

Analysis of Power System Harmonic Effects Based on Data Features and Gene Expression Programming

Yuhang Xia

Engineering Project Management
Center
Chengdu Power Supply Company
of Sichuan Electric Power
Company
Chengdu, China
yuhang_xia@sina.cn

Song Qing

Chengdu Power Supply Company
Chengdu Power Supply Company
of Sichuan Electric Power
Company
Chengdu, China
2248095592@qq.com

Yingjie Liu

Engineering Project Management
Center
Chengdu Power Supply Company
of Sichuan Electric Power
Company
Chengdu, China
1462331401@qq.com

Yijun Liu

High Voltage Cable Operation
and Maintenance Center
Chengdu Power Supply Company
of Sichuan Electric Power
Company
Chengdu, China
390892761@qq.com

Lin Zhou

Engineering Project Management Center
Chengdu Power Supply Company of Sichuan
Electric Power Company
Chengdu, China
uhey94@163.com

Jinshuo Jia

Group of Artificial Intelligence and System
Optimization
Chongqing University of Posts and
Telecommunications
Chongqing, China
s220101055@stu.cqupt.edu.cn

Zhengwen Huang

Department of Electronic and Computer
Engineering
Brunel University
London, U.K
zhengwen.huang@brunel.ac.uk

Abstract—In order to quantitatively recognize the harmonic impacts of harmonic source users at the common connection point (PCC), a method to calculate the system harmonic impedance and quantify the harmonic impacts by using gene expression programming to analyses data features was proposed in this paper. Firstly, the harmonic voltage and current data satisfying the analysis conditions were selected by using the time series segmentation method. The system harmonic impedance phase angle was obtained by using the characteristic of zero covariance of independent random variables, and then the harmonic phase angle was implanted into the existing data correlation analysis model. When the user harmonic current fluctuated, the accurate system harmonic impedance module was calculated by the user harmonic current fluctuation. When the user's harmonic current was stable, the harmonic impacts was quantified by the fluctuation of system harmonic voltage. In this paper, the user harmonic impedance was also given in thoughts in the entire process, which reduced the error caused by ignoring the user harmonic impedance in the traditional method. Some potential Evolutionary Algorithm based solutions which could be further applied in this domain was also reviewed. As a particular novelty of this work, the possibility and feasibility of employing gene expression programming into the conventional data correlation analysis work in power system. Simulation analysis and practical engineering examples showed that this method can effectively suppress the influence of system harmonic change and user harmonic impedance compared with the existing methods and obtained more accurate harmonic responsibility division results.

Keywords—Harmonic impacts, Gene Expression Programming, Data mining, Correlation analysis, System impedance

I. INTRODUCTION

In modern power system, nonlinear loads such as electronic devices, electrical and various electronic equipment are widely used in various sectors of the global industry, and the problem

of harmonic pollution in power systems, especially in distribution networks, has become increasingly serious. When the harmonic current flows into the power grid, harmonic voltage generated at the same time. The voltage distortion at the point of public connection (PCC point), will seriously affect the power quality. So, the valid assessment of the harmonic voltage responsibilities for power grid and consumer at the PCC point will be the premise of managing power quality [1].

At present, the method for determining harmonic responsibilities is mainly based on the estimation of system and user harmonic impedance, which can be mainly classified "intervention" and "non-intervention" method. The "intervention" method artificially injected harmonic current into power system for measuring the harmonic impedance of the grid [2-3]. This method could affect the normal operation of the power system and result large actual error [4]. The "non-intervention" method estimates the harmonic impedance through the measured value of the harmonic voltage and current at the point of common coupling (PCC) to quantitatively divide the harmonic responsibility. It includes: 1) "fluctuation method" [5-6], which was an estimation method based on the characteristic of voltage and current fluctuation at the measured point, while this method requires high accuracy of harmonic parameter and large enough fluctuation of the measured value. 2) Linear regression method: including robust regression method [7-10], etc., the essence of those is to construct the corresponding regression based on equivalent-circuit model with the law of Thevenin or Norton. The equation separates the real part and the imaginary part of the PCC voltage and current phasor to calculate the harmonic impedance. It is not an accurate solution to the original problem by least squares. So, the current linear regression method has the following limitations: the constant item in the regression model (the background harmonic voltage) was actually a large fluctuation, which resulted in a

large regression error [11-13]. The illustrated data correlation analysis method used the user harmonic current change to calculate the system harmonic impedance, but ignored the harmonic phase angle could produce large errors.

