Principal Component Regression and Its Application in Power Harmonic Emission Level Evaluation

Shuping Song¹, *, Jie Wang², a, Linheng Li³, b, Shuo Cheng¹, c and Zhijian Li¹, d

¹College of Electrical and Automation Engineering, Changshu Institute of Technology, Changshu 215500, China
²Ningneng Electric Design Limited Company, Nanjing 210000, China
³College of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255022, China

*Corresponding author e-mail: ssp17@126.com, awangjie_1980@yeah.net,
b825032418@qq.com, c18852933321@163.com, d1226004964@qq.com

Abstract. Accurate harmonic emission level evaluation is the basis of distinguishing the responsibility of harmonic pollution between power supply system and consumer. A new harmonic emission level evaluation method via principal component regression (PCR) is proposed in this paper. Firstly, the principle of PCR is analyzed. Then, harmonic impedance of supply system is evaluated by PCR-based method. Subsequently, harmonic emission level of supply system and consumer is evaluated according to the harmonic impedance of supply system. This technique has the advantage of high accuracy in estimating harmonic emission level. The effectiveness of the proposed method was verified by computer simulations and field test.

1. Introduction
The responsibility of harmonic pollution attribution between supply system and consumer must be clear to control harmonics [1]. Therefore, evaluation of harmonic emission level has become an important task [2].

At present, the research of power harmonic emission level evaluation is mainly focused on the contribution of system side and consumer to the point of common coupling (PCC) [3, 4]. In fact, harmonic emission level evaluation is performed by applying statistical methods, due to its statistical characteristics [5]. Early harmonic emission level evaluation method via statistics is the fluctuation method [6]. However, it is difficult for this method to obtain accurate results when the background harmonic fluctuation is small. And then, Binary linear regression [7], robust regression [8], partial least squares regression [9], partial ridge regression method [10] and so on are applied for harmonic emission level evaluation currently. The methods above improve the accuracy of power harmonic emission level evaluation to a certain extent. However, they encounter the problem of multicollinearity of independent variables. In addition, in some cases, these methods cannot extract the original variable information with strong explanatory ability to dependent variables, leading to unreasonable results of least squares regression. These problems affect the application of the existing methods in practice.

Accordingly, this paper proposes a PCR-based method to evaluate power harmonic emission level. The principal components are selected according to the amount of the information contained in this
method, which effectively overcomes the no orthogonality among the independent variables in modeling. Besides, the problem that the important independent variable information may be omitted is also avoided. Thus, the accuracy of harmonic emission level evaluation for supply system and consumer is effectively improved.

2. PCR Analysis

It is assumed that the problem studied has \( p \) independent variables, which are expressed as \{\( x_1, x_2, \ldots, x_p \}\). Then, the data of \( n \) samples can be recorded as follows:

\[
X = \begin{pmatrix}
  x_{11} & x_{12} & \cdots & x_{1p} \\
  x_{21} & x_{22} & \cdots & x_{2p} \\
  \vdots & \vdots & \ddots & \vdots \\
  x_{n1} & x_{n2} & \cdots & x_{np}
\end{pmatrix} = (x_1, x_2, \ldots, x_p)
\]  

(1)

The idea of PCR analysis is to transform this \( p \) independent variable into a linear combination of \( p \) independent variables [11]. The linear combination of the above \( p \) independent variables can be expressed by the following formulas:

\[
\begin{align*}
F_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\
F_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\
&\quad \ldots \\
F_p &= a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p
\end{align*}
\]  

(2)

Where \( a_{ij} \) is the principal component coefficient.

If \( F_1 \) is used to replace the original \( p \) variables, \( F_1 \) need to involve the information of the \( p \) variables as much as possible. Thus, the larger the variance \( \text{Var}(F_1) \), the better. Here, \( F_1 \) is called the first principal component. If \( F_1 \) cannot involve all information of variables, the second principal component \( F_2 \) should be introduced. \( F_1 \) is not correlated with \( F_2 \), that is, \( \text{Cov}(F_1, F_2) = 0 \). And so on, if there are \( p \) principal components, these principal components are not correlated with each other. At the same time, \( \text{Var}(F_1) > \text{Var}(F_2) > \cdots > \text{Var}(F_p) \). Usually, the first \( q \) principal components are selected in practice according to the cumulative contribution rate of principal component. The formula of contribution rate \( \mu \) is:

\[
\mu = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}
\]  

(3)

Where \( \lambda_i = \text{Var}(F_i) \).

