

Fault Prediction using a Grey-Markov Model from the Dissolved Gases Contents in Transformer Oils

Yang Liu¹, Yu Du¹, Zhiwu Wang¹, Guangming Feng¹, Shaowei Rao², Guoping Zou², and Shiyong Yang^{2*}

¹ Equipment Management Centre, Suzhou Nuclear Power Research Institute, Suzhou, Jiangsu Province, 215004, China

² College of Electrical Engineering, Zhejiang University, Hangzhou, Zhejiang Province, 310027, China

Abstract. A novel method to predict transformer fault by forecasting the variation trend of the dissolved gases content is proposed. After the content of each feature gas, such as hydrogen and methane, is obtained by the proposed forecasting model, the fault type can be diagnosed by the dissolved gas analysis (DGA) technologies. Firstly, the GM (1,1) grey model with unequal time interval is introduced to generate a general forecasting model for each feature gas. The introduced grey model with unequal time interval will enforce no constrain on the historical measurement data. Consequently, the time intervals of the two adjacent measuring points can be either constant or variant. To address the deficiency that the existing grey model is unable to describe the fluctuation of the predicted object in time domain, the Markov chain is introduced to improve the accuracy of the grey forecasting model. An adaptive method to automatically divide the state space based on the number of states and the relative error of the grey model is presented by using Fibonacci sequences. Practical measurements are used to verify the accuracy of the proposed forecasting model. The numerical results show that there is high probability (86%) that the proposed grey-Markov model acquires a smaller prediction residual as compared to the original GM(1,1) grey model.

1 Introduction

Power transformers play an important role in a power system to transform the voltage from one level to another. Consequently, it is essential to ensure a safe operation of the power transformer since a transformer fault may result in an interruption of the power supply. The conventional approaches to detect the incipient faults of a transformer include the state assessment and fault diagnosis. However, both approaches are based on a real-time measured or recent tested data. Once a fault has been identified, the transformer will be out of service as soon as possible to avoid an incipient fault deteriorating into a catastrophic one. As a result, an operator will suffer a heavy stress to maintain the continuity of the power supply when this situation appears. In this point of view, the prediction of transformer faults is indispensable to conduct a reasonable maintenance to ensure the safe operation of a transformer.

So far, some efforts have been devoted to solve the issues related to transformer fault predictions [1]-[13]. All existing approaches for predicting transformer fault are along the lines of forecasting the concentration of feature gases. Grey model is used to construct the forecasting model of the dissolved gases contents [1]-[3]. The cloud-based reasoning and semi-Markov model are used to predict and diagnose transformer faults [4]. Hidden Markov model together with Gaussian mixture model is applied to the dynamic fault prediction of power

transformers [5]. The prediction problem of transformer faults is converted into a multi-dimensional regression one and then solved by robust optimizations [6]. Recently, more and more machine learning methods, especially support vector machine (SVM) [7]-[12] and artificial neural network (ANN) [13], have been utilized in transformer fault predictions.

Grey model is developed from the grey theory [14]. More specifically, the GM(1,1) grey model is one of the most commonly used forecasting models. In a GM(1,1) grey model, it is assumed that the accumulated generating operation (AGO) sequences satisfy a first-order differential equation. The forecasting model of the predicted object can be obtained by solving this differential equation. Actually, GM(1,1) grey model has already been used to predict transformer faults, and it seems that a good performance has been obtained [1]-[3]. For example, A grey model with unequal time intervals is proposed, and an interpolation method is introduced to reconstruct the measuring points in [1]. A grey model for feature ratio is introduced, and then the future fault can be diagnosed with the base of the predicted feature ratio by using DGA technologies [2]. A modified grey model by replacing 1-AGO with 2-AGO is proposed in [3].

In this paper, a novel method to predict transformer faults is developed. More specially, a general forecasting model for dissolved gases is developed. The time interval of each two adjacent measuring points can be either constant or variant, and the trend of the original data can be either monotone or fluctuant in time domain. The

*Corresponding author's e-mail: eesyang@zju.edu.cn

Grey-Markov model is obtained by combining the grey model and Markov chain to reduce the modelling error and to improve the accuracy of the forecasted results. A method to divide the state space automatically with the aid of Fibonacci sequences is proposed. The comparison results show that the proposed grey-Markov model can improve the forecasting accuracy effectively.