Gene expression programming (GEP) was a novel adaptive evolutionary algorithm based on the structure and function of biological genes. GEP absorbed the advantages of genetic algorithms and genetic programming. Its notable feature was that it can use simple coding to solve complex problems. Several literatures had revealed the advantages of GEP in analytical ability of complex data association features. Khan proposed a kind method for optimization correlation among the Hadoop configuration parameters by GEP [14]. Naik built a brand-new equation by using GEP to quantized the correlation between converging angle, width ratio, relative distance, relative depth, aspect ratio and water surface profile [15].

In order to obtain the more accurate system harmonic impedance, this paper fully considered the harmonic phase angle, and then proposed a new technique for quantitative division of harmonic responsibilities, which used the change of system harmonic voltage (background harmonic voltage) to quantitatively calculate the user's harmonic voltage contribution at the PCC, and the user-side harmonic impedance contribution was taken into consideration during the calculation process. This paper also reviewed existing evolutionary algorithm based system modelling and optimization for power system. Theoretical simulation and engineering verification of the method in this paper showed that the proposed method can effectively determine the responsibility for harmonics in the power system with effectively manage the influence of system harmonic and user harmonic impedance changes compared with other existing correlation analysis method, and the evaluation results was accurate enough for further application.

For simplicity, the system impedance and user impedance hereinafter represented system harmonic impedance and user harmonic impedance, respectively.

II. TRADITIONAL CORRELATION ANALYSIS METHODS

The traditional correlation analysis uses the association between the harmonic voltage and current at the PCC, which to calculate the system impedance by system frequency and the user's harmonic current. Figure 1 is an equivalent circuit, and the subscripts 's' and 'c' refer to the system side and the user side, respectively. Obtain the measured value during the measurement period, and use Fourier transform to obtain the h-order harmonic voltage V_{pcc}^h and harmonic current I_{pcc}^h at the PCC point, and draw the image of $|V_{pcc}^h|$ about $|I_{pcc}^h|$ as shown in Figure 2, and use the linear regression method to get:

$$|V_{pcc}^h| = a |I_{pcc}^h| + b \quad (1)$$

According to formula (1), a is equal to the system impedance, and b represents the average value of the harmonic voltage at the PCC point when the user harmonic current is zero ($I_{pcc}^h=0$), which was the system harmonic voltage at the same time.

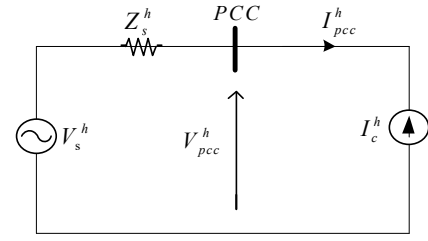


Fig. 1. Equivalent circuit when neglecting the consumer harmonic impedance

III. METHODOLOGY

A. the system harmonic impedance calculation

As seen in (1), harmonic phase angle should be taken into consideration, or it will cause errors in actual calculation. This error can be eliminated by the equation which represents the impact between the hth-harmonic voltage and current at the PCC. As follows.

$$V_{pcc}^h = V_s^h - Z_s^h \times I_{pcc}^h \quad (2)$$

Where, Z_s^h is the system impedance, and V_s^h is the system harmonic voltage. The Equation (2) can be further turn into the scalar form by considering the phase angle of each vector.

$$|V_{pcc}^h| = |V_s^h| \cos(\lambda^h) - |Z_s^h| \times |I_{pcc}^h| \cos(\gamma^h) \quad (3)$$

$$\gamma^h = \alpha^h + \beta^h \quad (4)$$

In the Equation (3) and (4), γ^h is the angle between the vector $-Z_s^h I_{pcc}^h$ and V_{pcc}^h and λ^h , α^h is the phase angle difference between vector V_{pcc}^h and I_{pcc}^h , β^h is the system impedance angle.

It can be seen that $-|Z_s^h| \times |I_{pcc}^h| \cos(\gamma^h)$ and $|V_s^h| \cos(\lambda^h)$ were the projection of $-Z_s^h \times I_{pcc}^h$ and V_s^h on the V_{pcc}^h , respectively.

So α^h can be defined as follows.

$$\alpha^h = \varphi(V_{pcc}^h) - \varphi(I_{pcc}^h) \quad (5)$$

$\varphi(V_{pcc}^h)$ and $\varphi(I_{pcc}^h)$ are the phase angles of vector V_{pcc}^h and I_{pcc}^h respectively.

In this article, the system impedance angle was estimated based on the property that the covariance of independent random variables is zero.

