Research Article

Data-Driven Fault Diagnosis Method for Power Transformers Using Modified Kriging Model

Yu Ding and Qiang Liu

School of Information and Control Engineering, Liaoning Shihua University, Fushun, Liaoning 113001, China

Correspondence should be addressed to Qiang Liu; neuliuqiang@163.com

Received 25 May 2017; Accepted 20 September 2017; Published 22 October 2017

Academic Editor: Wanquan Liu

Copyright © 2017 Yu Ding and Qiang Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A data-driven fault diagnosis method that combines Kriging model and neural network is presented and is further used for power transformers based on analysis of dissolved gases in oil. In order to improve modeling accuracy of Kriging model, a modified model that replaces the global model of Kriging model with BP neural network is presented and is further extended using linearity weighted aggregation method. The presented method integrates characteristics of the global approximation of the neural network technology and the localized departure of the Kriging model, which improves modeling accuracy. Finally, the validity of this method is demonstrated by several numerical computations of transformer fault diagnosis problems.

1. Introduction

Transformer is one of the most important equipment in power system [1], which is mainly used to transfer electrical energy between two or more circuits through electromagnetic induction. In the course of using this equipment, some factors such as electrical, thermal, and mechanical stresses may lead to irreversible damage to the insulating material [2]. In order to improve the reliability of power supply, fault diagnosis for transformer has drawn much attention from researchers, and many fault diagnosis methods have been widely proposed during the past decades.

At present, Dissolved Gas-in-oil Analysis (DGA) is a commonly used method to identify incipient failures of transformer fault [3]. With the development of artificial intelligence and computer technology, many fault diagnosis algorithms have been proposed based on DGA, such as neural network [4, 5], fuzzy logic [6, 7], expert system [8], support vector machine [2, 9], and rough set theory [10]. Given that existing methods have their own characteristics and some limitations, effective fault diagnosis methods that integrate advantages of existing technologies to improve the modeling accuracy still remain an open area of research.

As a classic modeling technology, Kriging model combines a global model plus localized departures to construct

approximation from sample data. It has been widely used in the field of Computer-Aided Engineering (CAE) [11, 12]. On the other hand, neural network technology is a well-known information processing paradigm and has been widely applied in various areas due to its advantages such as adaptive learning. In this paper, a data-driven fault diagnosis model based on Kriging model and BP neural network is constructed and then is used for transformer fault diagnosis problems based on DGA. In order to improve the modeling accuracy of Kriging model, BP neural network is used to modify the global model of Kriging model, where the modified formula is given in detail. The modified Kriging model combines global and adaptive learning abilities of neural network technology and retains the localized departures of the Kriging model in the meantime. It integrates the advantages of both methods, which thus effectively improve the modeling accuracy. Finally, some examples of transformer fault diagnosis show that the proposed method is effective and feasible.

2. Overall Design of Modified Kriging Model: A Hybrid Model

The Kriging model is an unbiased estimation model based on the minimum variance estimation of sample points and their

FIGURE 1: Overall design of hybrid model.

response values. The output can be viewed as a combination of a regression model and a stochastic process [13, 14]. The regression model is equivalent to the global simulation of the sample space, and the stochastic process is equivalent to local deviation.

In order to further improve the accuracy of the Kriging model, this paper proposes a transformer fault diagnosis method that combines Kriging model and neural network technology. The structure of this hybrid model is shown in Figure 1. The main steps for constructing the hybrid model are as follows:

- (1) Determine characteristic variables and fault types based on DGA method, and collect sample data and test data.
- (2) According to sample data, construct BP network model.
- (3) Construct the Kriging model based on sample data.
- (4) The global simulation of Kriging model is modified and updated by neural network, and then the hybrid model is constructed.

