Acta Technica *62* (2017), No. 5B, 523–530 

# Fault diagnosis model for internal combustion engines based on rvm hyperparameter optimization

Yang Bo[1,2], Shen Li-zhong[3], Han Feng-xia[1,4]

**Abstract.** To deal with the problems of low identification precision and long identification time of fault categories of internal combustion engines, we propose a fault diagnosis model based on RVM hyperparameter optimization. Firstly, an improved harmony search algorithm with stronger absorption and faster optimization-seeking speed is obtained via optimizing $HMCR$, $PAR$ and $BW$ searching approaches. Then, it is used to seek optimization of RVM hyperparameters to enhance the fault diagnosis capability. Six experimental conditions are set up: indirect misfire ($D1$), high gas-oil ratio ($D2$), low gas-oil ratio ($D3$), late ignition ($D4$), early ignition ($D5$) and normal operation ($D6$), to acquire corresponding $O_2$, CO, $CO_2$ and HC components, and taking the exhaust components and fault conditions as the input and output of RVM hyperparameters, we build the fault diagnosis model. Finally, to further verify the validity of fault diagnosis of IHS-RVM model, we make a comparison with IHS-SVM, IHS-BP and HS-BP models. Noise is added into the dataset to testify the anti-noise capability of this model. Results show that IHS-RVM model enjoys higher diagnosis precision (nearly close to 100%) and better anti-noise performance.

**Key words.** fault diagnosis; internal combustion engines;RVM; improved harmony search algorithm; exhaust component.

## 1. Introduction

A fine running of the internal combustion engine exerts impacts on the reliability and security of the power system[1]. An internal combustion engine is complex in structure and the change of its operating parameters mostly will shed the influence on the combustion into the exhaust [2~3]. In the emission process, each exhaust elements have some complex interrelation [4], which makes an obstacle to the accuracy

[1]Workshop 1 - Quality Development Institute, Kunming University of Science and Technology, Kunming Yunnan 650093, China

[2]Workshop 2 - KUNMING YUNNEI POWER CO., LTD., Kunming Yunnan 650224, China

[3]Workshop 3 - Yunnan Provincial Key Laboratory of IC Engine, Kunming University of Science and Technology, Kunming 650224, China

[4]Corresponding author: Han Feng-xia; e-mail: `hanfengxia@kmust.edu.cn`

of fault diagnosis of internal combustion engines. It is difficult for traditional multivariate statistical models like Fisher discriminant analysis (FDA) and canonical variate analysis (CVA), to achieve the satisfactory outcome. Intelligent algorithms such as artificial neural network (ANN), and support vector machine (SVM) are employed to build a non-linear model of exhaust emission and engine fault [5]. Relevance vector Machine (RVM) not only remains all advantages of SVM, but overcomes some innate shortage of SVM [6]. RVM has already been used in fault diagnosis of internal combustion engines, but its diagnostic precision needs improving due to some factors like hyperparameters and redundant information of engine faults.

In view of this above, this paper firstly makes an improvement of the acquisition of parameters *HMCR*, *PAR* and *BW* in Harmony Search (HS) algorithm, renamed as Improved Harmony Search (IHS). Then, hyperparameters of RVM are optimized based on IHS to build an IHS-RVM-based internal combustion engine fault diagnosis model. Finally, testing data of engine D09TCI -EU4 are used to verify the validity and effectiveness of IHS-RVM model.

## 2. BASIC PRINCIPLES AND METHODOLOGY

### *2.1. HS Algorithm and its Improvement*

HS algorithm is proposed by Geem, etc., HS algorithm has a simple principle with a few controllable parameters, easy to realize overall search. But in this algorithm, values of parameters *HMCR*, *PAR* and *BW* are fixed, which, to some extent, affects the search performance. The parameter optimization method[7-9] in this paper is as follows. Detailed improvement is as the following equation (1), (2), (3), (4).

$$HMCR = HMCR_{\max} - \frac{HMCR_{\max} - HMCR_{\min}}{NI_{max}} NI_d \tag{1}$$

$$PAR = PAR_{\min} + (PAR_{\max} - PAR_{\min}) \operatorname{div} \frac{\pi}{2} \times \arctan(IN_d) \tag{2}$$

$$BW = BW_{\max} - \frac{BW_{\max} - BW_{\min}}{T_{\max}} T_d \tag{3}$$

$$x^{(j)} = x^{\text{best}} \oplus x^{(j)} \tag{4}$$

where $NI_d$is current iteration number; $HMCR_{\max}$and $HMCR_{\min}$ are the maximum and minimum value of search probability respectively; $PAR_{\max}$ and $PAR_{\min}$ are the maximum and minimum value of the improvement; $BW_{\max}$ and $BW_{\min}$ are the maximum and minimum value of tone adjustment bandwidth; $x^{best}$ is the optimal harmony among HM; $\oplus$ is the symbol of XOR operation.