B. Estimation of system impedance angle based on the independent random variables with zero covariance

Based on the sample values, the average value of each quantity in Equation (2) can be calculated.

$$\frac{1}{N} \sum_{i=1}^N V_{pcc}^h = \frac{1}{N} \sum_{i=1}^N V_s^h - Z_s^h \frac{1}{N} \sum_{i=1}^N I_{pcc}^h \quad (6)$$

$$\text{Make } \Delta I_{pcc}^h = I_{pcc}^h - \frac{1}{N} \sum_{i=1}^N I_{pcc}^h, \quad \Delta V_{pcc}^h = V_{pcc}^h - \frac{1}{N} \sum_{i=1}^N V_{pcc}^h \text{ and } \Delta V_s^h = V_s^h - \frac{1}{N} \sum_{i=1}^N V_s^h$$

then Equation (2) minus Equation (6) can get:

$$\Delta V_s^h = \Delta V_{pcc}^h + Z_s^h \Delta I_{pcc}^h \quad (7)$$

For the user impedance Z_c^h is much larger than the system impedance Z_s^h , which can be considered that $I_{pcc}^h \approx I_c^h$. According to the probability theory, the covariance of two independent random vectors is zero [16], so the covariance of I_{pcc}^h and V_s^h were approximately 0. Therefore,

$$\text{Cov}(V_{s0}^h, I_{pcc}^h) = E \left[(I_{pcc}^h - E(I_{pcc}^h))^* (V_s^h - E(V_s^h)) \right] \quad (8)$$

If the sample size is large enough, the expectation can be approximated as the mean, so Equation (8) can be rewritten as follow:

$$\frac{1}{N} \sum_{i=1}^N (\Delta I_{pcc}^h)^* \Delta V_s^h \approx 0 \quad (9)$$

Multiply both sides of Equation (7) by ΔI_{pcc}^h and get the mean value:

$$\frac{1}{N} \sum_{i=1}^N ((\Delta I_{pcc}^h)^* (\Delta V_{pcc}^h + Z_s^h \Delta I_{pcc}^h)) = \frac{1}{N} \sum_{i=1}^N (\Delta I_{pcc}^h)^* \Delta V_s^h \approx 0 \quad (10)$$

Split Equation (10) into two parts, real part and imaginary part, which can get:

$$\tan^{-1}(\beta^h) = \frac{Z_{sy}^h}{Z_{sx}^h} = \frac{\sum^N (\Delta I_{pccx}^h \Delta V_{pccy}^h - \Delta I_{pccy}^h \Delta V_{pccx}^h)}{\sum^N (\Delta I_{pccx}^h \Delta V_{pccx}^h + \Delta I_{pccy}^h \Delta V_{pccy}^h)} \quad (11)$$

The quantities on the right-hand side of the equation can be obtained from sample values. If β^h was obtained, the system impedance also can be calculated.

C. Distinguish of harmonic responsibilities

The traditional division of harmonic responsibility ignores the user impedance, which will cause large errors when the user impedance is small. At the PCC point, to evaluate the user's contribution for the harmonic voltage with the change of V_s^h under the premise of considering the user's impedance was the

key to distinguish harmonic impacts. The relationship between user harmonic voltage V_c^h , V_{pcc}^h and I_{pcc}^h can be described as follow:

$$V_{pcc}^h = V_c^h + Z_c^h I_{pcc}^h \quad (12)$$

Obviously, the main component of I_{pcc}^h comes from $V_c^h \cdot I_{pcc}^h$ and V_c^h are not independent for each other. In order to satisfy the conditions of correlation analysis, substitute I_{pcc}^h with $I_{pcc}^h = I_s^h - (V_{pcc}^h / Z_s^h)$ into Equation (12), then can be obtained,

$$V_{pcc}^h = Z_c^h (I_s^h - V_{pcc}^h / Z_s^h) + V_c^h \quad (13)$$

Which can be transformed,

$$V_{pcc}^h = \frac{Z_c^h Z_s^h}{Z_c^h + Z_s^h} I_s^h + \frac{Z_s^h}{Z_c^h + Z_s^h} V_c^h \quad (14)$$

According the definition of V_c^h and I_s^h , $\frac{Z_s^h}{Z_c^h + Z_s^h} V_{c0}^h$ and I_s^h are mutual independent. Then I_s^h can be calculated by Figure 2.

$$I_s^h = \frac{V_{pcc}^h}{Z_s^h} - I_{pcc}^h \quad (15)$$

Then the harmonic voltage contribution on the user side and the system side at the PCC can be calculated.