3. Modeling Method of Hybrid Model

3.1. Kriging Model and Parameter Optimization. The Kriging model contains global simulation plus localized departures, and the basic principles of which can be briefly given as follows [13, 14]: set approximate function as $y(x)$, and the function between the response value and the independent variable of the system can be formulated as follows:

$$y(x) = F(\beta, x) + Z(x), \quad (1)$$

where $y(x)$ is the unknown function of interest, $F(\beta, x)$ is the regression model that is equivalent to the global simulation, and β is regression parameter; $Z(x)$ is a normal stochastic

process in which the mean value is 0 and the variance is denoted as σ^2 . It reflects the randomness of the response and is equivalent to partial divergence.

The covariance matrix of $Z(x)$ is formulated as follows:

$$\text{cov}[Z(x_i), Z(x_j)] = \delta^2 \mathbf{R}[R(\theta, x_i, x_j)], \quad (2)$$

where $\mathbf{R}$ is correlation matrix; the order of matrix is M . $i, j = 1, 2, \dots, M$; M is the number of sample points; x_i and x_j are the i th and the j th sample points; $R(\theta, x_i, x_j)$ is the correlation function between x_i and x_j . In this paper, we utilize the Gaussian correlation function:

$$R(\theta, x_i, x_j) = \exp\left(-\sum_{k=1}^N \theta_k |x_i^{(k)} - x_j^{(k)}|^2\right), \quad (3)$$

where N is the dimension of the problem; $x_i^{(k)}$ and $x_j^{(k)}$ are the k th dimensional components of the i th and the j th sample points, respectively; θ_k is the unknown related parameters of the interpolation model.

In general, θ_k can be replaced with a scalar θ . Thus, formula (3) can be formulated as

$$R(\theta, x_i, x_j) = \exp\left(-\theta \sum_{k=1}^N |x_i^{(k)} - x_j^{(k)}|^2\right). \quad (4)$$

Therefore, the estimated value of test point x can be given by the following equation:

$$\hat{y}(x) = \hat{\beta} + \mathbf{r}^T(x) \mathbf{R}^{-1}(\mathbf{y} - \mathbf{f}\hat{\beta}), \quad (5)$$

where $\hat{\beta}$ is the estimate of the global simulation; $\mathbf{y}$ is sample data response; $\mathbf{f}$ is M column vectors; $\mathbf{r}^T(x)$ is the correlation vector between observation point x and sample data, which can be formulated as follows:

$$\mathbf{r}^T(x) = (R(x, x_1), R(x, x_2), \dots, R(x, x_M))^T. \quad (6)$$

When $f(x)$ is a constant, $\hat{\beta}$ can be simplified and estimated by the following equation:

$$\hat{\beta} = (\mathbf{f}^T \mathbf{R}^{-1} \mathbf{f})^{-1} (\mathbf{f}^T \mathbf{R}^{-1} \mathbf{y}). \quad (7)$$

The parameter θ determines the accuracy of the Kriging model, which can be solved by the following optimization problem:

$$\min L(\theta) = |\mathbf{R}|^{1/M} \hat{\delta}^2, \quad (8)$$

where $\hat{\delta}^2$ is variance estimation, which can be determined by the following equation:

$$\hat{\delta}^2 = \frac{1}{M} (\mathbf{y} - \mathbf{f}\hat{\beta})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{f}\hat{\beta}). \quad (9)$$

To optimize parameter θ , intelligent optimization algorithms are commonly used. In this paper, a modified particle swarm optimization (MPSO) [15] is used to optimize the parameters of Kriging model. The key points of applying

MSPO to perform optimization are threefold: (1) make θ as encoding in real numbers; (2) take (8) as the objective function; and (3) with respect to the constraint condition of $\theta > 0$, the commonly used penalty function method is applied.

