Diesel Fault Prediction Based on Optimized Gray Neural Network

Han Bing, Yin Zhenjie

Abstract—In order to analyze the status of a diesel engine, as well as conduct fault prediction, a new prediction model based on a gray system is proposed in this paper, which takes advantage of the neural network and the genetic algorithm. The proposed GBPGA prediction model builds on the GM (1.5) model and uses a neural network, which is optimized by a genetic algorithm to construct the error compensator. We verify our proposed model on the diesel faulty simulation data and the experimental results show that GBPGA has the potential to employ fault prediction on diesel.

Keywords—Fault prediction, Neural network, GM (1.5), Genetic algorithm, GBPGA.

I. INTRODUCTION

AS the core equipment of ships, the diesel engine plays an important role in the safe operation of ships. The research of status analysis and fault prediction on diesel is gaining more and more attention. The common methods [1] of state analysis include mainly domain value analysis, time series analysis and frequency domain analysis, etc. In recent years, many researchers have focused on the establishment of a fault prediction model with high accuracy to forecast various faults. A certain prediction model [2], including a time series model, gray model and a neural network model has been tested and applied into actual fields.

Gray prediction establishes a gray model through the processing of raw data, then finds and masters the law of system development. After that, gray prediction anticipates the future state of a system [3]. The gray prediction model is tested and applied into the fault status analysis of a diesel engine on paper [4], which offers a new insight for research on fault prediction technology. However, the weak capability of nonlinear approximation of gray model makes the prediction effect unsatisfactory.

The neural network has a strong capability of nonlinear approximation with high reconstructing accuracy and a fast training rate, which is manages to make up for the deficiency of the gray model. Many efforts have been made in this area. X. B. Ye [5] created the gray neural network prediction model to analyze aircraft fault; the combination prediction model using a neural network and optimized gray model was proposed by X. Y. Lee [6] to compare the difference between combination models and single ones.

Though applying the neural network into the gray prediction model has resulted in a certain achievement, it

Han Bing and Yin Zhenjie are with the State Key Laboratory of Navigation and Safety Technology, Shanghai Ship and Shipping Research Institute, P.R. China (e-mail: hanbing@sssri.com).

tends to fall into local optimum when training the neural network, due to the limitation of the gradient algorithm. The local optima make the model unstable and different results might be obtained on the same input data. Therefore, introducing an optimization mechanism into the network to avoid such problems is especially necessary.

The optimization mechanism is inspired from the laws of nature and imitates the ways of natural law to solve problems in research [7]. It mainly contains genetic algorithm, a simulated annealing algorithm and swarm intelligence, etc. The genetic algorithm is widely used in lots of research for its advantages of global superiority and concurrency [8], which manages to solve the problem of local optimum originating from the neural network. Therefore, considering the characteristic of the diesel data, which are nonlinear, multiparameter, and of very few-labeled sample, the gray model is employed in this paper as it can achieve high classification accuracy when dealing with a few-labeled sample. A genetic algorithm, with global superiority and concurrency, is introduced to optimize the neural network. Then we use the optimized neural network as the error compensator of the gray model. The combination prediction model GBPGA is developed in this paper. We also test our model on the fault prediction of the diesel engine.

II. THE COMBINATION PREDICTION MODEL

A. Prediction Model of GM (1.5)

The grey prediction model GM uses accumulated raw data to establish the prediction model of the gray differential equation. It considers the raw data as a gray process which changes over time, then gradually turns whitened through the accumulation or subtraction generation. Finally, it builds corresponding differential equations model to predict [9].

GM (1.1) is usually applied to build the prediction model. The first 1 of which stands for the first-order equations, and the other one explains gray dimensions. Considering the multiparameter characteristic of fault on diesel engine, multidimensional model GM (1.5) is created in this paper.

With the development of the prediction system, the underlying information of raw data weakens gradually. At the same time, new information is added constantly. The present system can be reflected more accurately with the timely exclusion of outdated information. In addition, the computational burden is also decreased with the deletion of the old raw data. Therefore, the metabolism of the GM model in this paper is built with a high prediction precision, namely that the raw data keeps updating during the modeling.