Generally speaking, only when the cumulative contribution rate is more than 85\% [12], can we ensure that the integrated variables contain the vast majority of information of the original variables. Selecting principal components according to the contribution rate can avoid important information being deleted by mistake. In addition, when the value of \( F_i \) is close to zero, the corresponding independent variable is an approximate linear dependence relationship. Therefore, the measure above is able to neglect the smaller principal components and can solve the collinearity problem among independent variables.
After calculating the principal component expression and the sub-variables by using PCR, dependent variables are regressed to sub-variables of principal component. Then, the regression model among dependent variables and sub-variables of principal component is obtained.

3. Evaluation of Power Harmonic Emission Level based on PCR

Supply system and consumer are equivalent to the following circuit at PCC, shown in Figure 1. In Figure 1, $U_{sh}$ is the harmonic voltage source of system side; $I_{ch}$ is the equivalent harmonic current source of consumer side; $Z_{sh}$ is the harmonic impedance of system side; $Z_{ch}$ is the harmonic impedance of consumer; $I_{pcch}$ is the harmonic current at PCC; $U_{pcch}$ is the harmonic voltage at PCC.

![Figure 1. The equivalent circuit of system and consumer at PCC](image)

According to the circuit principle, the following formula can be achieved:

$$\dot{U}_{pcch} = \dot{U}_{sh} - \dot{I}_{pcch} Z_{sh}$$  (4)

Expanding (4) according to the real part and imaginary part, then:

$$U_{pcchx} = U_{shx} - I_{pcchx} Z_{shx} + I_{pcchy} Z_{shy}$$  (5)

$$U_{pcchy} = U_{shy} - I_{pcchx} Z_{shy} - I_{pcchy} Z_{shx}$$  (6)

Equation (5) and (6) are the multiple linear regression equations with the real part $I_{pcchx}$ and imaginary part $I_{pcchy}$ of $I_{pcch}$ as the independent variables and the real part $U_{pcchx}$ and imaginary part $U_{pcchy}$ of $U_{pcch}$ as the dependent variables. The two equations can be solved by using PCR to obtain regression coefficients $Z_{shx}$, $Z_{shy}$, $U_{shx}$ and $U_{shy}$. In practice, the data for analysis can be divided into several sections to reduce the influence of background harmonic fluctuation, and then the average values of $Z_{sh}$ and $U_{sh}$ are obtained.

Subsequently, the harmonic emission value of consumer can be obtained according to $Z_{sh}$ and $U_{sh}$. In the light of the superposition principle, Figure 1 can be divided into two circuits, which is shown in Figure 2.
From Figure 2,

$$U_{spcch} = \frac{Z_{ch}}{Z_{ch} + Z_{sh}} U_{sh} \quad (7)$$

$$U_{cpcch} = \frac{Z_{ch}Z_{sh}}{Z_{ch} + Z_{sh}} I_{sh} \quad (8)$$

$$U_{pcch} = U_{spcch} + U_{cpcch} \quad (9)$$

Consumer is usually regarded as harmonic current sources, whose resistance $Z_{ch}$ is large. Meanwhile, system side is equivalent to harmonic voltage source and its resistance $Z_{sh}$ is small. Thus, $|Z_{ch}| >> |Z_{sh}|$. So the harmonic emission level of consumer $U_{ch}$ can be expressed as:

$$U_{ch} = U_{pcch} - \frac{Z_{ch}U_{sh}}{Z_{ch} + Z_{sh}} \approx U_{pcch} - U_{sh} \quad (10)$$