The inertia weight and learning factors in MPSO [15] are updated as follows:

$$\omega = (\omega_s - \omega_e) \left(\frac{t}{T} \right) + (\omega_e - \omega_s) \left(\frac{2t}{T} \right) + \omega_s, \quad (10)$$

$$c1 = (c1_s - c1_e) \left(\frac{t}{T} \right) + (c1_e - c1_s) \left(\frac{2t}{T} \right) + c1_s, \quad (11)$$

$$c2 = (c2_s - c2_e) \left(\frac{t}{T} \right) + (c2_e - c2_s) \left(\frac{2t}{T} \right) + c2_s, \quad (12)$$

where ω_s and ω_e are the initial value and the final value of inertia weight, respectively; $c1_s$ and $c1_e$ and $c2_s$ and $c2_e$ are the initial value and the final value of learning factors $c1$ and $c2$, respectively; T is the maximum number of iterations; t is the current iteration number.

3.2. Combinations of Neural Network and Kriging Model.

The mapping relationship between the characteristic variables and fault types of transformer is very complex, which increases difficulty in improving high accuracy of Kriging model. On the other hand, neural network technology, for example, BP network, is a well-known information processing paradigm with some advantages such as adaptive learning and strong adaptability. The basic principle of BP network can be found in many references such as [4, 5, 16], details of which are not introduced here. In general, the output of BP network can be formulated as follows:

$$o_i = f \left(\sum_{j=0}^n \omega_{ij} x_j + q_i \right), \quad (13)$$

where o_i are the outputs of BP network, f is transfer function, ω_{ij} are network weights, q_i are network thresholds, and x_j are the outputs of the upper layer node.

To improve modeling accuracy, a modified Kriging model (hybrid model) is constructed by combining Kriging model and BP neural network technology, and the overall design of which has been shown in Figure 1. More specifically, the global model of Kriging model is replaced with BP neural network, which is given by (14):

$$\hat{y}(x) = o_i + \mathbf{r}^T(x) \mathbf{R}^{-1} (\mathbf{y} - \mathbf{f}\hat{\beta}). \quad (14)$$

Further, this modified method can be extended using linearity weighted aggregation method, which is formulated by (15):

$$\lambda = \gamma_1 o_i + \gamma_2 \hat{\beta}, \quad (15)$$

where λ is the modified global model and γ_1 and γ_2 are weighting coefficients.

Thus, the final output of the hybrid model can be given as follows:

$$\hat{y}(x) = \lambda + \mathbf{r}^T(x) \mathbf{R}^{-1} (\mathbf{y} - \mathbf{f}\hat{\beta}). \quad (16)$$

TABLE 1: Data distribution of each fault type.

Fault type	Coding mode	Sample data	Test data
Normal	Y_1	15	15
High temperature overheating	Y_2	25	25
Medium temperature overheating	Y_3	15	15
Low temperature overheating	Y_4	9	9
Partial discharge	Y_5	9	8
Low energy discharge	Y_6	20	16
High energy discharge	Y_7	20	23

TABLE 2: Comparisons of diagnostic results.

Test method	Average calculation time	Accuracy rate
BP network	5.06 s	72.07%
SVM	6.21 s	75.67%
Kriging model	2.49 s	81.08%
Hybrid method	7.84 s	91.89%

Obviously, the hybrid model is of generality. When $\gamma_1 = 0$ and $\gamma_2 = 1$, the hybrid model becomes original Kriging model and (16) can be rewritten by (5). When $\gamma_1 = 1$ and $\gamma_2 = 0$, the hybrid model can be formulated by (14), where the global model of Kriging model is replaced with BP neural network.

4. Application of Transformer Fault Diagnosis

4.1. Feature Variable and Fault Type. In general, the concentrations of five gases (H_2 , CH_4 , C_2H_6 , C_2H_4 , and C_2H_2) dissolved in transformer oil can be selected as characteristic variables based on DGA data samples. The corresponding fault types of the characteristic variables contain normal, high temperature overheating, medium temperature overheating, low temperature overheating, partial discharge, low energy discharge, and high energy discharge. In this paper, we select some DGA data published in [17, 18]. The distributions and coding of these fault data are shown in Table 1.