$$\%V_{c,pcc}^h = \frac{|((Z_s^h)/(Z_c^h + Z_s^h)) V_c^h|}{V_{pcc}^h} \times 100 \quad (16)$$

$$\%V_{s,pcc}^h = 100 - \%V_{c,pcc}^h \quad (17)$$

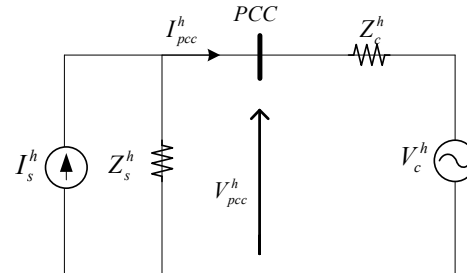


Fig. 2. Equivalent circuit when considering the consumer harmonic impedance

D. AI based modelling solutions

As a complex system, the data mining in power system could benefit from existing conventional/classical evolutionary algorithms Genetic Algorithm (GA) [17], Genetic Programming (GP) [18], Gene Expression Programming (GEP) [19].

The Genetic Algorithm employs bitstreams format container to represent information detected and recognized from the target problem space. It is known as a very good candidate for equally distributed space exploring. It is particularly good at searching an optimized or near optimized value for the system harmonic impedance calculation problem.

The Genetic Programming and Gene Expression Programming provide exceptional level of system modelling support. The data correlation analysis in complex system like power system which requires multi-objectives, accuracy, speed performance, flexibility, etc. optimization support could be fully boosted with the tree structured/formatted solution container. GP has the advantages on complex internal correlation mining. For quantitating the division of harmonic responsibilities, system harmonic impedance calculation problem could take the solution tree to main the near latest information obtained from the real time system [20].

If the searching space is getting complex, the near real time processing speed requirement of power system became another bottle neck. The fixed length of the gene in the chromosome of GEP addressed the bloat problem [21] in GP successfully. The head and tail structure in one gene is maintained and to express one part of the correlation detected from the target power system. The flexible structure (figure 3.) achieved an excellent balance on accuracy, speed performance for maintain the system model of the system harmonic impedance calculation problem. In Ref.[22] GEP was employed to generate an equal model to a sensor equivalent circuit. The method also provides implementation template for the system harmonic impedance calculation. In Ref.[23], a method of auto-modeling recognition was proposed to restructure the particle structure of Electrochemical Impedance Spectroscopy (EIS). The accuracy and optimization level adjust used in that work could give the researcher in a new idea to generate accurate model first then further exploring the best setting of the involved key parameter.

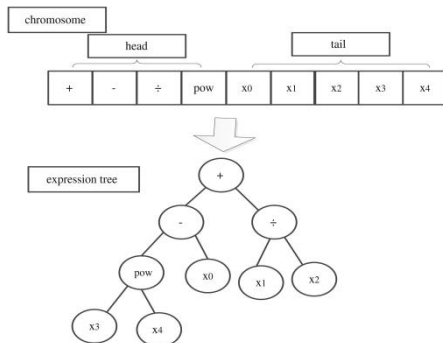


Fig. 3. The gene to expression tree structure

IV. DATA SELECTION SCHEME BASED ON TIME SERIES SEGMENTATION

The calculation of system impedance mainly uses the change of I_c^h , the division of harmonic responsibility uses the change of I_s^h . This paper adopts a data selection scheme based on time series segmentation, that is, to select the period in which each quantity (I_c^h or I_s^h) fluctuates greatly. In practice, a fluctuation threshold (10% in this paper) can be set in advance, and the period when the fluctuation exceeds this threshold is selected as the attention period.

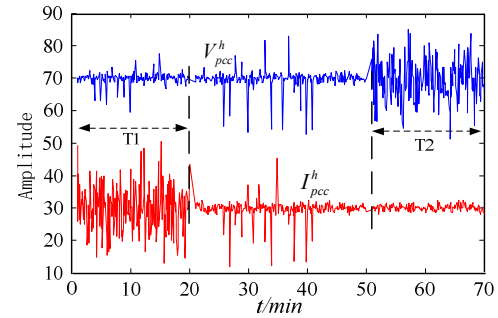


Fig. 4. Time series segmentation method

As shown in Figure 4, I_{pcc}^h (ignoring the user impedance, that is I_c^h) fluctuates greatly during the T1 period, and the measured value in the T1 period can be used to calculate the system impedance; in the T2 period, I_{pcc}^h is stable (I_c^h is basically no fluctuation), and V_{pcc}^h fluctuates greatly, it can be seen that the system harmonic voltage (system harmonic current I_s^h) fluctuates greatly in this section, and the data in T2 section can be selected to divide the harmonic responsibility.