2 Forecasting model for dissolved gases of transformers

2.1 Grey forecasting model

GM(1,1) grey forecasting model has its inherent advantage of using only a few measuring points in time domain to generate the GM(1,1) grey model. However, the traditional GM(1,1) grey model requires some evenly distributed measuring points in the time interval. When the measuring points with an equal time interval are unavailable, the unevenly distributed measuring points are generally converted into evenly distributed measuring points by interpolations. However, this procedure will incur additional errors to the measurement data, thereby reducing the reliability of the grey model. The GM(1,1) grey model with unequal time interval is built by redefining AGO in this paper. The procedure of constructing a GM(1,1) grey model is shown in figure 1.

The details of each step to generate a GM(1,1) grey model are explained in the following paragraphs.

- **Accumulated generating operation**

The redefined AGO is given as follows:

$$\begin{aligned} \Delta t(1) &= 1 \\ \Delta t(k) &= t(k) - t(k-1) \quad (k = 2, \dots, n) \\ x^{(1)}(k) &= \sum_{i=1}^k \Delta t(i) \times x^{(0)}(i) \quad (k = 1, 2, \dots, n) \end{aligned} \quad (1)$$

where $\Delta t(k)$ is the time interval between the k^{th} measuring point and the $(k-1)^{\text{th}}$ measuring point, $x^{(0)}(i)$ is the measurement value of the i^{th} measuring point, $x^{(1)}(k)$ is the k^{th} point of the AGO sequences, n is the number of the total measuring points. It should be noted that the time interval shown in equation (1) is normalized before making the AGO. The original data sequences $x^{(0)}$ and AGO sequences $x^{(1)}$ are as follows:

$$\begin{aligned} x^{(0)} &= (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \\ x^{(1)} &= (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)) \end{aligned} \quad (2)$$

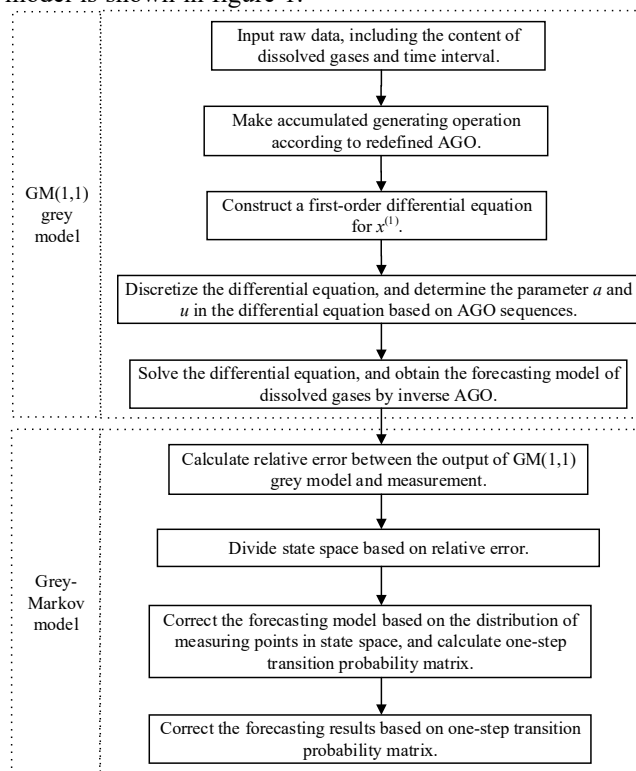


Figure 1 The flow chart in constructing the proposed forecasting model

- **A first-order differential equation for AGO sequences**

In GM(1,1) grey model, it is assumed that the trend of AGO sequences in time domain can be described by the following differential equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \quad (3)$$

where a and u are parameters to be determined later.

- **Discretization of the differential equation**

The purpose of discretizing the differential equation is to determine the value of parameters a and u in the equation. Replacing a differential by a difference, one obtains

$$\begin{aligned} dx^{(1)} &= x^{(1)}(k) - x^{(1)}(k-1) = \Delta t(k) \times x^{(0)}(k) \\ dt &= t(k) - t(k-1) = \Delta t(k) \quad (4) \\ \frac{dx^{(1)}}{dt} &= x^{(0)}(k) \quad (k = 2, \dots, n) \end{aligned}$$