4. Simulation
Matlab/Simulink is used for simulation. Simulation circuit is shown in Figure 3. In simulation, the impedance set value of fundamental and harmonic of supply system are shown in Table 1. The fundamental frequency of system is 49.9Hz, the fundamental initial phase is 10°, and the voltage amplitudes of harmonic are shown in Table 2. Besides, the current amplitudes of consumer harmonic current source are set as follows: $I_{c3}$ is 47A, $I_{c5}$ is 16A, $I_{c7}$ is 8A, $I_{c9}$ is 6A, and $I_{c11}$ is 3A. The fundamental impedance of consumer is set to 2.723+j11.345Ω.
Table 1. Settings and Estimates of Harmonic Impedance of System Side

| Harmonic order | Set value of $Z_{sh}$ (Ω) | Evaluated value of PLSM (Ω) | Evaluated value of PRRM (Ω) | Evaluated value of the proposed method (Ω) |
|----------------|---------------------------|----------------------------|----------------------------|---------------------------------|
| Fundamental    | 0.5+j1.371                | 0.471+j1.362               | 0.481+j1.389               | 0.5+j1.371                     |
| 3              | 0.5+j4.146                | 0.486+j4.012               | 0.480+j4.291               | 0.497+j4.181                   |
| 5              | 0.5+j6.910                | 0.516+j6.718               | 0.530+j6.503               | 0.515+j6.610                   |
| 7              | 0.5+j9.674                | 0.575+j9.970               | 0.547+j9.372               | 0.519+j9.728                   |
| 9              | 0.5+j12.438               | 0.468+j12.901              | 0.4772+j12.107             | 0.487+j12.173                  |
| 11             | 0.5+j15.202               | 0.521+j15.937              | 0.542+j16.152              | 0.531+j15.795                  |

Partial least squares method (PLSM), partial ridge regression method (PRRM), and the proposed method are used for simulation to estimate the harmonic impedance and the harmonic emission level of the system side respectively. The estimation results are given in Tables 1 and 2. In addition, the harmonic emission level of system is listed in the form of percentage relative error.

From Table 1 and 2, we can see that the proposed method has higher accuracy in evaluation of harmonic impedance and harmonic emission level, compared with the existing methods. For the evaluation of harmonic Impedance of system, the partial ridge regression method has the greatest error.

Figure 3. Simulink simulation circuit
among these three methods. The evaluation error of partial least squares method is small, but certain estimated results have large error, such as the real part of the 7th harmonic impedance. The estimated values of the proposed method are close to the set values, except for the 5th harmonic impedance estimates. On the estimation of harmonic emission level of system, the maximum relative error of the proposed method is only 1.75%, which is lower than the two existing methods. Therefore, the simulations indicate that the proposed method can accurately estimate the harmonic emission level.

5. Field test

Test was carried out in a chemical plant. The plant is powered by 10 kV buses of 35 kV substation. The minimal short circuit capacity of 10 kV is 175.3 MVA under the minimal operation mode, and the corresponding capacity is 241.6 MVA under the maximal operation mode. The capacity of power supply equipment at PCC is 2 × 20 MVA.

The test arrangement is as follows: Firstly, the 10 kV bus voltage of supply system and the 10 kV outgoing current of the plant are analysed with the measuring device. During the measurement, sampling period is 5 seconds, and harmonics greater than 20 times is filtered out. Sampling frequency is 5000Hz. Secondly, the interpolation FFT algorithm based on 5-term maximum side lobe attenuation window [13] is used for data analysis every 500 data points. Then, the harmonic voltage and current data at PCC are analysed by PCR-based method. In the process of analysis, every 500 voltage and current data is applied for harmonic emission level evaluation one time.

Due to space limitation, this paper only lists the evaluated harmonic impedance of the phase A 5th harmonic of system. The average value of 5th harmonic impedance at system side is 0.167+j2.471Ω. The real and imaginary parts of the impedance are shown in Figure 4. The abscissa \( n \) in Figure 5 represents the \( n \)th calculation.