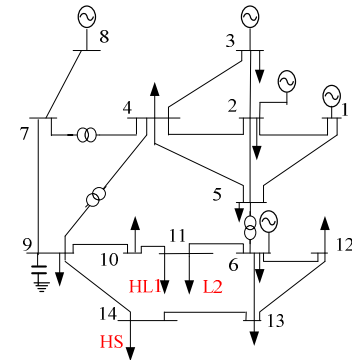


Fig. 5. IEEE 14-bus power system

V. SIMULATION EXPERIMENT

In Figure 5, the IEEE standard test system with 14-bus was used as the calculation and simulation example. Node 11 was selected as the concerned bus (PCC point), HL1 was the nonlinear harmonic source load concerned by the PCC point, L2 was the linear load, and HS was the system harmonic source load. The other parts except loads HL1 and L2 were taken as the system side. The results showed the harmonic voltage

contribution of the harmonic source load HL1 at bus No. 11 through simulation. Take the 5th harmonic as an example.

A. System harmonic impedance calculation

Firstly, the actual value $Z_s^5 = 2.87 \angle 79.75^\circ \Omega$ of the PCC point 5th harmonic system impedance was obtained by Digsilent software, and the userside impedance was ignored in this part.

Given the amplitude of the 5th harmonic current was from HS to 114.8A, the phase angle to -76.56° , and the user side impedance was neglected. First, specify N samples (I_c^5 values), perform harmonic power flow calculation for each sample value, and collect N harmonic voltage, current and their phase angle values at the PCC point. N samples of ΔV_{pccx}^5 , ΔV_{pccy}^5 , ΔI_{pccx}^5 , ΔI_{pccy}^5 can be calculated with the definition mentioned above. Use the Equation (9) to get the estimated value of the system harmonic phase angle $\beta_s^5 = 77.5^\circ$, and the absolute value of the error with the actual value β_{ac}^5 is only 2.25° , which can be ignored. After β_{es}^5 was got, γ can be calculated: $\gamma_i^5 = \alpha_i^5 + \beta_{es}^5$ ($1 \leq i \leq N$). After regression, $|V_{pcc}^5|$ to $|I_{pcc}^5 \cos \gamma^5|$, the system impedance modulus $|Z_s^5|$ is obtained.

Figure 6. and Figure 7. were the simulation results obtained by using method 1 and method 2 respectively. Method 1 and method 2 referred to the traditional correlation analysis and the method proposed in this paper respectively. It is known from the figure that the system impedance obtained by method one is 2.54Ω , and the system impedance obtained by method two is 2.75Ω . Compared with the actual value of 2.87Ω , the error of this method is 4.18%, which is obviously lower than the error of 11.49% of the traditional method.

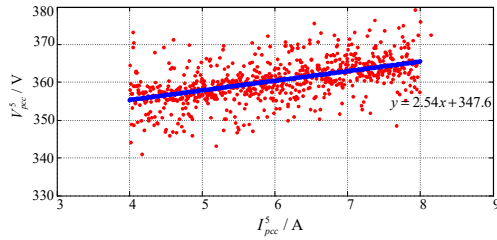


Fig. 6. Correlation between $|V_{pcc}^5|$ and $|I_{pcc}^5|$ of PCC

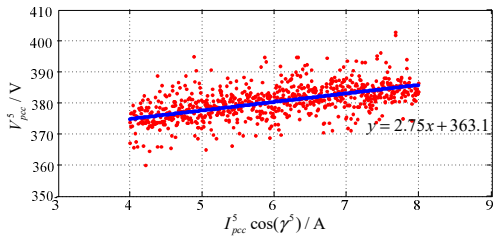


Fig. 7. Correlation between $|V_{pcc}^5|$ and $|I_{pcc}^5 \cos \gamma^5|$ of PCC

Table 1 showed the system impedance values $|Z_s^5|$ obtained by methods 1 and 2 when the user impedance was different

under the 5th harmonic. It can be seen from Table 1 that the accuracy of method 2 was obviously higher than that of method 1, and at the same time, the sensitivity of the system impedance value $|Z_s^5|$ obtained by method 2 to the change of user impedance $|Z_s^5|$ was lower than that of method 1.

TABLE I. SYSTEM 5TH HARMONIC IMPEDANCE $|Z_s^5|$ FOR DIFFERENT VALUES OF CONSUMER HARMONIC IMPEDANCE

	$ Z_s^5 = 78.2$	$ Z_s^5 = 108.7$	average error (%)
actual value	2.87	2.87	—
method 1	2.48	3.14	11.5
method 2	3.15	2.99	6.97

B. Quantitative division of harmonic responsibility

The classic curve of 5th harmonic current of system harmonic source could be adopted from Ref. [24]. If one minute corresponds to one sample point, then there are 1440 points in one day, as shown in Figure 8.

