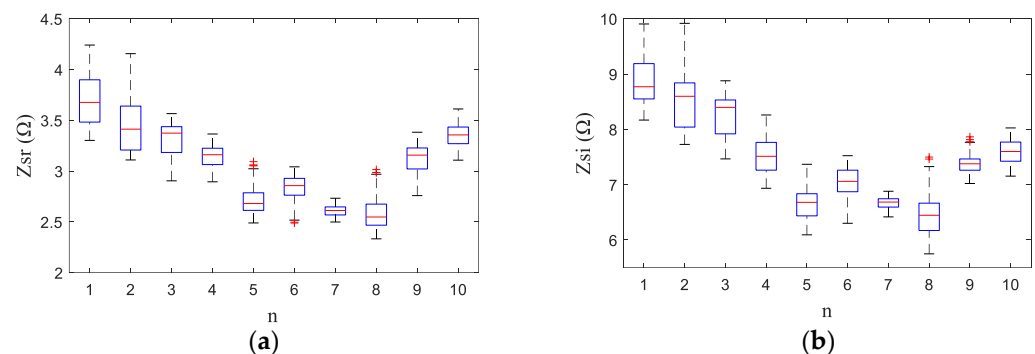


Figure 6. Results of method 3 in simulation I. (a) The result of the real part; (b) The result of the imaginary part.

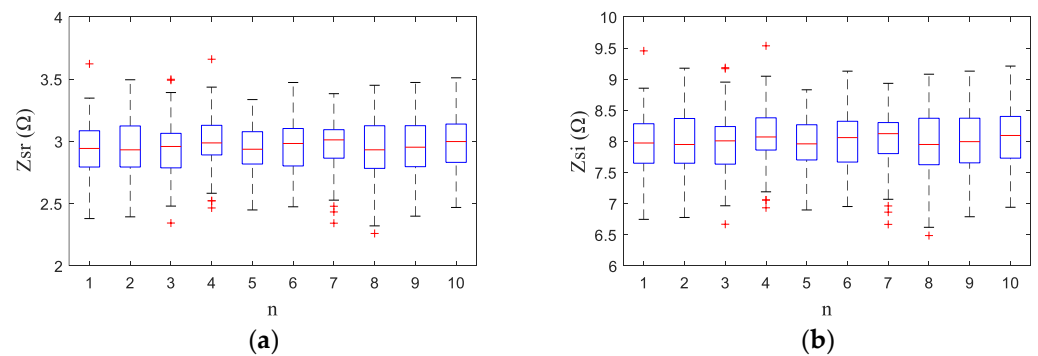


Figure 7. Results of method 4 in simulation I. (a) The result of the real part; (b) The result of the imaginary part.

Comparing Figures 3–6, it is not difficult to see that the results of method 1 and method 4 are less volatile. After careful comparison, it is found that the fluctuation range of the mean value of each segment in method 1 is smaller than that in method 4. Meanwhile, the fluctuation range of extreme values in method 1 is obviously smaller than that in method 4. Therefore, in this simulation, method 1 has the best stability among the four methods. To further emphasize the superiority of the proposed method, the error histogram of several methods when the value of k is different is made, in which the error size takes the absolute value, as shown in Figure 8.

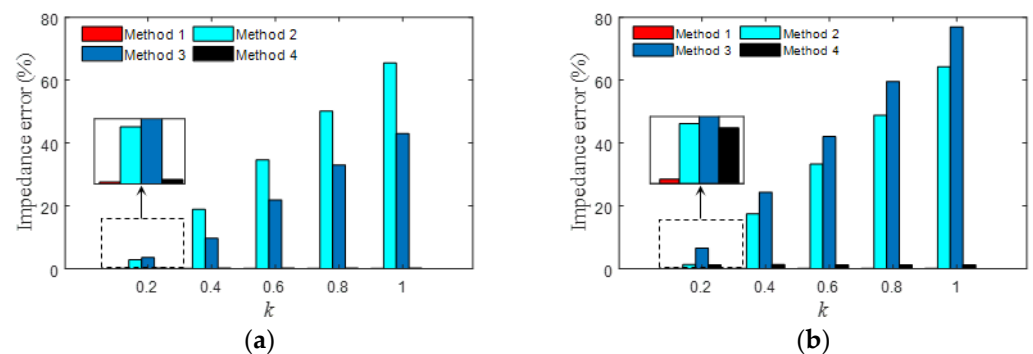


Figure 8. The error when k takes different values. (a) Impedance real part error; (b) Impedance imaginary part error.

As shown in Figure 8, if the value of k increases, the errors of method 2 and method 3 also increase. The error of the real impedance part of method 2 is smaller than that of method 3, and the error of the imaginary impedance part of method 3 is smaller than that of method 2. The error of method 1 and method 4 fluctuates very little when k changes. When $k = 0.2$, the error of method 2 is close to that of method 4. Since method 1 has less error in the model, a small box is used in the figure to show its error. When k takes different values, the mean value of the real part harmonic impedance obtained through various methods is recorded in Table 1, and the mean value of the imaginary part harmonic impedance is recorded in Table 2, in which three decimal places are reserved.

