



# A Novel Power Transformer Fault Diagnosis Model Based on Harris-Hawks-Optimization Algorithm Optimized Kernel Extreme Learning Machine

Xiaohui Han<sup>1</sup> · Shifeng Ma<sup>1</sup> · Zhewen Shi<sup>1</sup> · Guoqing An<sup>1</sup> · Zhenbin Du<sup>2</sup> · Chunlin Zhao<sup>3</sup>

Received: 25 July 2021 / Revised: 4 December 2021 / Accepted: 4 January 2022 / Published online: 20 January 2022  
© The Author(s) 2022

## Abstract

Dissolved gas analysis (DGA) method is widely used to detect the incipient fault of power transformers. This paper presents a novel DGA method for power transformer fault diagnosis based on Harris-Hawks-optimization (HHO) algorithm optimized kernel extreme learning machine (KELM). The non-code ratios of the gases are used as the characterizing vector for the KELM model, and the Harris-Hawks-optimization (HHO) algorithm is introduced to optimize the KELM parameters, which promotes the fault diagnostic performance of KELM. Based on dataset collected from IEC TC 10, the fault diagnosis capability of the proposed method is validated by different characterizing vectors and is compared with conventional KELM and other optimized KELM. Moreover, the generalization ability of the proposed method is confirmed by China DGA data. The results demonstrate that the proposed method is superior to other methods and is more effective and stable for power transformer fault diagnosis with high accuracy.

**Keywords** Dissolved gas analysis (DGA) · Kernel extreme learning machine (KELM) · Harris-Hawks optimization (HHO) · Transformer fault diagnosis

## 1 Introduction

The power transformer plays a vital role in power systems, and it serves as the connection of transmission and distribution networks at different voltage levels [1]. Once it fails, it will cause power interruption, which seriously influence economic production and living activities [2]. Thus, it is of great necessity to detect and diagnose incipient fault in power transformers [3].

In recent years, diagnosis incipient fault in power transformers has been studied by many scholars and experts and transformer fault diagnosis problem is solved with different

approaches. The aim of all these studies is to meet the reliability requirements of power supply in order to minimize device damages and reduce economic, social and personal losses. Various approaches have been used to solve transformer fault diagnosis problem such as frequency response analysis [4], the vibration analysis [5], dissolved gas analysis (DGA) [6], etc.

Among these methods, dissolved gas analysis (DGA) is a widely used technique to detect the incipient faults in power transformers [7, 8]. Many researchers have put forward a large number of DGA-based methods. Traditional methods based on DGA such as IEEE key gases [9], Rogers ratios [10], IEC standard code [11], Dornenburg ratios [12], Duval triangle [13], IEC 60599 [14], have been applied to transformer fault diagnosis.

In addition to these traditional methods, artificial intelligence (AI) methods based on DGA have been recently used in the field of transformer fault diagnosis, including artificial neural networks(ANN) [15], support vector machine(SVM) [16], fuzzy logic (FL) [17], adaptive neuro fuzzy inference system (ANFIS) [18], random forest (RF) [19], Bayesian network [20], gene expression programming(GEP) [21], time series analysis [22], expert system [23], association rule

✉ Guoqing An  
anguoqing99@sina.com

<sup>1</sup> College of Electric and Engineering, Hebei University of Science and Technology, Shi Jiazhuang 050018, China

<sup>2</sup> Hebei Provincial Key Laboratory of Electromagnetic & Structural Performance of Power Transmission and Transformation Equipment, Baoding Tianwei Baobian Electric Co., Ltd, Bao Ding 071056, China

<sup>3</sup> Hebei Weixun Electric Power Automation Equipment Co., Ltd, Heng Shui 053000, China

[24], set pair analysis [25], evidential reasoning [26], etc. The various proposed methods have produced good results and made an important contribution in the field of transformer fault diagnosis. These methods tend to build models that can bridge the relationship between machines data and machines health states [27].

Nowadays, deep learning theories have also been proposed for transform fault diagnosis with the development of calculation capacity [28]. Deep learning can realize end-to-end learning and avoid tedious feature engineering and build a black box model compared with traditional AI methods. However, it requires a lot of iterations and weight updates with time-consuming gradient computing. A new model based on extreme learning machine (ELM) was introduced to identify the state of transformer to solve the efficiency degradation caused by gradient update [29]. The ELM model outperformed ANN and other traditional methods. However, stochastic selection of the number of hidden layer nodes is easily leads to low stability and poor robustness of the model in the application [30].

By analyzing the reviewed articles, despite different diagnosis models have been presented by various scholars and experts, a precise diagnosis model is still needed yet. For this purpose, a novel fault diagnosis model based on HHO-KELM for power transformers is proposed in this study.

$$\| \mathbf{H}(w_1, \dots, w_L, b_1, \dots, b_L) \hat{\mathbf{B}} - \mathbf{T} \| = \min_{\hat{\mathbf{B}}} \| \mathbf{H}(w_1, \dots, w_L, b_1, \dots, b_L) \mathbf{B} - \mathbf{T} \| \tag{4}$$

The contributions of this study are as follows: (1) Non-code ratios set is selected as the characterizing vector for KELM model. (2) A parameter optimization method based on HHO is adopted to obtain the optimal model parameters of KELM automatically. (3) The established HHO-KELM model is applied to diagnose incipient fault in power transformers.

The remaining parts of the paper are organized as follows: Sect. 2 describes the proposed method; Sect. 3 presents the proposed fault diagnosis approach based on HHO-KELM; In Sect. 4, the results and analysis discussion are given. Section 5 concludes the summary.