![Graph](image.png)

(a) Real part of impedance  
(b) Imaginary part of Impedance

Figure 4. Evaluated value of the 5th harmonic impedance at system side

According to the maximal and minimal operation mode capacity of 10 kV bus above, the fundamental reference value of reactance of system is between 0.418Ω and 0.576Ω. For low-order harmonics of power system, the reactance on the same line satisfies \( L_1 = L_h/h \) [14]. Among them, \( L_h \) is the \( h \)th harmonic reactance and \( L_1 \) is the fundamental reactance. The 5th harmonic reactance \( L_{s5} \) obtained by the proposed method is 2.471Ω. According to \( L_1 = L_h/h \), \( L_1 \) is equal to 0.494Ω. The evaluated value is close to the average value in the range of [0.418Ω, 0.576Ω]. Therefore, the 5th harmonic reactance of system obtained by the proposed method meets the engineering practice.

6. Conclusion

A PCR-based method is proposed in this paper to evaluate the harmonic emission level of supply system side and consumer. This technique applies the complex relation of harmonic voltage, harmonic current and system side harmonic impedance at PCC to get regression equation, and then uses PCR analysis to
evaluate system side harmonic impedance. Finally, the harmonic emission levels of system and consumer are evaluated according to the principle of circuit superposition. This method removes the principal component with small variance in regression modeling and solves the problem of multiple correlations among independent variables. In addition, the presented method can extract principal components with strong explanatory ability to dependent variables in turn, avoiding the deletion of important original independent variable information by mistake. The simulation and test results show that the proposed method can estimate the harmonic emission level of power system with high accuracy.

Acknowledgments
This work was financially supported by the Science Fund of Changshu Institute of Technology KYZ2017099Z.

References
[1] S. Elphick, V. Gosbell, V. Smith, Methods for harmonic analysis and reporting in future grid applications, IEEE Trans. Power Del. 32 (2017) 989 - 995.
[2] W. Xu, Y. L. Liu, A method for determining customer and utility harmonic contribution at the Point of common coupling, IEEE Trans. Power Del. 15 (2000) 804 - 811.
[3] J. Hui, H. G. Yang, S. F. Lin, et al, Assessing utility harmonic impedance based on the covariance characteristic of random vectors, IEEE Trans. Power Del. 25 (2010) 1778 - 1786.
[4] S. Bhattacharyya, S. Cobben, P. Ribeiro, et al, Harmonic emission limits and responsibilities at a point of connection, IET Gener Transm Dis. 6 (2012) 256 - 264.
[5] J. Wang, Research on harmonic detection and emission level evaluation method of power grid, Southeast University, Nanjing, 2016, pp.1 - 7.
[6] H. G. Yang, P. Pirotte, Harmonics emission levels of industrial loads-statistical assessment, CIGRE. (1996) 15 - 24.
[7] W. Zhang, H. G. Yang. A method for assessing harmonic emission level based on binary linear regression, Proceedings of the CSEE. 24 (2004): 50 - 53.
[8] Q. Che, H. G. Yang, Assessing the harmonic emission level based on robust regression method, Proceedings of the CSEE. 24 (2004) 39 - 42. (In Chinese).
[9] S. Huang, Y. H. Xu, Assessing harmonic impedance and the harmonic emission level based on partial least-squares regression method, Proceedings of the CSEE. 27 (2007) 93 - 97.
[10] Y. L. Cui, H. Z. Ma, Y. M. Tang, Assessing harmonic impedance and the harmonic emission level based on partial ridge, Electric Power. 48 (2015) 137 - 140.
[11] T. Jolliffe, Principal component analysis, Springer series in Statistics, New York, 2002, pp. 1 - 21.
[12] X. Q. He, Applied multivariate statistical analysis, China Statistics Press, Beijing, 2010, pp. 56 - 87.
[13] S. P. Song, H. Z. Ma, G. Xu, Power harmonic analysis based on 5-term Maximum-sidelobe-decay window interpolation, Autom Electr Power Syst. 39 (2015) 83 - 88.
[14] J. C. Wu, Power supply system harmonics, China Electric Power Press, Beijing, 1988.