Table 1. Real part average impedance.

k	0.2	0.4	0.6	0.8	1.0
Reference impedance (Ω)	2.999	2.999	2.999	2.999	2.999
Method 1 (Ω)	2.996	2.996	2.996	2.996	2.996
Method 2 (Ω)	3.040	3.522	3.996	4.464	4.925
Method 3 (Ω)	3.195	3.729	4.260	4.786	5.306
Method 4 (Ω)	2.961	2.960	2.961	2.961	2.961

Table 2. Imaginary part average impedance.

k	0.2	0.4	0.6	0.8	1.0
Reference impedance (Ω)	7.995	7.995	7.995	7.995	7.995
Method 1 (Ω)	7.988	7.988	7.988	7.989	7.989
Method 2 (Ω)	8.223	9.506	10.766	12.005	13.226
Method 3 (Ω)	7.711	8.769	9.746	10.636	11.434
Method 4 (Ω)	8.012	8.012	8.012	8.012	8.012

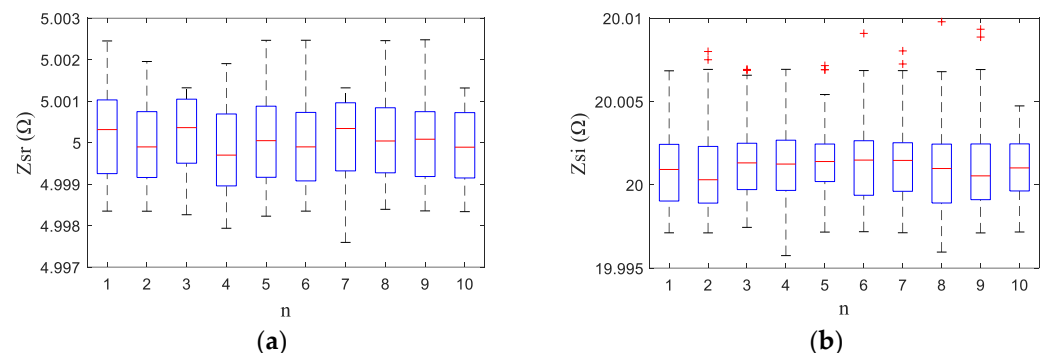
Based on the results in Tables 1 and 2, it is clear that, method 1 is the most stable and most accurate of the several methods. The real part and imaginary part of the calculation result of method 2 both increase with the increase in k value. In the error results of method 2, when $k = 1$, the maximum error of the real part is 64%, and the maximum error of the imaginary part is 65%. The trend of the calculation results of method 3 is consistent with that of method 2. When $k = 1$, the real part error of method 3 reaches the maximum value of 77%, and the imaginary part error reaches the maximum value of 43%. Method 4 is also very stable in this model, but slightly less accurate than method 1.

3.2. Simulation II: Uniformly Distributed Harmonic Impedance Estimation

In the light of the equivalent model in Figure 2, a uniformly distributed harmonic impedance estimation model is established. At this time, there is no correlation between the harmonic sources at both ends of the system. The following parameters are set in Simulation II.

- (1) Harmonic impedance on the utility side is $\dot{Z}_s$, which is uniformly distributed with mean $5 + j20$. Add a standard deviation of 0.01 to the real part and a standard deviation of 0.034 to the imaginary part, and the unit is Ω .
- (2) Harmonic impedance on the customer side is $\dot{Z}_c$, which is uniformly distributed with mean $40 + j296$. Add a standard deviation of 0.662 to the real part and a standard deviation of 1.664 to the imaginary part, and the unit is Ω .
- (3) Harmonic current emitted by the utility side is $\dot{I}_s$, which is set to a fixed value $0.75 + j0.75$, and the unit is A.
- (4) Harmonic current emitted by the customer side is $\dot{I}_c$, which is uniformly distributed with mean $3.73 + j3.74$. Add a standard deviation of 0.242 to the real part and a standard deviation of 0.238 to the imaginary part and the unit is A.

Consistent with Simulation I, the harmonic impedance in Simulation II can be calculated in four ways. Similarly, the calculation uses continuous calculation every 200 sampling points. In the calculation results of each method, take 1000 calculation results and use every 100 points as a data group to draw a box plot as shown in Figures 9–12.

**Figure 9.** Results of method 1 in simulation II. (a) Result of real part; (b) Result of imaginary part.

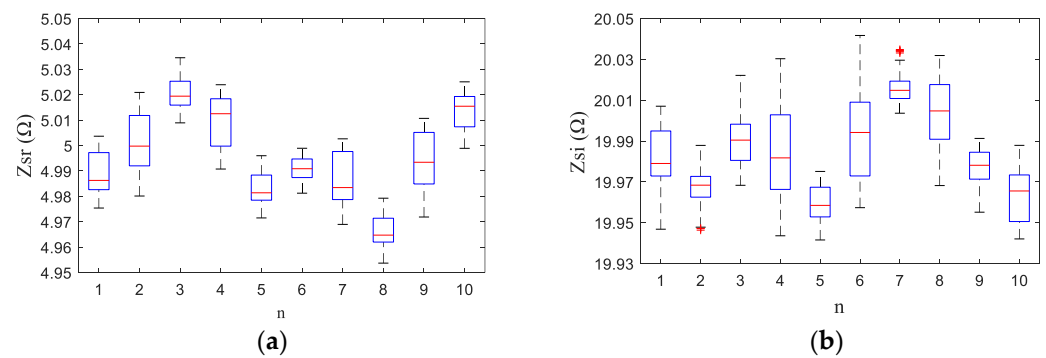


Figure 10. Results of method 2 in simulation II. (a) Result of real part; (b) Result of imaginary part.

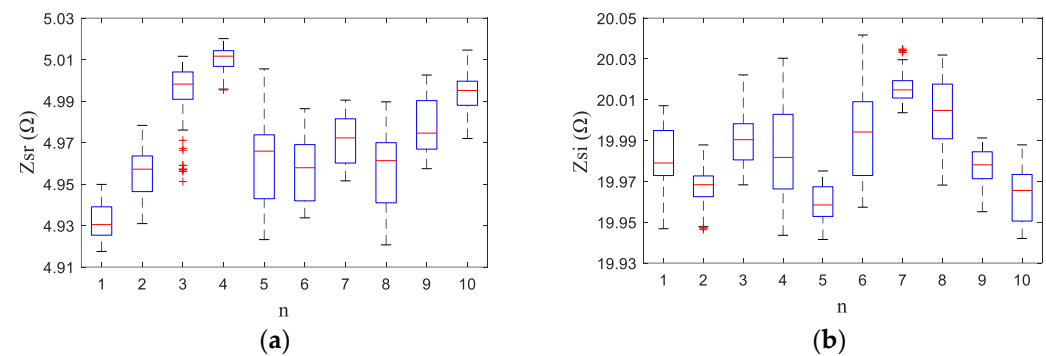


Figure 11. Results of method 3 in simulation II. (a) Result of real part; (b) Result of imaginary part.