## 2 The Proposed Method

### A. Kernel Extreme Learning Machine

Kernel extreme learning machine (KELM) [30] is a kind of machine learning algorithm with strongest ability and generalization ability which is proposed on the basis of ELM and kernel functions. On the basis of ELM, the learning method of KELM is described as follows:

Given  $N$  different samples  $(x_j, t_j) \in R^n \times R^m, j = 1, \dots, N$ , where  $x_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T$  is the input samples,  $t_j = [t_{j1}, t_{j2}, \dots, t_{jm}]^T$  is the encoded class label. The mathematical model of ELM can be described as below:

$$\sum_{i=1}^L \beta_i g(x_j, w_i, b_i) = t_j, \quad j = 1, \dots, N \tag{1}$$

where,  $L$  is the neurons number in the hidden layer,  $w_i, b_i, \beta_i$  are the input weights, hidden layer bias and output weights of the  $i$ th hidden neuron node respectively,  $g(x_i, w_i, b_i)$  is the activation function. Equation (1) can be written in a matrix form as:

$$\mathbf{HB} = \mathbf{T} \tag{2}$$

where  $\mathbf{B} = [\beta_1, \dots, \beta_L]^T$  represents the vector of output layer weights,  $\mathbf{T} = [t_1, \dots, t_N]^T$  is the corresponding coding class label, and  $\mathbf{H}$  is the hidden layer output matrix:

$$\mathbf{H} = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(x_1, w_1, b_1) \cdots g(x_1, w_L, b_L) \\ \vdots \vdots \vdots \\ g(x_N, w_1, b_1) \cdots g(x_N, w_L, b_L) \end{bmatrix} \tag{3}$$

The training of ELM is equivalent to solving the least squares solution of linear Eq. (2), which is defined as:

The least square solution of the output weight is calculated as:

$$\hat{\mathbf{B}} = \mathbf{H}^T (I/C + \mathbf{H}\mathbf{H}^T)^{-1} \mathbf{T} \tag{5}$$

where  $C$  is the regularization parameter. The output function for ELM is implemented as:

$$f(x) = h(x) \hat{\mathbf{B}} = h(x) \mathbf{H}^T (I/C + \mathbf{H}\mathbf{H}^T)^{-1} \mathbf{T} \tag{6}$$

When the hidden layer function  $h(x)$  is unknown, the kernel function matrix is calculated as:

$$\begin{cases} \Omega_{\text{ELM}} = \mathbf{H}\mathbf{H}^T \\ \Omega_{ij} = h(x_i) \cdot h(x_j) = K(x_i, x_j) \end{cases} \tag{7}$$

where  $K(x_i, x_j)$  represents the kernel function. The radial basis function (RBF) is applied in this paper, which is given by:

$$K(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{\sigma} \right) \tag{8}$$

Then, the output function of KELM can be obtained as:

$$f(x) = h(x)H^T(I/C + HH^T)^{-1}T = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T (I/C + \Omega_{ELM})^{-1}T \tag{9}$$

### B. Harris Hawks Optimization Algorithm

Harris hawks optimization (HHO) is a newly introduced metaheuristic optimization technique by Heidari in 2019 [31], which contains the phases of exploration and exploitation.

#### 1. Exploration Phase

In this stage, Harris hawks detect the prey through two strategies. The first strategy assumes that the hawks allocate close to the team members and the prey. In the second strategy, the hawks place on random trees. These two strategies can be described as:

$$X(t + 1) = \begin{cases} X_{rand}(t) - r_1|X_{rand}(t) - 2r_2X(t)| & K \geq 0.5 \\ (X_{prey}(t) - X_m(t)) - r_3(LB + r_4(UB - LB)) & K < 0.5 \end{cases} \tag{10}$$

where  $X(t)$  and  $X(t + 1)$  represent the hawk's positions at the current iteration  $t$  and the next iteration  $t + 1$ , respectively.  $X_{prey}(t)$  is the position of prey.  $X_{rand}(t)$  is a randomly selected hawk from the current position.  $r_1, r_2, r_3, r_4$  and  $K$  are randomly generated number between  $[0, 1]$ .  $LB$  and  $UB$  are the lower and upper bounds of control variables.  $X_m(t)$  is the average position of hawks which calculated as:

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \tag{11}$$

where,  $N$  is the total number of hawks.  $X_i(t)$  is the position of each hawk in iteration  $t$ .

#### 2. Ransition from Exploration to Exploitation

The change from the exploration to exploitation in the HHO can be modeled based on the prey escaping energy as:

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \tag{12}$$

where,  $E$  indicates the escaping energy of the prey.  $E_0$  is the initial energy of the prey which is generated randomly between  $[-1, 1]$ .  $T$  is the maximum number of iterations. Therefore, if  $E < 1$ , this means the prey is weak, hence the hawks should exploit close to the prey location, and if  $E > 1$ , this means the prey has enough energy to escape, so the hawks should continue exploring the prey location.

### 3. Exploitation Phase

The exploitation process of HHO has been modeled based on the escaping energy of prey  $E$  and its chance for escaping  $r$ . There are four chasings besiege in the exploitation phase of HHO.

a: *Soft Besiege* ( $r \geq 0.5$  and  $|E| \geq 0.5$ ).