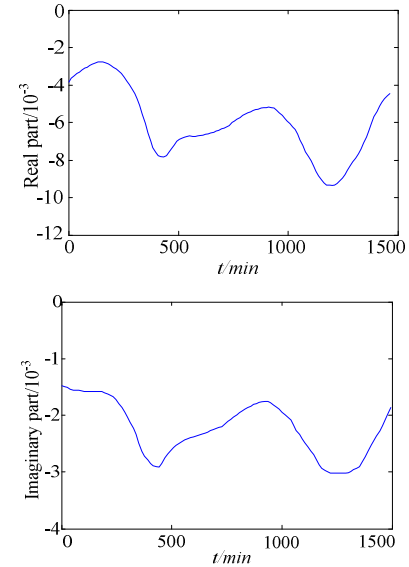


Fig. 8. 5th harmonic current injection by harmonic source HS

Assume that the user harmonic current is a fixed value, the magnitude is 10A, and the phase angle is -74.25° . The system impedance is $2.87 \angle 79.75^\circ \Omega$, and the user impedance is $78.2 \angle 90^\circ \Omega$. The simulation results obtained were shown in Figure 9, which was obviously seen that the user harmonic voltage was 29.2V at the PCC point. Taking the 5th harmonic voltage at the PCC point as 50V as the focus point, the user harmonic voltage contribution is 58.4%, which is very close to the actual value of 55.4%. The linear regression method ignoring the user impedance (method 1), the traditional correlation analysis (method 2), and the method in this paper (method 3) was used to calculate the harmonic liability to prove the accuracy of the method in this paper.

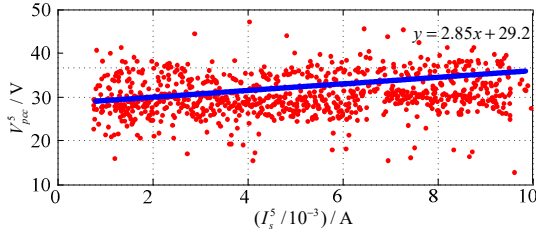


Fig. 9. Correlation between V_{pcc}^5 of PCC and I_s^5

Table 2. shows the user harmonic voltage contribution rates obtained by three methods when the user impedance is different. It can be seen from each column in the table that the accuracy of the method in this paper is obviously higher than that of methods 1 and 2, and it is more obvious as the user harmonic impedance decreases. Because methods 1 and 2 ignore user impedance in the calculation process ($Z_c^5 = \infty$) in the Norton equivalent circuit, the smaller Z_c^5 is, the greater the error it produces. Since the method in this paper considers the user impedance, even when Z_c^5 is small, the result is within the error range.

TABLE II. LOAD 5TH HARMONIC VOLTAGE CONTRIBUTION AT THE PCC (% $V_{c,pcc}^5$) FOR DIFFERENT VALUES OF CONSUMER HARMONIC IMPEDANCE

	% $V_{c,pcc}^5$			average error (%)
	$Z_c^5 = 43.3 \angle 90^\circ$	$Z_c^5 = 78.2 \angle 90^\circ$	$Z_c^5 = 108.7 \angle 90^\circ$	
actual value	53.7	55.3	55.8	—
method 1	27.1	49.3	51.6	22.64
method 2	39.6	41.5	44.3	23.94
method 3	45.9	58.4	53.7	7.96

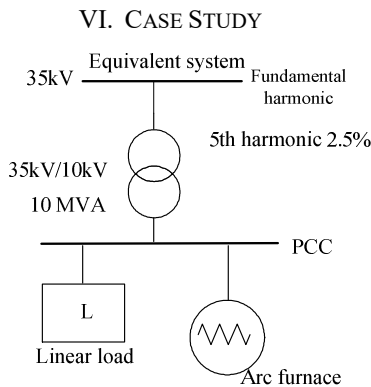


Fig. 10. Equivalent circuit of the 35/10 kV industrial power fundamental harmonic systems

Figure 10 was the circuit equivalent diagram of 35kV/10kV industrial system, the short-circuit capacity is 10MVA, the 35kV bus contains 2.5% of the 5th harmonic component, and the linear load and nonlinear load electric arc furnace was accessed

through 10kV distribution system. The sampling frequency was 3200 Hz, and fast Fourier transform was performed on the sampling data every minute to obtain the measured values of each harmonic. The measured 5th harmonic voltage V_{pcc}^5 and current I_{pcc}^5 at the PCC point within 300 minutes on a certain day were shown in Figure 11. The load runs normally for the first 240 minutes, and the load stops for the next 60 minutes.