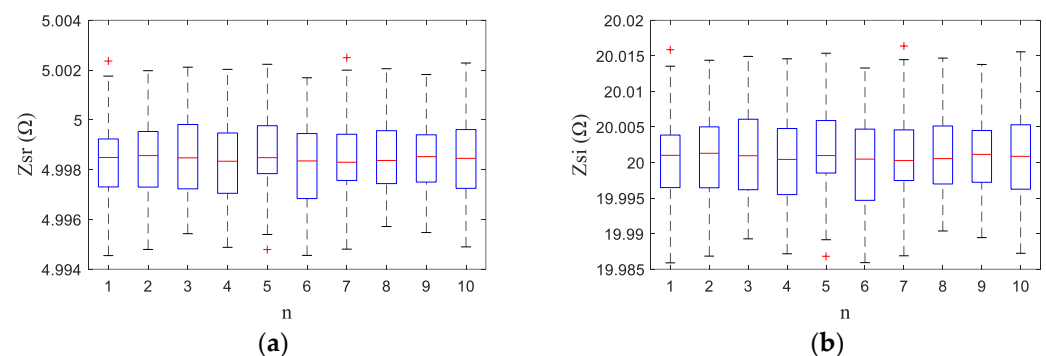


Figure 12. Results of method 4 in simulation II. (a) Result of real part; (b) Result of imaginary part.

Comparing Figures 8–11, method 1 and method 4 are relatively stable. Although there are some outliers in the imaginary part of method 1, the fluctuation amplitude is still smaller than method 4, and the volatility of method 2 and method 3 is larger. To further highlight the superiority of the proposed method in this paper, the error histogram of several methods is drawn, in which the absolute value of the error is taken, as shown in Figure 12.

Figure 13 shows that methods 1 and 4 are more accurate than methods 2 and 3. It is worth noting that the errors of the four methods in the second simulation are very small, all less than 1%. The average impedance values calculated by several methods in Simulation II are shown in Table 3, and the data in the table are kept to three decimal places.

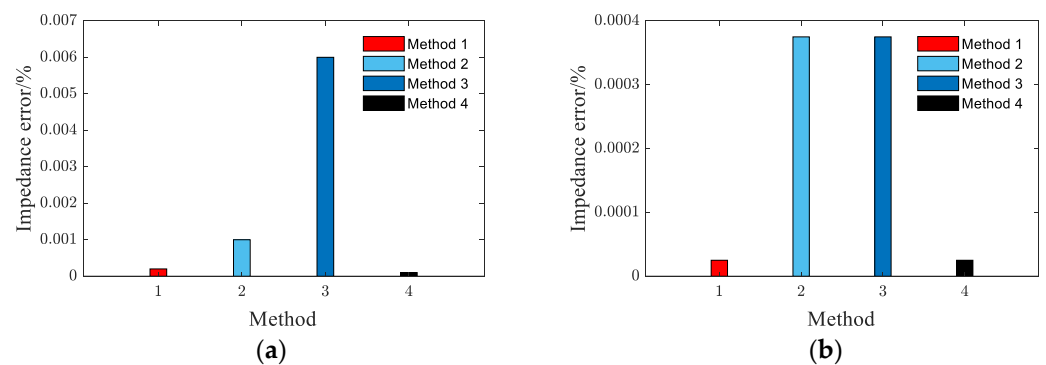


Figure 13. The error of simulation II. (a) Impedance real part error; (b) Impedance imaginary part error.

Table 3. Average impedance.

Average Impedance (Ω)	Real Part Impedance	Imaginary Part Impedance
Reference impedance	4.999	20.002
Method 1	5.000	20.001
Method 2	4.994	19.987
Method 3	4.969	19.987
Method 4	4.998	20.001

Longitudinal comparison of data in Tables 1–3 in Simulation I and Simulation II shows that method 1 (the method in this paper) has significant superiority in estimating the harmonic impedance of Gaussian distribution. In simulation II, a good result is obtained when the sample fitting degree is not as good as Gaussian distribution, which demonstrates the generalization performance of this method. The results of method 2 and method 3 are closer in simulation II. The change trend of the two methods is the same as that of Simulation II, and with the increase in k , the error also increases. The reason for this phenomenon is that there is a correlation between the harmonic sources at both ends of the system. The results of method 4 demonstrate the advantages of the SVM method in simulations I and II. When there is a correlation between the harmonic sources on the utility side and the customer side, it will not be significantly affected, and is only slightly lower than method 1 in terms of accuracy and stability.

4. Case Analysis

Based on the 110kV traction power supply substation, the power quality analyzer is capable of accurately collecting harmonic voltages and currents at the PCC. A schematic diagram of harmonic measurement is shown in Figure 14. In the figure, the fundamental wave voltage amplitude test error of the PQ monitor is: $<\pm 0.5\%$ of Full Scale (F.S.), and the fundamental wave current amplitude test error is: $<\pm 1.0\%$ F.S.

The unit measurement time interval is 3 s, and the total measurement time is 10 h. fast Fourier transform is used to obtain each harmonic, and this paper takes the 3rd harmonic as an example. The traction power supply load will inevitably have low load points and no-load points. The fundamental current is analyzed to remove such points that affect the calculation. Figure 15 shows 600 load samples, which were used as experimental targets, and the third harmonic voltage and current with load are represented. Figure 16 shows the phase angles of the third harmonic voltage and current with load, expressed in the radian system.