In soft besiege phase, the prey tries to escape using random jumps, but the hawks surround it softly. This process can be expressed as:

$$X(t + 1) = \Delta X(t) - E \left| KX_{prey}(t) - X(t) \right| \tag{13}$$

$$X(t + 1) = X_{prey}(t) - X(t) \tag{14}$$

where  $\Delta X(t)$  is the distance between the prey position and the hawks' position in iteration  $t$ .  $K = 2(1 - r_5)$  indicates the strength of the prey randomly jumping during the escaping.  $r_5$  is a random number ranged from  $[0, 1]$ .

b: *Hard Besiege* ( $r \geq 0.5$  and  $|E| < 0.5$ ).

In hard besiege mode, the prey becomes too exhausted to escape. As a result, the hawks effortlessly catch the prey, and then pounce on it. This move can be modeled as:

$$X(t + 1) = X_{prey}(t) - E|\Delta X(t)| \tag{15}$$

c: *Soft Besiege with Progressive Rapid Dives* ( $r < 0.5$  and  $|E| \geq 0.5$ ).

In soft besieges with progressive rapid dives approach, the prey able to escape and the hawks softly surround it. To model this besiege, a levy fight (LF) concept is used as follows:

$$H = X_{prey}(t) - E \left| KX_{prey}(t) - X(t) \right| \tag{16}$$

where  $H$  is the soft besiege positions. The hawks dive based on the LF as:

$$G = H + S \times LF(D) \tag{17}$$

where  $D$  is the dimension of problem.  $S$  is a random vector by size  $1 \times D$ . The LF is calculated as:

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}} \tag{18}$$

$$\sigma = \left( \frac{\Gamma(1 + \beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right)^{\frac{1}{\beta}} \tag{19}$$

where  $\beta$  is a constant value set to 1.5.  $u$  and  $v$  are random values between  $[0, 1]$ .

Finally, the updating position of hawks is calculated as:

$$X(t + 1) = \begin{cases} H & \text{if } F(H) < F(X(t)) \\ G & \text{if } F(G) < F(X(t)) \end{cases} \quad (20)$$

d: *Hard Besiege with Progressive Rapid Dives* ( $r < 0.5$  and  $|E| < 0.5$ ).

In hard besieges with progressive rapid dives case, the behavior of the prey is similar to its behavior in the case of a soft besiege. The position of team members is updated by reducing the distance between their average position and the prey position. This motion is described as:

$$X(t + 1) = \begin{cases} H' & \text{if } F(H') < F(X(t)) \\ G' & \text{if } F(G') < F(X(t)) \end{cases} \quad (21)$$

$$H' = X_{prey}(t) - E |KX_{prey}(t) - X_m(t)| \quad (22)$$

$$G' = H' + S \times LF(D) \quad (23)$$

where  $X_m(t)$  is obtained from Eq. (11).

### C. The Proposed HHO-KELM Method

The diagnosis performance of KELM is affected by the penalty factor  $C$  and kernel parameter  $\sigma$ . In order to obtain the optimal parameters of the KELM, we optimize the penalty factor  $C$  and the parameter  $\sigma$  of the kernel function by HHO. The specific steps of HHO optimizing KELM parameters are described as follows:

Step 1: Initialize the KELM parameters  $C$ ,  $\sigma$  and the parameters of HHO, including the number of search agents  $N$ , the number of iterations  $T$  and  $D$ -dimensional space.

Step 2: Calculate the objective function value of each hawks through the Eq. (9) and update the position of  $X_{prey}$ . The fitness function of the HHO in the optimization process is the mean square error (MSE), which is:

$$MSE = \frac{1}{M} \sum_{i=1}^M (y_D(i) - y(i))^2 \quad (24)$$

where  $y_D$  and  $y$  represent the measure value and the observed value.  $M$  represents the number of observation samples.

Step 3: Update the initial energy, jump strength and escaping energy of each individual in each iteration.

Step 4: Determine whether the number of iterations is equal to the maximum value. If it is not, execute  $t = t + 1$  and return to step 2. Otherwise, go to step 5.

Step 5: Export HHO optimal solutions (values of  $C$  and  $\sigma$ ).

The flow chart of this procedure is described in Fig. 1.