When discretizing the differential equation (3), $x^{(1)}$ is substituted by an equal-weighted background value, and its expression is as follows:

$$x^{(1)} = \frac{1}{2}(x^{(1)}(k-1) + x^{(1)}(k)) \quad (5)$$

● **Determination of parameters a and u**

According to equation (4) and equation (5), when the value of k goes from 2 to n , following equation sets can be obtained:

$$\begin{aligned} x^{(0)}(2) + \frac{1}{2}a \times [x^{(1)}(1) + x^{(1)}(2)] &= u \\ x^{(0)}(3) + \frac{1}{2}a \times [x^{(1)}(2) + x^{(1)}(3)] &= u \\ &\vdots \\ x^{(0)}(n) + \frac{1}{2}a \times [x^{(1)}(n-1) + x^{(1)}(n)] &= u \end{aligned} \quad (6)$$

Rewrite these equations in a form of matrixes:

$$Y = B\theta \quad (7)$$

where

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -\frac{1}{2}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(x^{(1)}(n-1) + x^{(1)}(n)) & 1 \end{bmatrix}, \theta = \begin{bmatrix} a \\ u \end{bmatrix} \quad (8)$$

The parameters a and u can be obtained by solving $\theta = (B^T B)^{-1} B^T Y$.

● **Solution of the differential equation**

Substituting the value of parameters a and u into equation (3), the solution of the differential equation (3) can be obtained by solving an ordinary first-order differential equation, and its solution is:

$$\hat{x}^{(1)}(k) = [x^{(0)}(1) - \frac{u}{a}]e^{-a(t(k)-t(1))} + \frac{u}{a} \quad (k = 1, 2, \dots, n) \quad (9)$$

The forecasting model of the predicted object (the content of dissolved gases) can then be obtained by the inverse AGO:

$$\hat{x}^{(0)}(k) = \frac{\hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)}{\Delta t(k)} \quad (k = 2, 3, \dots, n) \quad (10)$$

Apparently, when $k=1$, according to the definition of the AGO and equation (9), there exists $\hat{x}^{(0)}(1) = \hat{x}^{(1)}(1) = x^{(0)}(1)$, which means that there is no modelling error in the first point in GM(1,1) grey model.

2.2 Grey-Markov forecasting model

Usually, the GM(1,1) grey model can not completely describe the trend of the predicted object. In other words, there are modelling errors between the outputs of the grey model and measurement results. The modelling errors would become more obvious when the original data fluctuates in time domain, since the essence of the output of GM(1,1) grey model is an exponential function. However, the content of dissolved gases is not always maintaining the trend of growth [15]. For this reason, it is necessary to modify the GM(1,1) grey model so that it can be applied to the case where the predicted object does not vary monotonously in time domain.

For the purpose of reducing modelling errors, the GM(1,1) grey model is combined with the Markov chain to generate a grey-Markov model. Grey-Markov model is firstly proposed in [16] and shows good performances in forecasts [17]. After grey model is built, the grey-Markov model can be generated following the procedures shown in figure 1. The detail of each step is described as:

● **Calculation of modelling errors**

The modelling errors between the outputs of GM(1,1) grey model and the measurement results are given by:

$$\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) \quad (k = 1, 2, \dots, n) \quad (11)$$

where $x^{(0)}(k)$ and $\hat{x}^{(0)}(k)$ are the measurement value and the output value of the grey model at the k^{th} point, respectively.

● **State space division**

If the number of states in the state space is recorded as m , the range of each state can be expressed as:

$$\begin{aligned} \text{The first state: } & \hat{x}^{(0)}(i) + \lambda_1 \bar{x}^{(0)} \leq x^{(0)}(i) < \hat{x}^{(0)}(i) + \lambda_2 \bar{x}^{(0)} \\ \text{The second state: } & \hat{x}^{(0)}(i) + \lambda_2 \bar{x}^{(0)} \leq x^{(0)}(i) < \hat{x}^{(0)}(i) + \lambda_3 \bar{x}^{(0)} \quad (12) \\ & \vdots \\ \text{The } m\text{th state: } & \hat{x}^{(0)}(i) + \lambda_m \bar{x}^{(0)} \leq x^{(0)}(i) < \hat{x}^{(0)}(i) + \lambda_{m+1} \bar{x}^{(0)} \end{aligned}$$

where $\bar{x}^{(0)}$ is the average of the measurement values and is calculated ones from the following equation:

$$\bar{x}^{(0)} = \sum_{i=1}^n x^{(0)}(i) / n \quad (13)$$

Parameter λ_j ($j = 1, 2, \dots, m+1$) satisfies following condition:

$$\lambda_1 < \lambda_2 < \dots < \lambda_{m+1} \quad (14)$$

It should be stressed that the value of λ_j should be determined according to the modelling errors. In this

paper, a new method to determine the value of λ_j is developed.