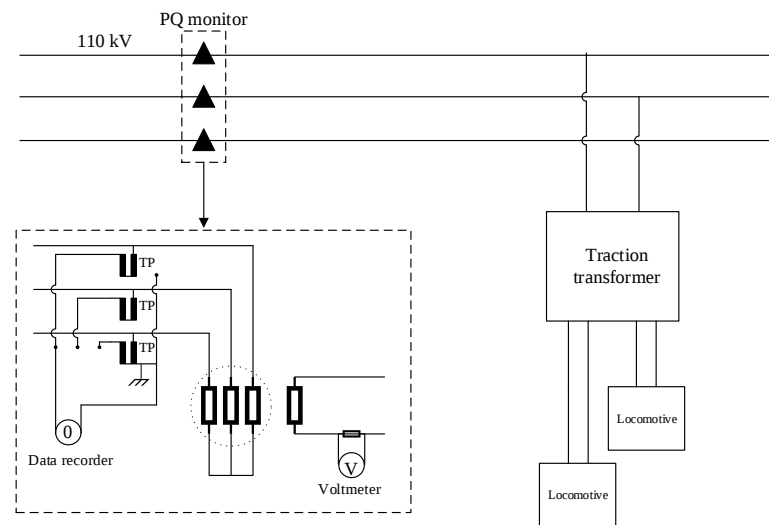


Figure 14. Schematic diagram of harmonic measurement.

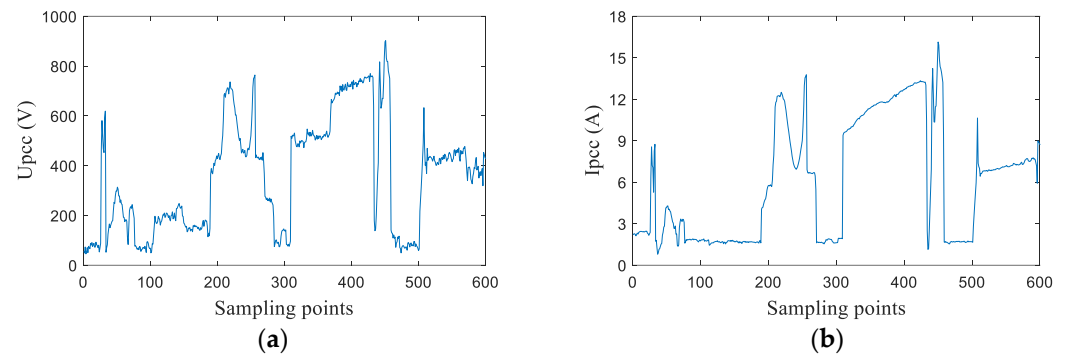


Figure 15. Third harmonic with load. (a) Harmonic voltage; (b) Harmonic current.

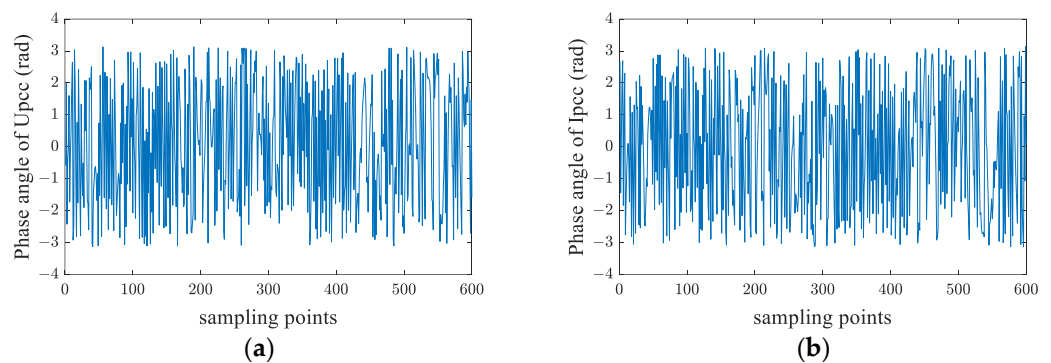


Figure 16. Phase angle of the 3rd harmonic with load. (a) Harmonic voltage; (b) Harmonic current.

In the case analysis, four methods consistent with the simulation study are used to calculate the 3rd harmonic impedance. The differences of each method in case analysis should be more directly reflected, and the results should be displayed by three-dimensional images. The differences of each method in the case study should be presented more directly, and the results should be presented in the form of three-dimensional images. Mean value, mean error, standard deviation, maximum deviation of mean and calculation time are used to evaluate the results. The calculation formula of the evaluation index is as follows:

$$Error_{ave} = \frac{\sum_{i=1}^N |Z_{s,i} - \overline{Z}_s|}{N} \quad (16)$$

$$Error_{std} = \sqrt{\frac{1}{N} \sum_{i=1}^N |Z_{s,i} - \bar{Z}_s|} \quad (17)$$

$$Error_{max} = \max |Z_{s,i} - \bar{Z}_s| \quad (18)$$

Among them, $Error_{ave}$ represents the average error; $Error_{std}$ represents the standard deviation; $Error_{max}$ represents the maximum deviation of the mean; $Z_{s,i}$ represents the calculated harmonic impedance; $\bar{Z}_s$ represents the average harmonic impedance of this method.

In method 1, the harmonic impedance of the utility side is estimated by Gaussian mixture regression, and the result is obtained by interpolation. Figure 17 shows the calculation results of method 1, while Table 4 shows the index parameters of method 1.