### 2.2. RVM Algorithm

Assume the given internal combustion engine fault sample training set is $\{\mathbf{x}_n, \mathbf{t}_n\}_{n=1}^{N}$. The relation between input vector of the said fault sample training set $\{\mathbf{x}_n\}_{n=1}^{N}$ and the targeted value $t_n$ can be given by

$$t_n = y\left(\mathbf{x}_n; \omega\right) + \varepsilon_n \tag{5}$$

Bayes probability model (eq. 6) is introduced to explain the effect of noise $\varepsilon_n$ on RVM fault classifier.

$$p\left(t_*\middle| t\right) = \int p\left(t_*\middle| w, \sigma^2\right) p\left(w, \sigma^2\middle| t\right) dw d\sigma^2 \tag{6}$$

Since the integrand meets Gaussian distribution, Laplace's method is used to approximate $p\left(w\middle| t, \partial_{\mathrm{MP}}, \sigma_{\mathrm{MP}}^2\right)$ to Gaussian distribution, given by

$$p\left(t_n = 1\middle| \omega^{\mathrm{T}}\phi\left(x_n\right)\right) = \left(1 + \exp\left(-\omega^{\mathrm{T}}\phi\left(x_n\right)\right)\right)^{-1} \tag{7}$$

The predicted probability of the fault classification outcome is given by

$$p\left(t\middle| w\right) = \prod_{i=1}^{N} \sigma\left[y\left(\mathbf{x}_n; \omega\right)\right]^{t_i} \left\{1 - \sigma\left[y\left(\mathbf{x}_n; \omega\right)\right]\right\}^{1-t_i} \tag{8}$$

In equation (5)~(8), $K\left(\mathbf{x}, \mathbf{x}_n\right)$ is the kernel function selected by fault classifier of internal combustion engines; $\{\omega_n\}_{n=0}^{N}$ refers to the weight of the model; $\varepsilon_n$ falls into the Gaussian distribution of N(0, $\sigma^2$) and it is the noise; $t_*$ is the predicted value; A few weights $\omega_i$ are stable and near some finite value, and the corresponding sample vector $\mathbf{x}_i$ is the relevant vector.

## 3. IHS-RVM FAULT DIAGNOSIS MODEL

### 3.1. IHS-based RVM Hyperparameters Optimization

This paper employs IHS algorithm to acquire the optimal hyperparameter combination of RVM diagnosis model, we use $V_{acc}$as the targeted function, which is given by

$$f\left(x\right) = V_{acc} = \frac{V'_a}{V_a} \tag{9}$$

where $V_a'$ is the number of correct fault classification; $V_a$ is the number of fault sample.

Detailed RVM hyperparameters optimization procedures are as follows. **Step1:** set the initial range of *HMCR, HMS,PAR, BW*, $NI_{max}$, $r$, $y$ *and* $N$ (harmonic vector dimension). **Step2:** there are *HMS* harmony generated randomly from the solution domain. Make the classification accuracy $V_{acc}$ the fitness value of each harmony in the harmony base. **Step3:** choose randomly a new harmony from HM. **Step4:** if the number of iteration reaches $NI_{max}$, the algorithm is ended, then output the optimal hyperparameter combination $(r, y)$ of RVM fault diagnosis model. **Step5:** after obtaining the optimal hyperparameter combination, output the vector of optimal solution of the hyperparameter and construct the RVM diagnosis model of hyperparameter optimization.

### *3.2. IHS-RVM Fault Diagnosis Procedures*

IHS-RVM fault diagnosis procedures are as follows. **Step1:** divide the normalized data set into a training set and a testing set. The front 2/3 data is used in IHS-RVM training set of fault model; the rest 1/3 is taken into the testing set. Use HC, CO, $CO_2$and $O_2$as the input of classifier and fault type the output. **Step2:** as to fault diagnosis training set of internal combustion engines, Gaussian kernel function is adopted as the kernel function of RVM fault diagnosis model. **Step3:** in accordance with the optimizing procedures. **Step4:** after the construction of IHS-RVM model with hyperparameter optimization, model test is carried out in line with the testing set.