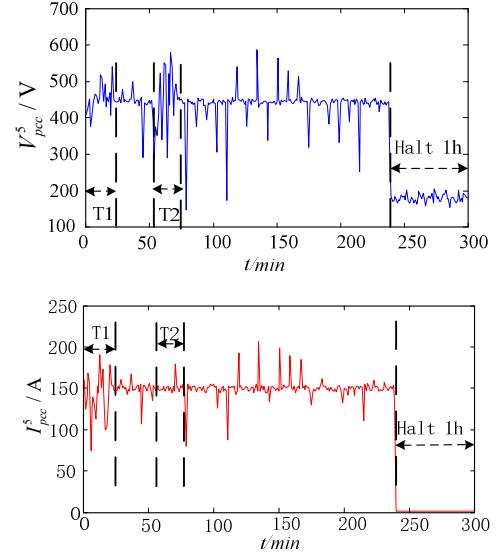


Fig. 11. Magnitude waveforms of the 5th harmonic voltage and current at the PCC

It can be seen from the above data selection scheme that in the T1 segment in Figure 11, the harmonic current fluctuates greatly during the start-up process of the user, so the data in the T1 segment can be used to calculate the system impedance. The 5th harmonic impedance angle of the system was 67.89° by the method in this paper, the numerical value was 2.16Ω .

Select the data in section T2 to calculate the user's harmonic voltage contribution. Table 3 shows the contribution of 5th harmonic voltage for the system and the user at PCC point under the 95% probability value of V_{pcc}^5 was 446.5V, respectively using the traditional correlation analysis and the method in this paper.

TABLE III. COMPARISON OF THE ESTIMATED 95% PROBABILITY VALUES OF $V_{c,pcc}^5$ AND $V_{s,pcc}^5$

Responsibility distinguishing method	$V_{c,pcc}^5$	$V_{s,pcc}^5$
Traditional correlation analysis method	247.1	204.6
Method in this article	267.9	178.6

The 5th harmonic voltage when the load is out of service in the measured data is taken as the contribution of the system harmonic voltage, as shown in Figure 12, and 95% probability value was 175.2 V. Comparing the data in Table 3, it can be seen

that the calculated value of the method in this paper was the most accurate, which proved that the harmonic voltage generated by the user at the PCC point was 267.9V is credible.

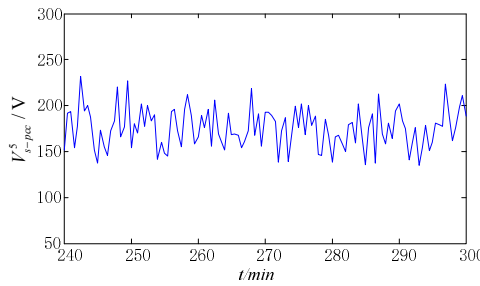


Fig. 12. $V_{s,pcc}^5$ measured at the PCC

VII. CONCLUSION

Accurate calculation of system harmonic impedance is the premise of realizing the quantitative division of harmonic responsibilities. In view of the lack of ignoring the harmonic phase angle when the traditional correlation analysis method is used to solve the system impedance, this paper proposes to take the harmonic phase angle into the equation to obtain a more accurate system harmonic impedance. In addition, in view of the shortcomings of neglecting user impedance and assuming that the system harmonic voltage was the fixed value when the traditional linear regression method divides the harmonic responsibility (voltage contribution), this paper proposed the fluctuation of the system harmonic voltage and combine the correlation analysis method to obtain the quantitative harmonic of the user at the PCC point. Some classical Evolutionary Algorithms were also reviewed with existing successful cases for implementation in power system area. The simulation analysis and actual engineering examples have verified that the method in this paper can obtain more accurate system impedance than the existing methods, and the division of harmonic responsibilities has obtained relatively satisfactory results. Furthermore, the adoption of gene expression coding can further eliminate the interference of irrelevant factors, make the calculation results converge faster, and even be applied to the analysis of correlation relationships among multiple databases and influencing factors.

ACKNOWLEDGMENT

This research was funded by State Grid Sichuan Electric Power Company of State Grid Corporation of China, grant number 511904230005.