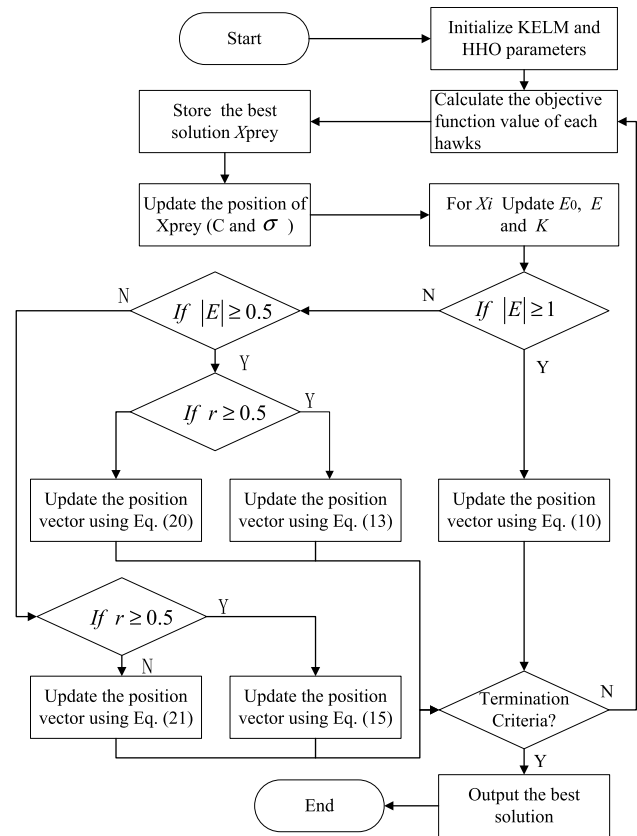


Fig. 1 The flowchart of HHO optimizing KELM parameters

## 3 Transformer Fault Diagnosis

### 3.1 Model Based on HHO-KELM

#### A. Transformer Fault Types and Input Characterizing Vector Selection

Faults inside transformers are classified in thermal faults and electrical faults according to [32]. Thermal faults are low and medium thermal (LM-T) and high thermal (H-T) categorized by means of temperature in the range of TLM-T < 700 °C and TH-T > 700 °C, respectively. Electrical faults are low-energy discharge (LE-D) and high-energy discharge (HE-D). In addition this work takes the normal state (NS) into account to distinguish between normal mode and failure mode.

The above incipient faults caused by oil and insulation paper decomposition can generate several kinds of gases: hydrogen (H<sub>2</sub>), methane (CH<sub>4</sub>), acetylene (C<sub>2</sub>H<sub>2</sub>), ethylene (C<sub>2</sub>H<sub>4</sub>), ethane (C<sub>2</sub>H<sub>6</sub>), carbon monoxide (CO) and carbon dioxide (CO<sub>2</sub>). Based on this gases, Non-code ratios are taken as the input characterizing parameter of the diagnostic model in this paper. Non-code ratios include CH<sub>4</sub>/H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>/C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>/C<sub>2</sub>H<sub>6</sub>, C<sub>2</sub>H<sub>2</sub>/(C<sub>1</sub> + C<sub>2</sub>), H<sub>2</sub>/(H<sub>2</sub> + C<sub>1</sub> + C<sub>2</sub>), C<sub>2</sub>H<sub>4</sub>/

$(C_1 + C_2)$ ,  $CH_4/(C_1 + C_2)$ ,  $C_2H_6/(C_1 + C_2)$  and  $(CH_4 + C_2H_4)/(C_1 + C_2)$ . Of these,  $C_1$  represents  $CH_4$  and  $C_2$  represents the sum of  $C_2H_4$ ,  $C_2H_6$  and  $C_2H_2$ .

**B. Transformer Fault Diagnosis Model Based on HHO-KELM**

In this study, a fault diagnosis model based on HHO-KELM for transformers is established. It can be divided into three parts: data collection and processing, the training of the network, and the performance testing and evaluation of the network model.

In the data collection and processing part, some dissolved gas content of transformer is firstly collected from the transformer faults records as DGA data, the collected DGA data are then transformed to non-code ratios. When the ratio with 0/0 is set to “0” and c/0 to 20, where c is not “0” [33]. The obtained DGA ratios are then normalized to clear up the influence of order of magnitude, shown as:

$$x_{si} = \frac{x_i - x_{i\min}}{x_{i\max} - x_{i\min}} \tag{25}$$

where  $x_{si}$  and  $x_i$  are the DGA ratios before and after normalization,  $x_{i\max}$  and  $x_{i\min}$  and are the maximum and minimum value of a DGA ratio.

In the network training part, a training dataset is constructed by using DGA records with transformer fault types. The characterizing vectors are input to the KELM network. The KELM adopts the HHO algorithm to obtain the model parameters. The learnt feature shows the relationships between the DGA ratios and the corresponding fault patterns.

After the training process, the diagnostic parameters are obtained. Then, the types of faults are classified for the samples in the testing set or the transformer under investigation with the trained diagnostic parameters. Accuracy is taken into consideration to appraise the capability of fault diagnosis approach which is computed by:

$$A = \frac{N_{correct}}{N_{total}} \tag{26}$$

where  $N_{correct}$  is the number of samples correctly diagnosed,  $N_{total}$  is the total number of samples in the whole dataset.

The implementation process of the proposed fault diagnosis model is described as follows:

- Step1: Obtain the DGA data of transformers through the acquisition device.
- Step2: Transform DGA data to Non-code ratios.
- Step3: Normalize the obtained DGA ratios and construct a feature vector.
- Step4: Divide the data into a training set and testing set. The training set is used to train a diagnostic model.

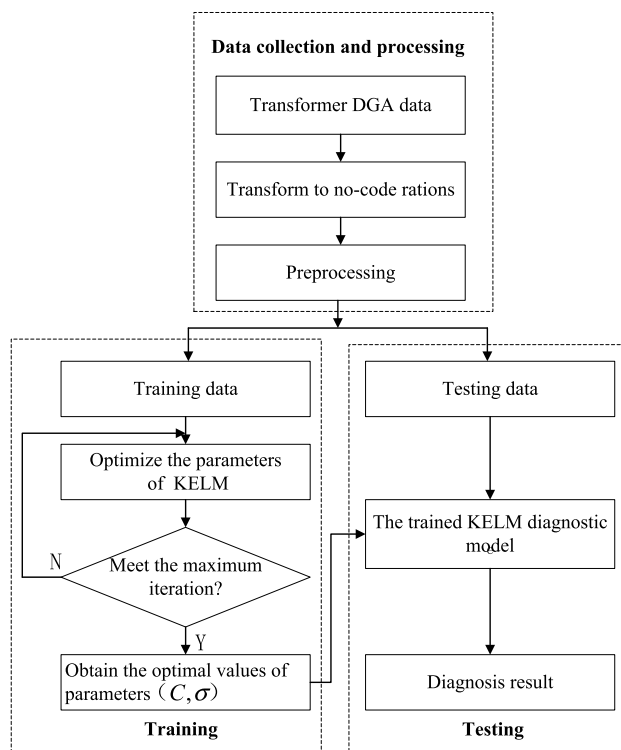


Fig. 2 The flow chart of the proposed fault diagnosis model

Table 1 Test data of IEC TC 10 database

Fault types	LM-T	H-T	LE-D	HE-D	NS
IEC TC 10 database	10	14	23	45	26

Step5: The testing set is selected to test and evaluate the performance of the proposed model.