## 4. CASE ANALYSIS OF FAULT DIAGNOSIS

### *4.1. Exhaust testing data collection*

Misfire fault experiment is carried out in Kunming Yunnei Power Co., Ltd. In reference to the paper [5], we set 6 states of engine: indirect misfire ($D1$), high gas-oil ratio($D2$), low gas-oil ratio ($D3$), late ignition ($D4$), early ignition ($D5$) and normal operation ($D6$).Exhaust analyzer (MEXA-584L) is used to obtain the exhaust emission data under the 6 conditions with a rotational speed of 4000r/min. Each test is for one state and obtain 15 pieces of data. Initial data of gas content that the sensor can test is the net weight of each gas. Normalize the initial data, make up the sample data as shown in Table 1.

Table 1.Gas composition and content (part)

| Group | $O_2$/% | CO/% | $CO_2$/% | HC/(mg•$L^{-1}$) | Fault diagnosis |
|---|---|---|---|---|---|
| 1 | 0.0384 | 0.1089 | 0.8196 | 0.4371 | $D5$ |
| 2 | 0.0289 | 0.0793 | 0.6578 | 0.2324 | $D2$ |
| ··· | ······ | ······ | ······ | ······ | ······ |
| 90 | 0.0519 | 0.9999 | 0.6104 | 0.9998 | $D3$ |

### *4.2. Fault Diagnosis Result Analysis*

In order to fully assess the diagnosis performance of the model, we make comparison among models of HS-BP, HS-SVM, IHS-SVM and IHS-RVM and take diagnosis time and classification accuracy as the evaluation indicators. Fault results are shown in Fig. 1.

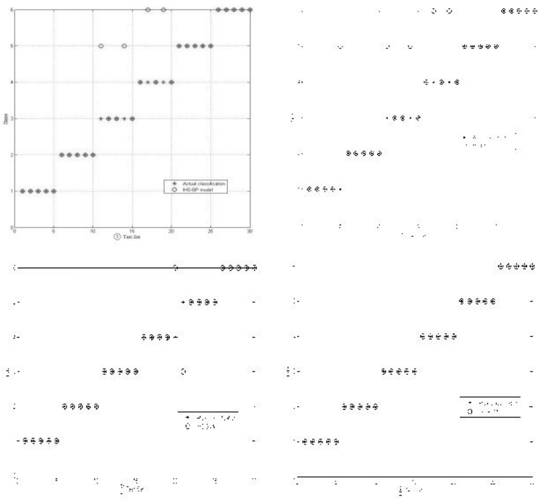

Fig. 1. Fault diagnosis outcome of each model

It can be learned from Fig. 1 that IHS-RVM enjoys good searching and its classification accuracy can arrive at 100%, which proves the validity of this model. Diagnosis results of other models are presented in table 2.

Table 2. Comparison of internal combustion engine fault diagnosis model

| Diagnosis model | HS-BP | IHS-BP | IHS-SVM | IHS- |
|---|---|---|---|---|
| Diagnosis time /s | 8.8762E-4 | 6.3178E-4 | 2.2125E-4 | 1.901 |
| Classification accuracy /% | 83.33 | 86.67 | 93.33 | 10 |

From table 2, we can learn that IHS-RVM model enjoys the best diagnosis performance with diagnosis time of 1.9012E-4 and classification accuracy of 100%. IHS-RVM is better than IHS-SVM fault diagnosis model, which is resulted from that RVM adopts probability Bayesian framework and overcomes the shortage of SVM like Mercer condition limitation and thereby it has a better classification performance[10, 11].

### *4.3. Data Noise Disturbance Analysis*

When the rotational speed is higher, noise data tested by the sensor will intensify. In the fault diagnosis model in existing literature, most exhaust data are obtained from the fixed rotational speed in the lab, with little consideration of noise's disturbance on the diagnosis accuracy[12]. This experiment is a whole vehicle test. Testing data are acquired at the static stall. We pay little attention to the effect of speed change on fault diagnosis accuracy. Thus, data noise is added to testify the applicability of IHS-RVM fault diagnosis model. Diagnosis results of each model are listed in table 3.

Table 3 Effect of noise on the mode

| Noise intensity | Standard | 0.05 | 0.01 | 0.15 | 0.20 | Noise intensity | Standard | 0.05 | 0.01 | 0.15 | 0.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| HS-BP | 5 | 6 | 7 | 10 | 13 | IHS-SVM | 2 | 3 | 5 | 8 | 11 |
| IHS-BP | 4 | 5 | 7 | 9 | 12 | IHS-RVM | 0 | 1 | 2 | 4 | 6 |

Note: Numbers listed in the table refers to the number of fault classification mistakes of the testing samples.