Firstly, the value of λ_{m+1} is determined from the following equations:

$$\lambda_{m+1} = 1.2 \times \max(|\delta_i|) \quad (i = 1, 2, \dots, n)$$

$$\delta_i = \frac{x^{(0)}(i) - \hat{x}^{(0)}(i)}{x^{(0)}(i)} \quad (15)$$

The principle of the state space division is that the width of each state subspace is proportional to its distance between the subspace and centerline (the line corresponding to $x^{(0)}(i)$). Besides, the state subspaces are symmetric based on the centerline. So, m is an even number.

In the case of $m=2$, there is no state space division problem. Therefore, one will discuss the division only when $m \geq 4$. In this study, the division of the state space is realized with the aid of Fibonacci sequences. Here, the recurrence formula of Fibonacci sequences is simply reviewed, and its recurrence formula is as follows:

$$\begin{aligned} F(0) &= 0 \\ F(1) &= 1 \\ F(k) &= F(k-1) + F(k-2) \end{aligned} \quad (16)$$

The procedure for dividing the state space is as follows: starting with the fourth number ($F(3)$) of the Fibonacci sequences, then taking $m/2$ numbers from $F(3)$, and using these $m/2$ numbers for state space bandwidth allocation. Firstly, the half-space ($\lambda \geq 0$) is divided as follows:

$$\lambda_{\frac{m}{2}+1} = 0$$

$$\lambda_{\frac{m}{2}+d} = \lambda_{\frac{m}{2}+d-1} + \lambda_{m+1} \times \frac{F(1+d)}{\sum_{k=3}^{\frac{m}{2}} F(k)} \quad (2 \leq d \leq \frac{m}{2} + 1) \quad (17)$$

Then, another half-space can be determined according to the symmetry of the state space. In order to make this process visible, an example of the state space division is shown in figure 2. In this example, there are four state subspaces, and each one is a banding zone.

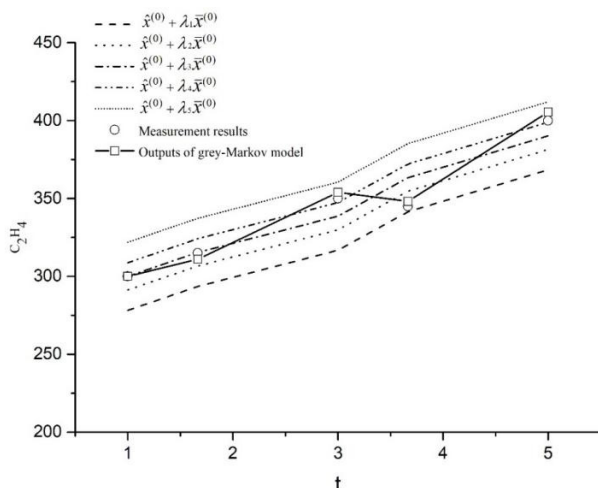


Figure 2 An example of state space division

● Calculation of the transition probability matrix

The procedure of calculating the transition matrix is detailed as: firstly, evaluate the number of the original data sequences $x^{(0)}(i)$ ($i=1, 2, \dots, n-1$) contained in each state subspace, and record as n_k ($k=1, 2, \dots, m$); Symbol n_{ks} represents the number of the original data sequences transiting from the k^{th} state into the s^{th} state within one step; then the one-step transition probability matrix can be calculated as follows:

$$P = \begin{bmatrix} \frac{n_{11}}{n_1} & \frac{n_{12}}{n_1} & \dots & \frac{n_{1m}}{n_1} \\ \frac{n_{21}}{n_2} & \frac{n_{22}}{n_2} & \dots & \frac{n_{2m}}{n_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{n_{m1}}{n_m} & \frac{n_{m2}}{n_m} & \dots & \frac{n_{mm}}{n_m} \end{bmatrix} \quad (18)$$

● Correction of the outputs of the grey model and forecasting results

Supposing that $x^{(0)}(i)$ falls into the j^{th} state, the output of the forecasting model is corrected as:

$$\tilde{x}^{(0)}(i) = \hat{x}^{(0)}(i) + \frac{1}{2}(\lambda_j \bar{x}^{(0)} + \lambda_{j+1} \bar{x}^{(0)}) \quad (19)$$

The state of the forecasting results can be determined based on the state of the last data point and the one-step transition probability matrix.