The flow chart of the proposed transformer fault diagnosis model is shown in Fig. 2.

**4 Results and Analysis**

**A. Test Data**

DGA samples are collected from IEC TC 10 database [11], as shown in Table 1. These samples are used to evaluate the performance of the proposed fault diagnosis method.

**B. Model Parameters**

The relevant parameters of HHO and KELM for transformer fault diagnosis are detailed in Table 2.

**C. Transformer Fault Diagnosis Results**

**Table 2** Parameters of KELM and HHO

Algorithms	Parameters	Values
HHO	The population size $N$	20
	The maximum iteration number $T$	100
	Dimension $D$	2
KELM	The range of penalty factor $C$	[0.1, 50]
	The range of kernel parameter $\sigma$	[0.1, 10]

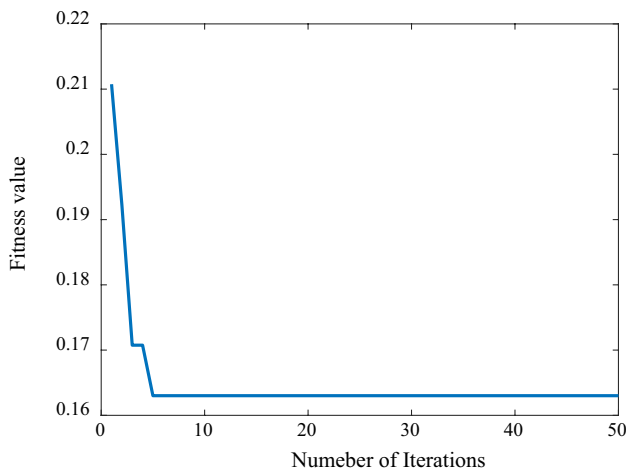
In order to investigate the accuracy of HHO-KELM for transformer fault diagnosis, 93 samples out of total 118 cases are used to train KELM structure parameters and the rest 25 samples (5 samples of each fault) are employed to test diagnosis capability of the trained KELM.

Figure 3 reports the typical HHO fitness variation curves. It is found that the converge time is 11.2693 s and the fitness value (mean square error) reaches 0.1708 after the second iteration, and then remain flat for a while. After the fifth iteration, the fitness value is reduced to 0.1630, which is the best value. This illustrates that the HHO algorithm can converge to the best solution very quickly.

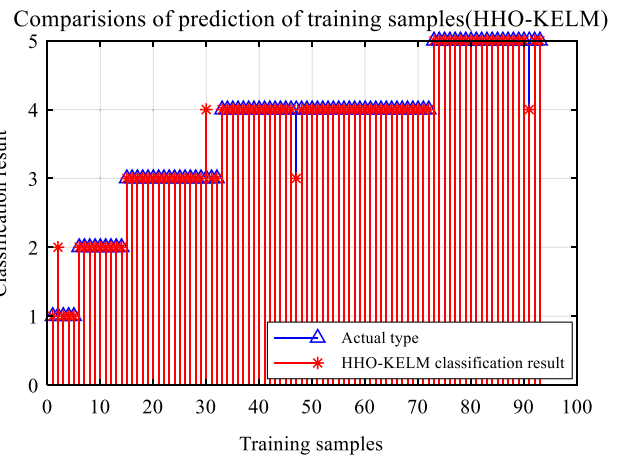
The diagnosis results based on HHO-KELM are present as Fig. 4. It can be seen that HHO-KELM has four errors in training samples, and the classification accuracy is 95.7% (89/93). In the test samples, three errors are generated in the type of faults, and the classification accuracy is 88% (22/25). This indicates that the proposed method has good accuracy.

**D. Comparisons with Different Input Characterizing Vector**

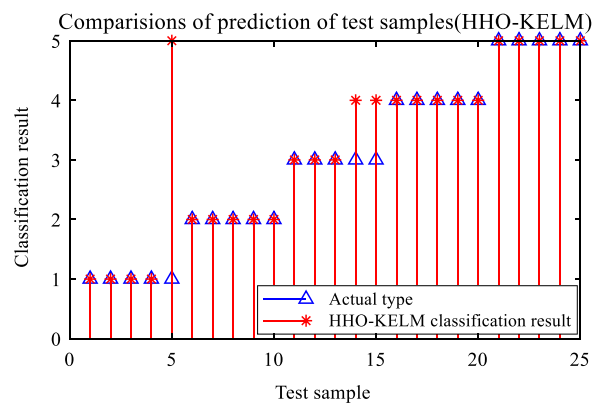
In order to further determine the proposed method, the input characterizing vectors of HHO-KELM are divided into



**Fig. 3** The fitness variation curves of HHO



**(a) The diagnosis results on training samples**



**(b) The diagnosis results on test samples**

**Fig. 4** The diagnosis results using HHO-KELM model. **a** The diagnosis results on training samples. **b** The diagnosis results on test samples

four categories: (1) All the DGA data ( $H_2$ ,  $CH_4$ ,  $C_2H_2$ ,  $C_2H_4$ ,  $C_2H_6$ ,  $CO$ ,  $CO_2$ ); (2) IEC Ratios ( $CH_4/H_2$ ,  $C_2H_2/C_2H_4$  and  $C_2H_4/C_2H_6$ ); (3) Rogers Ratios ( $CH_4/H_2$ ,  $C_2H_2/C_2H_4$ ,  $C_2H_4/C_2H_6$  and  $C_2H_6/CH_4$ ); (4) Non-code Ratios. The above four different types of characterizing parameters are modeled in KELM for diagnosis and the fault diagnosis results are reported in Table 3 and Fig. 5.