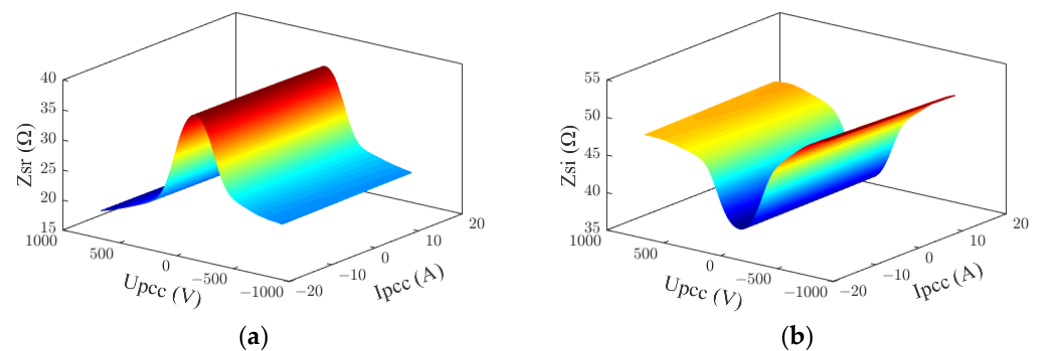


Figure 17. Results of method 1 in 3rd harmonic. (a) Result of real part; (b) Result of imaginary part.

Table 4. Evaluation indexes of method 1 in 3rd harmonic.

Evaluation Indexes	Mean (Ω)	$Error_{ave}$ (Ω)	$Error_{std}$ (Ω)	$Error_{max}$ (Ω)	t (s)
Real part	28.83	5.71	6.27	11.08	29
Imaginary part	42.83	3.91	4.24	8.76	28

Through methods 2, 3 and 4, the real and imaginary parts of harmonic impedance are estimated, as shown in Figures 18–20, respectively. The parameters of various indicators are shown in Tables 5–7.

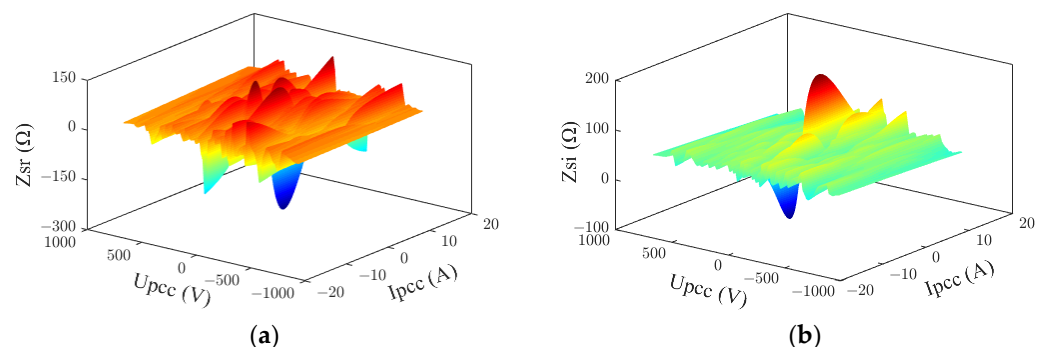


Figure 18. Results of method 2 in 3rd harmonic. (a) Result of real part; (b) Result of imaginary part.

Table 5. Evaluation indexes of method 2 in 3rd harmonic.

Evaluation Indexes	Mean (Ω)	$Error_{ave}$ (Ω)	$Error_{std}$ (Ω)	$Error_{max}$ (Ω)	t (s)
Real part	26.44	19.66	25.43	65.20	1.7
Imaginary part	42.77	9.28	12.32	36.99	1.7

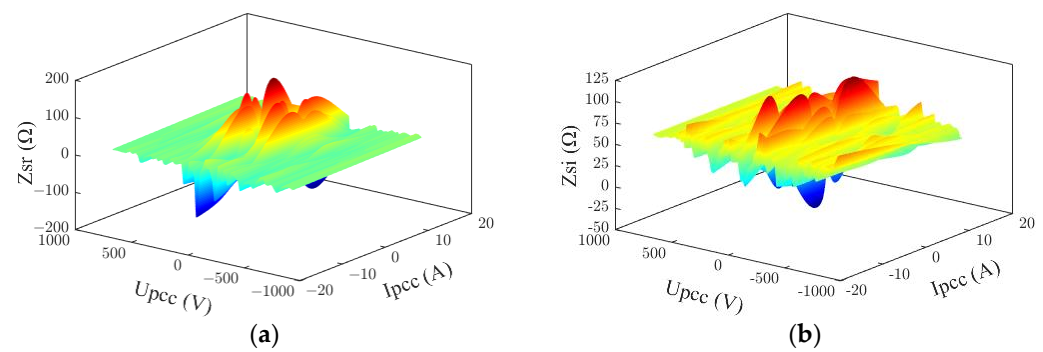


Figure 19. Results of method 3 in 3rd harmonic. (a) Result of real part; (b) Result of imaginary part.

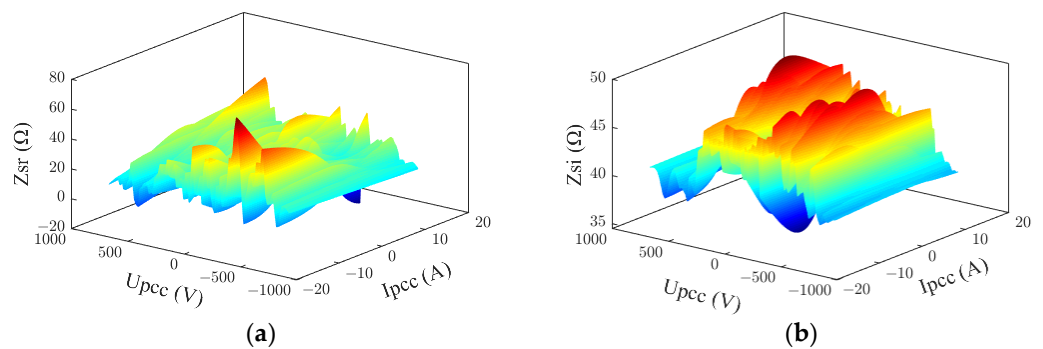


Figure 20. Results of method 4 in 3rd harmonic. (a) Result of real part; (b) Result of imaginary part.