3Application

In order to verify the accuracy of the proposed method, some practical measurements are used to test the proposed forecasting model. The data shown in table 1 is extracted from [15]. The time interval of these measurements is variant. The seven feature gases result in seven forecasting models. A distinguishing feature of these original data is that the original data sequences fluctuate in time domain, resulting a big challenge in accurately forecasting the results from existing models and methods.

There are six measuring points for each feature gas. The first five measuring points are used to construct the forecasting model, and the last measuring point is used as the testing point. Table 1 shows the results obtained from the GM(1,1) grey model and the grey-Markov model. By comparing the modelling errors of the GM(1,1) grey model and the grey-Markov model, it is found that the proposed grey-Markov model has smaller modelling errors than the original GM(1,1) grey model in all of these seven cases. In view of the forecasted performance, apparently, the grey-Markov model shows a better forecast accuracy, since the grey-Markov model achieves a smaller prediction residual for tall six cases.

Moreover, if one observes the results shown in table 1 carefully, it is not hard to find that the outputs of the

GM(1,1) grey model change monotonously after the second point. This phenomenon can be easily understood since the essence of the outputs of the GM(1,1) grey model is an exponential function. However, the outputs of the proposed grey-Markov model fluctuate in time domain, which is similar to the original data sequences. Consequently, it can be concluded that the proposed grey-Markov model can trace the fluctuation of the predicted object.

4Conclusion

A novel forecasting model to predict the variation trend of the dissolved gases content in transformers is proposed. A general forecasting model for feature gases is developed based on the GM(1,1) grey model with unequal time interval to facilitate its application in uneven time interval measurements. In order to avoid the deficiency of being unable to describe the fluctuation of the predicted object in time domain of a GM(1,1) grey model, the GM(1,1) grey model is combined with a Markov chain to generate a grey-Markov model. The numerical results as reported show that the proposed grey-Markov model can achieve not only smaller modelling error but also better forecast performances.

Table 1 The comparison of the results obtained from GM(1,1) grey model and grey-Markov model

Date	1994-6-1	1994-6-3	1994-6-7	1994-6-9	1994-6-13	1994-6-17	Modeling error	Prediction residual
	1	2	3	4	5	6 (Forecasted)		
H_2	120	110	135	115	140	125	\	\
$\hat{x}^{(0)}$	120.00	114.62	121.05	127.80	134.97	145.12	5.85%	16.10%
$\tilde{x}^{(0)}$	120.00	111.31	132.64	116.21	138.29	141.81	1.04%	13.45%
CO	360	370	380	380	450	420	\	\
$\hat{x}^{(0)}$	360.00	358.69	381.79	406.21	432.37	469.76	2.87%	11.85%
$\tilde{x}^{(0)}$	360.00	365.11	375.37	383.73	454.85	447.28	0.92%	6.50%
CO_2	1500	1450	1500	1500	1700	1600	\	\
$\hat{x}^{(0)}$	1500.00	1424.70	1497.10	1572.75	1652.68	1765.29	1.91%	10.33%
$\tilde{x}^{(0)}$	1500.00	1442.51	1514.91	1510.42	1715.01	1702.95	0.62%	6.43%
CH_4	160	170	185	185	220	200	\	\
$\hat{x}^{(0)}$	160.00	167.62	181.59	196.58	212.96	236.83	2.54%	18.42%
$\tilde{x}^{(0)}$	160.00	170.38	184.35	186.91	222.63	227.15	0.56%	13.58%
C_2H_6	55	58	61	61	77	74	\	\
$\hat{x}^{(0)}$	55.00	55.63	60.98	66.78	73.20	82.68	3.71%	11.73%
$\tilde{x}^{(0)}$	55.00	57.05	62.40	61.81	78.16	77.71	1.36%	5.01%
C_2H_4	300	315	350	345	400	375	\	\
$\hat{x}^{(0)}$	300.00	315.38	338.63	363.38	390.16	428.80	2.23%	14.35%
$\tilde{x}^{(0)}$	300.00	311.01	353.93	348.07	405.46	413.50	0.93%	10.27%
C_2H_2	1.4	1.4	2.1	1.7	1.6	1.5	\	\
$\hat{x}^{(0)}$	1.40	1.63	1.63	1.62	1.61	1.60	8.83%	6.67%
$\tilde{x}^{(0)}$	1.40	1.34	1.93	1.32	1.53	1.30	7.82%	13.33%

$\hat{x}^{(0)}$: The result from GM(1,1) grey model, $\tilde{x}^{(0)}$: The result from grey-Markov model.