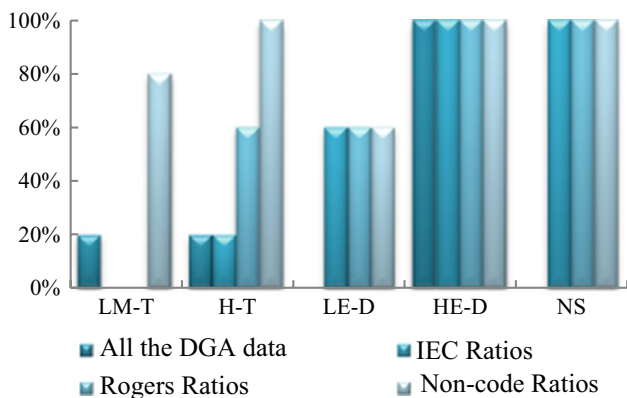
As seen in Table 3 and Fig. 5, HHO-KELM with Non-code ratios can detect LM-T better than others. HHO-KELM with IEC ratios and Rogers ratios can not detect LM-T properly. HHO-KELM with Non-code ratios has the best performance for H-T as well. The accuracy of HHO-KELM with Non-code ratios for LE-D is same to HHO-KELM with IEC ratios and Rogers ratios. All methods except HHO-KELM with all the DGA data, can detect HE-D and N-S properly.

In short, the HHO-KELM with Non-code ratios shows the highest total accuracy. With the input vector of Non-code ratios, the average accuracy of HHO-KELM is increased significantly compared to HHO-KELM with all the DGA data,



**Table 3** HHO-KELM diagnosis results using different characterizing vectors

Faults	All the DGA data (%)	IEC Ratios (%)	Rogers Ratios (%)	Non-code Ratios (%)
LM-T	20	0	0	80
H-T	20	20	60	100
LE-D	0	60	60	60
HE-D	100	100	100	100
NS	0	100	100	100
Average	28	56	64	88



**Fig. 5** HHO-KELM diagnosis results using different characterizing vectors

IEC ratios and Rogers ratios, the average accuracy for testing samples is increased by 60%, 32% and 24%, respectively.

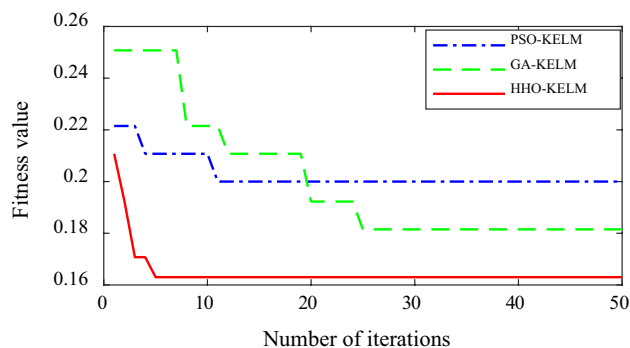
**E. Comparisons with Different Diagnostic Methods**

In order to verify the diagnosis performance, the ELM, KELM, GA-KELM and PSO-KELM with the input vector of Non-code ratios are selected to compare with the HHO-KELM.

Figure 6 shows the relationship between the fitness value (mean square error) and the number of iterations for different algorithms in the optimization process for KELM. It can be seen from Fig. 6 that the HHO-ELM model reduces the number of iterations and can find a stable solution close to the best goal. The reason the original ELM, KELM are not represented in the iteration number versus fitness plot is that the original ELM, KELM did not incorporate an intelligent optimization algorithm and did not ensure fitness.

The fault diagnosis results for five different methods are shown in Table 4 and Fig. 7.

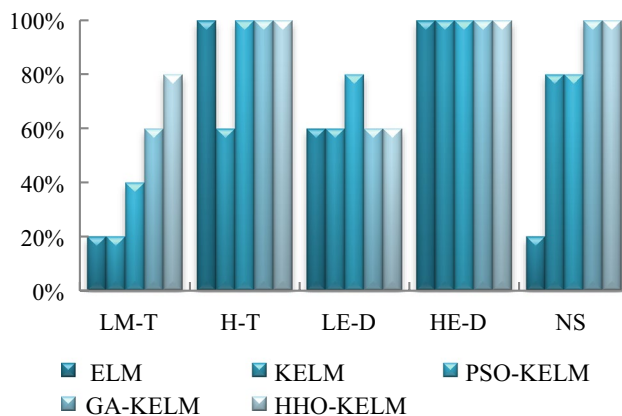
According to Table 4 and Fig. 7, HHO-KELM has better performance than other methods for detection of LM-T. The worst method for detecting LM-T is ELM and KELM. ELM,