4.2. Parameter Setting. In this paper, BP network structure is set as three-layer network. The selected sample data have 5 characteristic variables; thus the number of nodes in the input layer of the network is set as 5; the number of nodes in the output layer is set as 1; the number of hidden layer nodes is set as 8 by trial and error. The initial weights and thresholds of the network are randomly initialized, and Log-sigmoid is selected as function transfer function.

Parameters of MPSO are set as follows: $\omega_s = 0.95$, $\omega_e = 0.55$, $c1_s = c2_e = 2.25$, $c1_e = c2_s = 0.75$, $T = 40$, and population size $P = 20$. Figure 2 shows the MPSO convergence curve of objective function.

4.3. Analysis of Examples. (1) Table 2 shows the comparisons between the proposed method and other methods. As far as these test examples are concerned, the proposed method

TABLE 3: Some test results ($\mu\text{L/L}$).

Serial number	H_2	CH_4	C_2H_6	C_2H_4	C_2H_2	Fault type	
						Actual fault	Fault analysis
(1)	93	58	43	37	0	Middle temperature overheating	Middle temperature overheating
(2)	139	52	6.8	63	9.6	Middle and low temperature overheating	Middle temperature overheating
(3)	19.6	320.7	279.2	574.7	0	High temperature overheating	High temperature overheating
(4)	457	79	54	190	2.4	High temperature overheating	High temperature overheating
(5)	279	41	18.1	42	31.8	High energy discharge	High energy discharge
(6)	14.67	3.68	10.54	2.71	0.2	Normal	Normal
(7)	345	112.25	27.5	51.5	58.75	Low energy discharge	Low energy discharge
(8)	217.5	40	4.9	51.8	67.5	High energy discharge	High energy discharge
(9)	181	262	210	528	0	Middle and low temperature overheating	Low temperature overheating
(10)	44.3	17.3	3.6	23.3	10.4	High energy discharge	High energy discharge
(11)	172.9	334.1	172.9	812.5	37.7	Partial discharge	Partial discharge
(12)	12.26	8.86	13.95	18.21	0	Normal	Low energy discharge
(13)	56	78	18	173	0	Low temperature overheating	Low temperature overheating

FIGURE 2: Convergence curve of MPSO.

can effectively improve the accuracy of transformer fault diagnosis.

(2) Tables 2 and 1 also show that the calculation time of presented method is averagely about 7.84 s for test data including 111 sample points (hardware configuration: CPU i5, RAM 4 G; programming software: Matlab), which demonstrates the efficiency of the presented method.

(3) Table 3 lists some test results using the presented method (due to limitation of paper length, the complete results are not listed here). The results show that diagnosis accuracy is basically satisfactory.

5. Conclusions

In this paper, a data-driven fault diagnosis model based on Kriging model and neural network is proposed. The proposed model is based on the Kriging model and integrates neural network technology. Meanwhile, the localized departures of Kriging model are retained.

The presented hybrid model is further used for power transformer fault diagnosis problems based on DGA method. Some numerical computations of transformer fault diagnosis problems are conducted, and the results show the feasibility and efficiency of the proposed method. In addition, the presented modified Kriging model is of some potential application value in other areas such as power system and engineering machinery.

Conflicts of Interest

The authors declare that the mentioned received funding did not lead to any conflicts of interest regarding the publication of this manuscript and there are not any possible conflicts of interest in the manuscript.

Acknowledgments

This study is supported by Program for Liaoning Excellent Talents in University (Grant no. LJQ2014037) and Natural Science Foundation of Liaoning Province of China (Grant no. 20170540589).