Table 6. Evaluation indexes of method 3 in 3rd harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	22.34	23.08	31.25	106.42	1.2
Imaginary part	45.51	12.94	18.23	59.68	1.2

Table 7. Evaluation indexes of method 4 in 3rd harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	22.10	7.18	8.27	20.89	3.9
Imaginary part	42.01	2.36	2.71	6.85	3.9

By comparing Figures 17–20 and Tables 4–7, it can be seen that each evaluation index of method 1 is significantly better than those of the other three methods, except the calculation time. Due to the measured data from the actual traction power supply system, the linearity of the data is low, and the result of method 2 has a very large fluctuation, which is greatly affected by the background harmonics. Method 3 is the worst among the methods used in this paper. The maximum deviation of the mean of the real part is nearly 4 times higher than its mean, and the maximum deviation of the mean of the imaginary part is nearly 1 times higher than its mean. Because the harmonic source on the utility side has a certain relationship with the harmonic source on the customer side, this method is abnormal. In method 4, some peak values appear in the operation, falling into the local optimum, and the results show great instability.

In order to further reflect the universality of the method in this paper, the 7th harmonic of the measured data is analyzed again. The 7th harmonic impedance obtained by the four methods is shown in Figures 21–24. At the same time, the parameters of the 7th harmonic impedance of the four methods are shown in Tables 8–11.

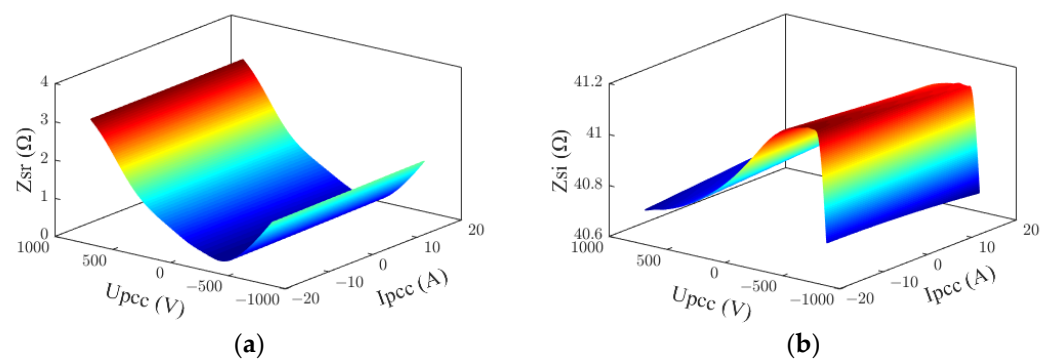


Figure 21. Results of method 1 in 7th harmonic. (a) Real part; (b) Imaginary part.

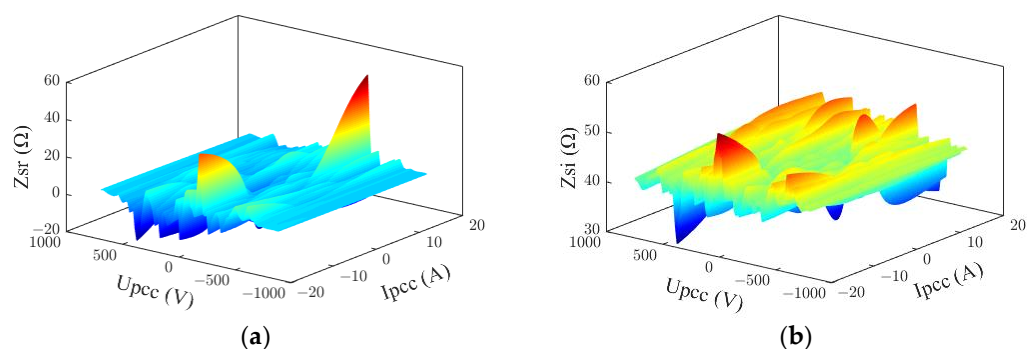


Figure 22. Results of method 2 in 7th harmonic. (a) Real part; (b) Imaginary part.

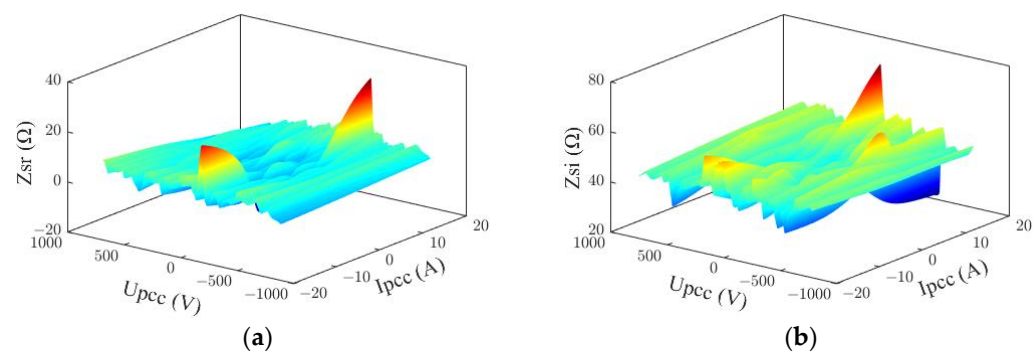


Figure 23. Results of method 3 in 7th harmonic. (a) Real part; (b) Imaginary part.