**Fig. 6** The relationship between the fitness value and the number of iterations for different algorithms

**Table 4** Diagnosis results using different methods

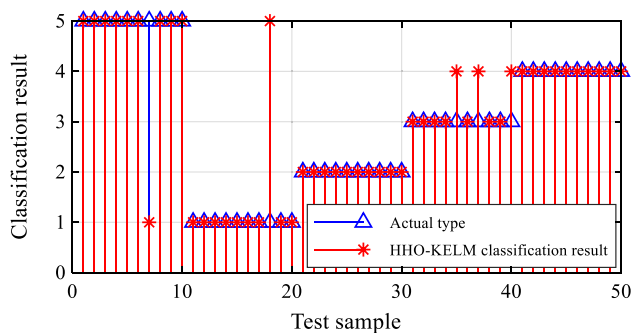
Faults	ELM (%)	KELM (%)	PSO-KELM (%)	GA-KELM (%)	HHO-KELM (%)
LM-T	20	20	40	60	80
H-T	100	60	100	100	100
LE-D	60	60	80	60	60
HE-D	100	100	100	100	100
NS	20	80	80	100	100
Average	60	64	80	84	88



**Fig. 7** Diagnosis results using different methods

PSO-KELM, GA-KELM and HHO-KELM can effectively diagnose H-T fault. However, HHO-KELM method is weak in detecting LE-D and is lower than PSO-KELM. All methods have better performance for HE-D. For N-S, GA-KELM and HHO-KELM have better accuracy than ELM, KELM and PSO-KELM.

It can be concluded that HHO-KELM has the best average accuracy. Compared with ELM, KELM, GA-KELM and PSO-KELM, the average fault diagnosis accuracy of



**Fig. 8** Fault diagnosis results of DGA samples in China

HHO-KELM is increased by 28%, 24%, 8%, and 4% for testing set, which show that the HHO-KELM has higher classification accuracy than those of the other compared method.

## F. Fault Diagnosis Using China Transformers DGA Data

In order to investigate fault diagnosis capability of the proposed method comprehensively, 180 DGA samples of transformers in China are diagnosed by HHO-KELM. The 180 samples are divided into two parts, of which 130 samples are used to train HHO-KELM model and the other 50 samples (10 samples of each fault) are tested with the trained model. Figure 8 presents a high accuracy of 90% (46/50) for China DGA data, which shows that the proposed method has better performance.

## 5 Conclusion

In this work, a fault diagnosis method based on HHO-KELM is presented to diagnose the different fault states of transformers. The proposed model is solved using non-code ratios as the input vector. Furthermore, HHO algorithm is implemented to optimize the KELM model parameters. Finally, some numerical analysis are proposed in order to demonstrate the validity of the proposed strategy. The comparison results show that the average test accuracy of the Non-code ratios feature set is 60%, 32% and 24% higher than that of the DGA full data, the IEC ratios and the Rogers ratios feature quantity respectively. Also, the accuracy of the HHO-KELM model for transformer fault diagnosis reaches 88% which is higher than the model of ELM, KELM, PSO-KELM model and GA-KELM. Considering the presented results, it can be said that the proposed method outperforms all compared diagnosis models and is an accurate model for solving such problems.

However, some shortcomings still existence in the study, such as the insufficient data samples. Additional datasets

could be considered in future work in order to validate the proposed method by means of fault diagnosis accuracy.

**Acknowledgements** The authors would like to express their appreciation of the support by Hebei provincial Key Research and Development Program (20312101D).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Duval M (1989) Dissolved gas analysis: it can save your transformer. *IEEE Electr Insul Mag* 5:22–27
2. Li YC, Ma LQ (2019) Fault diagnosis of power transformer based on improved particle swarm optimization OS-ELM. *Arch Electr Eng* 68(1):161–172
3. Mahmoudi N, Samimi MH, Mohseni H (2019) Experiences with transformer diagnosis by DGA: case studies. *IET Gener Transm Distrib* 13:5431–5439
4. Premalatha V, Sruthi SS, Vikash M et al (2012) Instrument transformer winding fault analysis using frequency response analysis. *IEEE international conference on advanced communication control and computing technologies (ICACCCT)*, Syed Ammal Engn Coll, Ramanathapuram, INDIA, August, 370–374
5. Bagheri M, Zollanvari A, Nezhivenko S et al (2018) Transformer fault condition prognosis using vibration signals over cloud environment. *IEEE Access* 6:9862–9874
6. Apte S, Somalwar R, Wajirabadkar A et al (2018) Incipient fault diagnosis of transformer by DGA using fuzzy logic. *2018 IEEE international conference on power electronics, drives and energy systems (PEDES)*, Chennai, India, December, 1–5
7. Benhamed K, Mooman A, Younes A, Shaban K, El-Hag A (2018) Feature selection for effective health index diagnoses of power transformers. *IEEE Trans Power Deliv* 33(6):3223–3226
8. Jiang J (2018) TDLAS-based detection of dissolved methane in power transformer oil and field application. *IEEE Sens J* 18(6):2318–2325
9. IEEE draft guide for the interpretation of gases generated in oil-immersed transformers. *IEEE Standard C57.104*, 2008
10. Rogers RR (1978) IEEE and IEC codes to interpret incipient faults in transformers using gas in oil analysis. *IEEE Trans Dielectr Electr Insul* 13:349–354
11. Guide for the interpretation of dissolved gas analysis and gas-free (2007). *IEC Standard*, 60599
12. Bacha K, Souahlia S, Gossa M (2012) Power transformer fault diagnosis based on dissolved gas analysis by support vector machine. *Electr Power Syst Res* 83(1):73–79
13. Irungu GK, Akumu AO, Munda JL (2016) A new fault diagnostic technique in oil-filled electrical equipment; the dual of Duval triangle. *IEEE Trans Dielectr Electr Insul* 23(6):3405–3410