References

- [1] J. Pleite, C. Gonzalez, J. Vazquez, and A. Lazaro, "Power Transformer Core Fault Diagnosis Using Frequency Response Analysis," in *Proceedings of the MELECON 2006 - 2006 IEEE Mediterranean Electrotechnical Conference*, pp. 1126–1129, Benalmadena, Spain.
- [2] S.-W. Fei and X.-B. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Systems with Applications*, vol. 36, no. 8, pp. 11352–11357, 2009.
- [3] S. S. M. Ghoneim and I. B. M. Taha, "A new approach of DGA interpretation technique for transformer fault diagnosis," *International Journal of Electrical Power & Energy Systems*, vol. 81, pp. 265–274, 2016.
- [4] Y. J. Sun, S. Zhang, C. X. Miao et al., "Improved BP neural network for transformer fault diagnosis," *Journal of China*

- University of Mining & Technology*, vol. 17, no. 01, pp. 138–142, 2007.
- [5] S. Seifeddine, B. Khmais, and C. Abdelkader, “Power transformer fault diagnosis based on dissolved gas analysis by artificial neural network,” in *Proceedings of the 2012 1st International Conference on Renewable Energies and Vehicular Technology, REVET 2012*, pp. 230–236, March 2012.
 - [6] S. M. Islam, T. Wu, and G. Ledwich, “A novel fuzzy logic approach to transformer fault diagnosis,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 7, no. 2, pp. 177–186, 2000.
 - [7] Y.-C. Huang and H.-C. Sun, “Dissolved gas analysis of mineral oil for power transformer fault diagnosis using fuzzy logic,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 20, no. 3, pp. 974–981, 2013.
 - [8] C. E. Lin, J. M. Ling, and C. L. Huang, “An expert system for transformer fault diagnosis using dissolved gas analysis,” *IEEE Transactions on Power Delivery*, vol. 8, no. 1, pp. 231–238, 1993.
 - [9] Y.-C. Xiao, X.-H. Chen, and H.-J. Zhu, “Application of genetic support vector machine in power transformer fault diagnosis,” *Shanghai Jiaotong Daxue Xuebao/Journal of Shanghai Jiaotong University*, vol. 41, no. 11, pp. 1878–1886, 2007.
 - [10] A.-H. Zhou, H. Song, H. Xiao, and X.-H. Zeng, “Power transformer fault diagnosis based on integrated of rough set theory and neural network,” in *Proceedings of the 2nd International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2012*, pp. 1463–1465, January 2012.
 - [11] S. Jeong, M. Murayama, and K. Yamamoto, “Efficient optimization design method using kriging model,” *Journal of Aircraft*, vol. 42, no. 2, pp. 413–420, 2005.
 - [12] T. W. Simpson, T. M. Mauery, J. J. Korte, and F. Mistree, “Kriging models for global approximation in simulation-based multidisciplinary design optimization,” *AIAA Journal*, vol. 39, no. 12, pp. 2233–2241, 2001.
 - [13] S. Dey, T. Mukhopadhyay, and S. Adhikari, “Stochastic free vibration analyses of composite shallow doubly curved shells - A Kriging model approach,” *Composites Part B: Engineering*, vol. 70, no. 5, pp. 99–112, 2015.
 - [14] Z. Chen, S. Peng, X. Li et al., “An important boundary sampling method for reliability-based design optimization using kriging model,” *Structural and Multidisciplinary Optimization*, vol. 52, no. 1, pp. 55–70, 2015.
 - [15] Q. Liu and C. Wang, “Pipe-assembly approach for aero-engines by modified particle swarm optimization,” *Assembly Automation*, vol. 30, no. 4, pp. 365–377, 2010.
 - [16] H. M. Guo and A. Yang, “Diagnosis of transformer fault based on GA-BP neural network,” *Coal Mine Machinery*, vol. 36, no. 7, pp. 318–320, 2015.
 - [17] Y. B. Tang, *Data-Driven Fault Diagnosis and Prediction for Large-Scale Power Transformer*, Central South University, Changsha, China, 2013.
 - [18] W. Gao, *Study on Fault Diagnosis for Power Transformer Based on Support Vector Machine of Artificial Immune Algorithm*, Taiyuan University of Technology, Taiyuan, China, 2012.