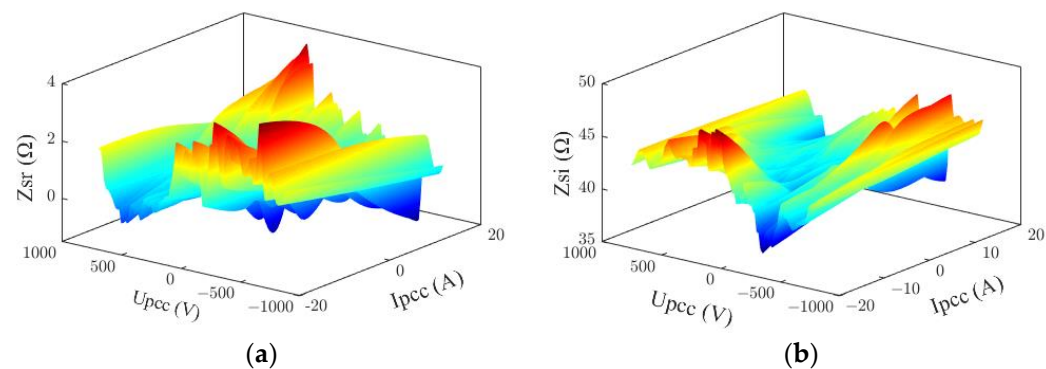


Figure 24. Results of method 4 in 7th harmonic. (a) Real part; (b) Imaginary part.

Table 8. Evaluation indexes of method 1 in 7th harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	0.863	0.436	0.560	2.303	1.4
Imaginary part	40.936	0.094	0.109	0.240	1.4

Table 9. Evaluation indexes of method 2 in 7th harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	0.877	2.905	4.439	15.854	1.5
Imaginary part	41.234	1.900	2.255	6.241	1.5

Table 10. Evaluation indexes of method 3 in 7th harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	2.681	2.220	2.855	12.832	1.1
Imaginary part	42.314	2.768	3.671	8.689	1.1

Table 11. Evaluation indexes of method 4 in 7th harmonic.

Evaluation Indexes	Mean (Ω)	Error _{ave} (Ω)	Error _{std} (Ω)	Error _{max} (Ω)	t (s)
Real part	2.049	0.428	0.588	2.654	4.6
Imaginary part	41.374	0.902	1.239	5.589	4.6

Previously, the literature [20] proposed a method for harmonic impedance estimation via O-GPR. This method uses the Bayesian framework to optimize Gaussian process regression, and the effect is very significant. It is not difficult to find that GMR results are similar to O-GPR results after visualizing the results. Likewise, the results of the two methods in the calculation of harmonic impedance are symmetrical on the positive and negative half axis. In terms of mean error, standard deviation and mean maximum deviation, GMR is close to O-GPR. However, in terms of running time, GMR is significantly better than O-GPR.

5. Conclusions

The background harmonics of the railway power supply system are characterized by strong fluctuation and correlation between harmonics sources. The existing methods of harmonic impedance estimation may be biased. A GMR method for estimating the harmonic impedance of the utility side of the railway power supply system is presented in this paper. First, the GMM model is constructed and the hyperparameters are solved by the EM algorithm to obtain the GMR model and the weights of each Gaussian component. Then, substitute the harmonic voltage and harmonic current data to solve the harmonic impedance. Finally, the validity is verified by two different harmonic impedance estimation models and an actual case of a railway power supply system. The following conclusions are drawn:

- (1) This method is not sensitive to the correlation of harmonic sources and is suitable for occasions where the correlation of harmonic sources is strong. When the correlation between the two harmonic sources changes, the method can obtain better evaluation results and has good generalization performance.
- (2) Compared with previous similar methods, such as the GPR method, its timeliness is also better optimized.
- (3) In the Gaussian distributed harmonic and uniformly distributed harmonic impedance models, the stability performance of the GMR method is better than that of the BLR, IRV, and SVM methods compared in this paper.
- (4) In the case analysis, the average error and standard deviation of this method are about one-fifth to one-third of those of the binary linear regression (BLR) method and the

independent random vector (IRV) method. At the same time, its index is slightly better than that of the support vector machine (SVM) method.

Author Contributions: W.T. and Y.X. mainly carried out, completed this work, simulated and analyzed; Y.X. provided and analyzed the field data; Y.X. supplemented and perfected this work in the later stage; W.T. wrote the text; Y.X. and W.T. reviewed the text; all authors revised and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported in part by the Key Scientific Research fund of Xihua University (Z1520909), the Southwest jiaotong University Traction Power State Key Laboratory Open Issue (TPL1907).