14. Duval M, Depabla A (2001) Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. *IEEE Electr Insul Mag* 17(2):31–41
15. Sherif SM, Ibrahim BM, Nagy IE (2016) Integrated ANN-based proactive fault diagnostic scheme for power transformer using dissolved gas analysis. *IEEE Trans Dielectr Electr Insul* 23:1838–1845
16. Li JZ, Zhang QG, Wang K, Wang JY (2016) Optimal dissolved gas ratios selected by genetic algorithm for power transformer fault diagnosis based on support vector machine. *IEEE Trans Dielectr Electr Insul* 23(2):1198–1206
17. Abu-Siada A, Hmood S, Islam S (2013) A new fuzzy logic approach for consistent interpretation of dissolved gas-in-oil analysis. *IEEE Trans Dielectr Electr Insul* 20:2343–2349
18. Khan SA, Eqbal MD, Islam T (2015) A comprehensive comparative study of DGA based transformer fault diagnosis using fuzzy logic and ANFIS models. *IEEE Trans Dielectr Electr Insul* 22:590–596
19. Dai CX, Liu ZG, Hu KT (2016) Fault diagnosis approach of traction transformers in high-speed railway combining kernel principal component analysis with random forest. *IET Electr Syst Transp* 6(3):202–206
20. Carita AJQ, Leita LC, Junior APPM, Godoy RB, Sauer L (2013) Bayesian networks applied to failure diagnosis in power transformer. *IEEE Lat Am Trans* 11(4):1705–1082
21. Malik H, Mishra S (2016) Application of gene expression programming (GEP) in power transformers fault diagnosis using DGA. *IEEE Trans Ind Appl* 52(6):4556–4565
22. Wang XL, Li QM, Yang R, Li CR, Zhang Y (2015) Diagnosis of solid insulation deterioration for power transformers with dissolved gas analysis-based time series correlation. *IET Sci Meas Technol* 9(4):393–399
23. Sarkar S, Sharma T, Baral A, Chatterjee B, Chakravorti D (2014) An expert system approach for transformer insulation diagnosis combining conventional diagnostic tests and PDC, RVM data. *IEEE Trans Dielectr Electr Insul* 21(2):882–891
24. Yang Z, Tang WH, Shintemirov A, and Wu QH (2009) Association rule mining-based dissolved gas analysis for fault diagnosis of power transformers. *IEEE Trans Syst Man Cybern Part C: Appl Rev* 39(6):597–610
25. Li L, Cheng Y, Xie LJ, Jiang LQ, Ma N, Lu M (2015) An integrated method of set pair analysis and association rule for fault diagnosis of power transformers. *IEEE Trans Dielectr Electr Insul* 22(4):2368–2378
26. Liao RJ, Zheng HB, Grzybowski S, Yang LJ, Zhang YY, Liao YX (2011) An integrated decision-making model for condition assessment of power transformers using fuzzy approach and evidential reasoning. *IEEE Trans Power Deliv* 26(2):1111–1118
27. Qi G, Zhu Z, Erqinhu K (2018) Fault-diagnosis for reciprocating compressors using big data and machine learning. *Simul Model Pract Theory* 80:104–127
28. Dai JJ, Song H, Sheng G, Jiang X (2017) Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network. *IEEE Trans Dielectr Electr Insul* 24(5):2828–2835
29. Li JM, Li XH, Dai DM (2020) Research on credit risk measurement of small and micro enterprises based on the integrated algorithm of improved GSO and ELM, *Math Probl Eng*. <https://doi.org/10.1155/2020/3490536>
30. Chen P, Zhao XQ, Zhu QX (2020) A novel classification method based on ICGOA-KELM for fault diagnosis of rolling bearing. *Appl Intell* 50(9):2833–2847
31. Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H (2019) Harris hawks optimization: algorithm and applications. *Futur Gener Comput Syst* 97:849–872
32. Zheng HB, Liao RJ, Grzybowski S, Yang L J (2011) Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimization. *IET Electr Power Appl* 5(9):691–696
33. Li XH, Wu H, Wu D (2011) DGA Interpretation scheme derived from case study. *IEEE Trans Power Delivery* 26:1292–1293

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Xiaohui Han** was born in Hebei, China. She graduated from China Agricultural University and obtained her Ph.D. degree in 2017. She is a lecturer with the College of Electrical Engineering, Hebei University of Science and Technology. Her research interests include fault diagnosis of transformers and power system protective relaying.

**Shifeng Ma** was born in Hebei, China. He graduated from Hebei University and obtained his B.E. degree in 2020. His research interests include condition monitoring and fault diagnosis of transformers.

**Zhewen Shi** was born in Hebei, China. He graduated from Hebei University of Science and Technology and obtained his B.E. degree in 2019. His research interests include condition monitoring and fault diagnosis of transformers.

**Guoqing An** was born in Hebei, China. He graduated from Hebei University of Technology and obtained his Ph.D. degree in 2013. He is currently an associate professor with the College of Electrical Engineering, Hebei University of Science and Technology. His main research interests include equipment condition monitoring and fault diagnosis.

**Zhenbin Du** was born in Hebei, China. He is currently a senior engineer with the Hebei provincial key laboratory of electromagnetic & structural performance of power transmission and transformation equipment. His main research interests include transformer design and simulation.

**Chunlin Zhao** was born in Hebei, China. She is currently an assistant engineer with the Hebei Weixun Electric Power Automation Equipment Co., Ltd. Her main research interests include condition monitoring and fault diagnosis of transformers.