We can learn from table 3 that when noise intensity is changing, IHS-RVM fault classification accuracy is better in performance than HS-BP, IHS-BP and IHS-RVM models. In our future study, experiments are redesigned to obtain exhaust data samples with different vehicle speed, and build RVM fault diagnosis model of hyperparameters optimization. In this way, a dynamic monitoring of internal combustion engine with different vehicle buffer can be realized and the real-time operation status of the engine is displayed.

## 5. CONCLUSION

Aase study has proved the diagnosis performance of IHS-RVM model, and made an analysis of diagnosis accuracy, diagnosis time and anti-noise performance. Conclusions are as follows. (1) The IHS enjoys a favorable performance of parameter optimization. With the same sample amount, IHS-BP model is of shorter time than HS-BP model in diagnosis, with a difference of 2.5584 E-4 in time. (2) RVM model after hyperparameter optimization is of higher prediction precision. It decreases the effect of hyperparameter on diagnosis accuracy by using IHS to optimize RVM hyperparameters. The diagnosis accuracy can reach 100.00% under the non-noise circumstance. When the intensity of noise is 0.2, error rate is 6.67%, which indicates this model has an excellent diagnosis accuracy with a stable performance. (3) This model has a bright market prospect. Using exhaust data to make a fault diagnosis will considerably decrease the effect of vehicle operation situation on the performance of fault diagnosis models, and exhaust data is easily accessible and cost little.

To some extent, there is some weakness of the real-time diagnosis in this research. New experimental design will be made to obtain exhaust data with different vehicle speed and construct the corresponding RVM models with hyperparameter optimization, i.e., a dynamic monitor of real-time change of vehicles can be realized and real-time fault diagnosis and recognition will be under control.

**References**

[1] J. D. Wu, C. H. Liu: *Investigation of engine fault diagnosis using discrete wavelet transform and neural network.* Expert Systems with Applications *35* (2008) No. 3, 1200-1213.

[2] X. Y. Xu, W. P. Yang, X. Lv, X. H. Ma: *Fault diagnosis for a car engine based on support vector machine.* Journal of Vibration & Shock *32* (2013) No. 6, 143-146.

[3] A. Widodo, E. Y. Kim, J D. Son, B. S. Yang, A. C. C. Tan, D. S. Gu, B. K. Chio, J. Mathew: *Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine.* Expert Systems with Applications *36* (2009) No. 3,7252-7261.

[4] A. Widodo, B. S. Yang: *Support vector machine in machine condition monitoring and fault diagnosis.* Mechanical Systems & Signal Processing *21*(2008) No. 6,2560-2574.

[5] C. M. Vong, P. K. Wong, W. F. Ip: *A new framework of simultaneous-fault diagnosis using pairwise probabilistic multi-label classification for time-dependent patterns.* IEEE transactions on industrial electronics *60* (2013)No. 8, 3372-3385.

[6] M. E. Tipping: *Sparse Bayesian Learning and the Relevance Vector Machine.*Journal of Machine Learning Research *1*(2001) No. 3, 211-244.

[7] M. Mahdavi, M. Fesanghary, E. Damangir: *An improved harmony search algorithm for solving optimization problems.* Applied Mathematics & Computation *188* (2007)No. 2, 1567-1579.

[8] V. T. Tran, B. S. Yang, F. Gu, A. Ball: *Thermal image enhancement using bi-dimensional empirical mode decomposition in combination with relevance vector machine for rotating machinery fault diagnosis.* Mechanical Systems and Signal Processing *38* (2013) No. 2, 601-614.

[9] R. Diao, Q. Shen: *Feature selection with harmony search.* IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) *42* (2012) No. 6, 1509-1523.

[10] C. Zhang, Z. Ni, L. Ni, N. Tang: *Feature selection method based on multi-fractal dimension and harmony search algorithm and its application.*International Journal of Systems Science *47*, (2016) No. 14, 3476-3486.

[11] I. Psorakis, T. Damoulas, M. Girolami: *Multiclass relevance vector machines: sparsity and accuracy.* IEEE Trans Neural Netw *21*(2010) No. 10, 1588-1598.

[12] Y. Shatnawi, M. Al-Khassaweneh: *Fault diagnosis in internal combustion engines using extension neural network.*IEEE Transactions on Industrial Electronics*61* (2014), No. 3, 1434-1443.

Received November 16, 2017