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Zihan, H.; Can, W.; Yonghua, S. Frequency Regulation From Electrified Railway. *IEEE Trans. Power Syst.* **2022**, *37*, 2414–2431.
2. Chen, L.; Chen, M.; Chen, Y.; Chen, Y.; Cheng, Y.; Zhao, N. Modelling and control of a novel AT-fed co-phase traction power supply system for electrified railway. *Int. J. Electr. Power Energy Syst.* **2021**, *125*, 106405. [\[CrossRef\]](#)
3. Zhao, E.; Han, Y.; Lin, X.; Liu, E.; Yang, P.; Zalhaf, A.S. Harmonic characteristics and control strategies of grid-connected photovoltaic inverters under weak grid conditions. *Int. J. Electr. Power Energy Syst.* **2022**, *148*, 108280. [\[CrossRef\]](#)
4. Farhad, S.; Kaveh, M.; Wolfgang, S. On the dominant harmonic source identification—Part I: Review of methods. *IEEE Trans. Power Deliv.* **2018**, *33*, 1268–1277.
5. Xiao, X.; Zhou, R.; Ma, X.; Xu, R. A novel method for estimating utility harmonic impedance based probabilistic evaluation. *IET Gener. Transm. Distrib.* **2022**, *16*, 1438–1448. [\[CrossRef\]](#)
6. Xian, Z.; Xianyong, X.; Yang, W. Harmonic impedance measurement based on an improved binary regression algorithm and dynamic time warping distance. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106907.
7. Monteiro, H.L.; Duque, C.A.; Silva, L.R.; Meyer, J.; Stiegler, R.; Testa, A.; Ribeiro, P.F. Harmonic impedance measurement based on short time current injections. *Electr. Power Syst. Res.* **2017**, *148*, 108–116. [\[CrossRef\]](#)
8. Sun, Y.; Li, S.; Xu, Q.; Xie, X.; Jin, Z.; Shi, F.; Zhang, H. Harmonic contribution evaluation based on the distribution-level pmus. *IEEE Trans. Power Deliv.* **2021**, *36*, 909–919. [\[CrossRef\]](#)
9. Sumner, M.; Palethorpe, B.; Thomas, D.W.; Zanchetta, P.; Di Piazza, M.C. A technique for power supply harmonic impedance estimation using a controlled voltage disturbance. *IEEE Trans. Power Electron.* **2002**, *17*, 207–215. [\[CrossRef\]](#)
10. Farhad, S.; Kaveh, M.; Wolfgang, S. On the dominant harmonic source identification—Part II: Application and interpretation of methods. *IEEE Trans. Power Deliv.* **2018**, *33*, 1278–1287.
11. Papič, I.; Matvoz, D.; Špelko, A.; Xu, W.; Wang, Y.; Mueller, D.; Miller, C.; Ribeiro, P.F.; Langella, R.; Testa, A. A benchmark test system to evaluate methods of harmonic contribution determination. *IEEE Trans. Power Deliv.* **2019**, *34*, 23–31. [\[CrossRef\]](#)
12. Yang, H.; Pirote, P.; Robert, A. Assessing the harmonic emission level from one particular customer. In Proceedings of the PQA'94, Amsterdam, The Netherlands, 24–27 October 1994; pp. 160–166.
13. Zang, T.; He, Z.; 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. [\[CrossRef\]](#)
14. Shu, Q.; Wu, Y.; Xu, F.; Zheng, H. Estimate utility harmonic impedance via the correlation of harmonic measurements in different time intervals. *IEEE Trans. Power Deliv.* **2020**, *35*, 2060–2067. [\[CrossRef\]](#)
15. Xiao, X.; Zheng, X.; Wang, Y.; Xu, S.; Zheng, Z. A Method for Utility Harmonic Impedance Estimation Based on Constrained Complex Independent Component Analysis. *Energies* **2018**, *11*, 2247. [\[CrossRef\]](#)
16. Tang, Z.; Li, H.; Xu, F.; Shu, Q.; Jiang, Y. A harmonic impedance estimation method based on the cauchy mixed model. *Math. Probl. Eng.* **2020**, *2020*, 1580475. [\[CrossRef\]](#)
17. Serfontein, D.; Rens, J.; Botha, G.; Desmet, J. Continuous event-based harmonic impedance assessment using online measurements. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 2214–2220. [\[CrossRef\]](#)
18. Fei, Q.; Li, J.-W.; Li, Y.-G.; Sun, W.; Li, Z.-L. 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.
19. Shu, Q.; Fan, Y.; Xu, F.; Wang, C.; He, J. A harmonic impedance estimation method based on AR model and Burg algorithm. *Electr. Power Syst. Res.* **2022**, *202*, 107568. [\[CrossRef\]](#)
20. Zhao, X.; Yang, H. A new method to calculate the utility harmonic impedance based on FastICA. *IEEE Trans. Power Deliv.* **2016**, *31*, 381–388. [\[CrossRef\]](#)
21. Xia, Y.; Tang, W. Study on the estimation of harmonic impedance based on Bayesian optimized Gaussian process regression. *Int. J. Electr. Power Energy Syst.* **2022**, *142*, 108294. [\[CrossRef\]](#)

22. Huang, X.; Nie, P.; Gong, H. A new assessment method of customer harmonic emission level. In Proceedings of the 2010 Asia-Pacific Power and Energy Engineering Conference, Chengdu, China, 28–31 March 2010; pp. 1–5.
23. Xu, F.; Wang, C.; Shu, Q.; Ma, Z.; Zheng, H.; Peng, A. Study on the estimation of utility harmonic impedance based on minimum norm of impedance difference. *IEEE Access* **2020**, *8*, 207389–207396. [[CrossRef](#)]
24. De Andrade, G.V., Jr.; Naidu, S.R.; Neri, M.G.G.; da Costa, E.G. Estimation of the Utility's and Consumer's contribution to harmonic distortion. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 3817–3823. [[CrossRef](#)]
25. Pan, P.; Hu, H.; Yang, X.; Blaabjerg, F.; Wang, X.; He, Z. Impedance measurement of traction network and electric train for stability analysis in high-speed railways. *IEEE Trans. Power Electron.* **2018**, *33*, 10086–10100. [[CrossRef](#)]
26. Xiao, X.; Yang, H. The method of estimating customers harmonic emission level based on bilinear regression. In Proceedings of the 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies, Hong Kong, China, 5–8 April 2004; pp. 662–665.
27. Hui, J.; Yang, H.; Lin, S.; Ye, M. Assessing utility harmonic impedance based on the covariance characteristic of random vectors. *IEEE Trans. Power Deliv.* **2010**, *25*, 1778–1786. [[CrossRef](#)]
28. Xia, Y.; Tang, W.; Lin, X. Assessing the harmonic impedance based on least squares support vector machine. In Proceedings of the 2021 4th International Conference on Energy, Electrical and Power Engineering, Chongqing, China, 23–25 April 2021; pp. 987–992.
29. Hanmin, S.; Jian, X.; Peng, W. Lithium Iron Phosphate battery electric vehicle state-of-charge estimation based on evolutionary gaussian mixture regression. *IEEE Trans. Ind. Electron.* **2017**, *64*, 544–551.
30. Jin, H.; Shi, L.; Chen, X.; Qian, B.; Yang, B.; Jin, H. Probabilistic wind power forecasting using selective ensemble of finite mixture Gaussian process regression models. *Renew. Energy* **2021**, *174*, 1–18. [[CrossRef](#)]
31. Stauffer, C.; Grimson, W.E.L. Learning patterns of activity using real-time tracking. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 747–757. [[CrossRef](#)]