Acknowledgments

This work was supported by the Smart Nuclear Power Project of China General Nuclear Power Group under project number: R-2020SZEM21TF.

References

1. Wang, M.H., Hung, C.P. (2003) Novel grey model for the prediction of trend of dissolved gases in oil-filled power apparatus. *Electric Power Systems Research*, 67: 53-58.

2. Wang, J., Liu, J.X. (2007) Transformer fault prediction using new grey prediction model. *Journal of North China Electric Power University*, 34: 10-14.
3. Fei, S.W., Sun, Y. (2008) Fault Prediction of Power Transformer by Combination of Rough Sets and Grey Theory. *Proceeding of The CSEE*, 28: 154-160.
4. Zhou, Q., Sun, C., An, W.D., Li, J., Liao, J.S., Zou, D. (2015) Transformer Failure Prediction Based on Cloud Reasoning and Weighted Implicit Semi-Markov Model. *High Voltage Engineering*, 41: 2268-2275.
5. Jiang, J., Chen, R., Chen, M., Wang, W., Zhang, C. (2019) Dynamic fault prediction of power transformers based on hidden Markov model of dissolved gases analysis. *IEEE Transactions on Power Delivery*, 34: 1393-1400.

6. Kim, Y., Goyal, A., Kumar, T. (2016) Predictive modeling of dissolved gas concentration in oil-immersed substation transformers. In: 2016 IEEE Smart Energy Grid Engineering (SEGE). Ontario. pp. 261-267.
7. Kari, T., Gao, W., Tuluhong, A., Yaermaimaiti, Y., Zhang, Z. (2018) Mixed kernel function support vector regression with genetic algorithm for forecasting dissolved gas content in power transformers. *Energies*, 11: 2437.
8. Liao, R.J., Bian, J.P., Yang, L.J., Grzybowski, S., Wang, Y.Y., Li, J. (2012) Forecasting dissolved gases content in power transformer oil based on weakening buffer operator and least square support vector machine-Markov. *IET Generation, Transmission & Distribution*, 6: 142-151.
9. Zheng, H., Zhang, Y., Liu, J., Wei, H., Zhao, J., Liao, R. (2018) A novel model based on wavelet LS-SVM integrated improved PSO algorithm for forecasting of dissolved gas contents in power transformers. *Electric Power Systems Research*, 155: 196-205.
10. Huang, X.B., Jiang, W.T., Zhu, Y.C., Tian, Y. (2020) Transformer Fault Prediction Based on Time Series and Support Vector Machine. *High Voltage Engineering*, 46:2530-2538.
11. Zeng, B., Guo, J., Zhang, F., Zhu, W., Xiao, Z., Huang, S., Fan, P. (2020) Prediction model for dissolved gas concentration in transformer oil based on modified grey wolf optimizer and LSSVM with grey relational analysis and empirical mode decomposition. *Energies*, 13: 422.
12. Yuan, F., Guo, J., Xiao, Z., Zeng, B., Zhu, W., Huang, S. (2020) An Interval Forecasting Model Based on Phase Space Reconstruction and Weighted Least Squares Support Vector Machine for Time Series of Dissolved Gas Content in Transformer Oil. *Energies*, 13: 1687.
13. Pereira, F. H., Bezerra, F. E., Junior, S., Santos, J., Chabu, I., Souza, G. F. M. D., Micerino, F., Nabeta, S.I. (2018) Nonlinear autoregressive neural network models for prediction of transformer oil-dissolved gas concentrations. *Energies*, 11: 1691.
14. Deng, J.L. (1989) Introduction to grey system theory. *The Journal of grey system*, 1: 1-24.
15. Luo, Y.B., Yu, P., Song, B., Peng, Z.H., Lin, X.M. (2001) Prediction of The Gas Dissolved in Power Transformer Oil by The Grey Model. *Proceeding of the CSEE*, 21: 65-69.
16. He, Y., Bao, Y.D. (1992) Grey-Markov Forecasting Model and its Application. *System Engineering Theory and Practice*, 4: 59-63.
17. Huang, M., He, Y., Cen, H. (2007) Predictive analysis on electric-power supply and demand in China. *Renewable Energy*, 32: 1165-1174.