REFERENCES

- [1] Jia Xiufang, Hua Huichun, Cao Dongsheng, et al. Determining harmonic contributions based on complex least squares method[J]. Proceedings of the CSEE, 2013, 33(4): 149-155.
- [2] Zhao J, Yang H, Ma X, et al. Evaluation of Harmonic Contributions for Multi Harmonic Sources System Based on Mixed Entropy Screening and an Improved Independent Component Analysis Method[J]. Entropy, 2020, 22(3): 323-340.
- [3] Shu Q, Zhao S, Xu F. A Novel Estimation Method of Utility Harmonic Impedance Based on Short-term Impedance Minimum Variance Criterion[J]. IET GENER TRANSM DIS, 2020, 14(15): 13-23.
- [4] Tian Shuya, Jia Qingquan, Lin Lijuan, et al. Collaborative Allocation of Voltage Detection Active Power Filter and Static Var Generator for

- Harmonic and Voltage Comprehensive Mitigation[J]. Automation of Electric Power Systems, 2021, 45(24): 149-157.
- [5] Zhao J, Yang H, Pan A, et al. An improved complex ICA based method for wind farm harmonic emission levels evaluation[J]. Electric power systems research, 2020, 179(2): 1-14.
- [6] Naude H, Beukes J, Minnaar U. The Impact of Traction Load Harmonic Current Emissions on the Harmonic Assessment of Renewable Power Plants[C]// IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, IEEE, 2019, 4764-4769.
- [7] Hui Jin, Yang Honggeng, Ye Maoqing. Assessing harmonic emission level based on the impedance gathering trend discrimination[J]. Proceedings of the CSEE, 2011, 31(10): 73-80.
- [8] Shu Q, He J, Wang C. A Method for Estimating the Utility Harmonic Impedance Based on Semiparametric Estimation[J]. IEEE T INSTRUM MEAS, 2021, 70: 1-9.
- [9] Zang T, He Z, Ling F, et al. Adaptive method for harmonic contribution assessment based on hierarchical K-means clustering and Bayesian partial least squares regression[J]. IET GENER TRANSM DIS, 2016, 10(13): 3220-3227.
- [10] Mazin H E, Xu W, Biao H. Determining the harmonic impacts of multiple harmonic-producing loads[J]. IEEE T POWER DELIVER, 2011, 26(2): 1187-1195.
- [11] Hua Huichun, Jia Xiufang, Zhang Shaoguang. Neighborhood Multi-Point Measurement Method for Harmonic Contribution Determination [J]. Power System Technology, 2014,38(2): 502-508.
- [12] Bulychева E, Yanchenko S. Real-time harmonic identification under varying grid conditions[J]. Serbian Journal of Electrical Engineering, 2021, 18(1): 29-48.
- [13] Eslami A, Negnevitsky M, Franklin E , et al. Review of AI applications in harmonic analysis in power systems[J]. RENEW SUST ENERG REV, 2022, 154(2): 1-26.
- [14] Mukhtaj Khan, Zhengwen Huang, Maozhen, Li, et al. Optimizing hadoop parameter settings with gene expression programming guided PSO[J]. Concurrency and Computation: Practice and Experience, 2017, 29(3): 1-21.
- [15] Naik B, Kaushik V, Kumar M. Water surface profile in converging compound channel using gene expression programming[J]. Water science & technology: Water supply, 2022, 22(5): 5221-5236.
- [16] Xu F, Wang W, Zheng H, et al. Harmonic impedance estimation considering the correlation between harmonic sources[J]. ELECTR POW SYST RES, 2022,209(8): 1-10.
- [17] Bäck, T. Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms [M]. Oxford, UK: Oxford Univ. Press, 1996: 235-320.
- [18] Koza, J.R. Genetic Programming: On the Programming of Computers by Means of Natural Selection [M]. Massachusetts: MIT Press, 1992: 185.
- [19] C. Ferreira, Gene Expression Programming: A New Adaptive Algorithm for Solving Problems [J]. Complex Systems, 2001, 13(2): 87-129.
- [20] Khan, M., Huang, Z., Li, M., Taylor, G. A., Ashton, P. M., & Khan, M. Optimizing hadoop performance for big data analytics in smart grid [J]. Mathematical Problems in Engineering, 2017, 1-11.
- [21] de Jong, E. D., Watson, R. A., and Pollack, J. B. Reducing bloat and promoting diversity using multi-objective methods [C]. San Francisco, CA, USA: Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation, 2001, 11-18.
- [22] Janeiro, Fernando M., José Santos, and Pedro M. Ramos. Gene expression programming in sensor characterization: Numerical results and experimental validation [J]. IEEE Transactions on Instrumentation and Measurement, 2012, 62(5): 1373-1381.
- [23] Wan, Jianfen, Yin Hongrun, Zhu Chengjun, et al. A Hybrid Genetic Expression Programming and Genetic Algorithm (GEP-GA) of Auto-modeling Electrical Equivalent Circuit for Particle Structure Measurement with Electrochemical Impedance Spectroscopy (EIS) [J]. IEEE Sensors Journal, 2023, 23(5): 4344-4351.
- [24] Gursoy E, Niebur D. Harmonic load identification using complex independent component analysis[J]. IEEE T POWER DELIVER, 2009, 24(1): 285-292.