The expression of the discrete gray differential equation of

the prediction model is as follows:

Assuming the raw dataset:

$$X = (x_i(1), x_i(2),...,x_i(m)), \quad i = 0,1,...,n \qquad (1)$$

Accumulate the sequence to obtain the monotone increasing one. The new sequence is more prone to meeting the statistical law, which is expressed as:

$$X^{(1)} = (x_i^1(1), x_i^1(2),...,x_i^1(n)) \qquad (2)$$

$$x_i^1(k) = \sum_j^k x_i(j) \qquad (3)$$

Create the gray differential equation with the monotone increasing sequence as follows:

$$\frac{dX^{(1)}}{dt} = aX^{(1)} + u \qquad (4)$$

The value of parameter matrix a and gray matrix u, depends on the raw data and the structure form as:

$$a_i = (L_i^T L_i)^{-1} L_i^T Y_i, \quad a_i = (a_{i1}, a_{i2}, \cdots, a_{in}, u_i)^T \qquad (5)$$

$$L_i = \begin{pmatrix} 0.5(x_i^1(2) + x_i^1(1)) & 1 \\ 0.5(x_i^1(3) + x_i^1(2)) & 1 \\ \cdots & \cdots \\ 0.5(x_i^1(m) + x_i^1(m-1)) & 1 \end{pmatrix} \qquad (6)$$

$$Y_i = (x_i^1(2), x_i^1(3), \cdots, x_i^1(m))^T \qquad (7)$$

According to gray systematic theory, the discrete equation of prediction [10] is expressed as:

$$\hat{x}^{(1)}(k) = (x^{(1)}(1) - \frac{u}{a})e^{-a(k-1)} + \frac{u}{a} \qquad (8)$$
$$k = 1,2,\cdots$$

B. Genetic Algorithm Optimization

The genetic algorithm is introduced to optimize the neural network due to its advantage of global superiority and concurrency.

The optimization procedure [11] is as follows:

1. Pop-Initialization

Individual S is randomly generated, in order to ensure the genetic algorithm model is superior to the traditional one.

2 Computing the Fitness Function

In the genetic algorithm, the size of the individual fitness determines the probability of individuals that is inherited in the next generation. The fitness function is the only information to guide the search, so its goodness is the key to success for the algorithm. In order to minimize the residual

error σ, the fitness function is explained as:

$$f(\lambda) = \frac{1}{1 + \sum \sigma} = \frac{1}{1 + \sum_{k=1}^{n} \left| \hat{x}^{(0)}(k) - x^{(0)}(k) \right|} \qquad (9)$$

The search will continue until accuracy is achieved.

3. Selection Operation

Choose two paternal parents using gambling wheel selection according to the probability p_i. Two daughters are produced when cross them. The selection probability of each chromosome is p_i.

$$p_i = f_i / \sum f_i \qquad (10)$$

4. Crossover Operation

The search capability will be improved after the crossing operation. According to the crossover probability, two individuals were randomly selected and randomly divided into two groups. The process of search will stall if the probability is too small, so set it between 0.5 and 1.0. The probability is 0.7 in this paper.

5. Mutation Operation

The gene mutation operator changes the gene mutation probability through determining all the individuals randomly, whose purpose is to enhance the local search ability and increase the diversity of the population. In the parameters of mutation operation, mutation probability influences the behavior and performance of the genetic algorithm directly. Usually the mutation probability is between 0.01 and 0.2, and it takes 0.15 in this paper.

6. Termination

Iteration loops until the accuracy or maximum algebra is reached.

C. The Neural Network Error Compensator

The neural network error compensation is adopted to identify the mapping relationship between the raw data and the prediction error, and then to correct the prediction value. The error compensation process, based on the neural network, can be done by training the neural network on a certain number of training samples to get the expected network for calculating the errors.

The three-layer BP neural network is adopted to establish the error compensation model as it can approximate nonlinear function with arbitrary precision [12]. The node numbers of the input layer, hidden layer and output layer depend on the object and its complexity. For the training progress, the prediction and error value of the grey prediction model is taken as samples of the neural network compensator. According to the characteristics of the simulation data of diesel fault, a $5 \times 6 \times 5$ network is adopted, namely the input layer and output layer are both set up to five neurons, and the hidden layer six neurons. The connection weight matrix

between the input layer and the hidden layer is expressed as $W = (w_{ji})_{5 \times 6}$, and the one between the hidden layer and the output layer is $W = (w_{ji})_{5 \times 6}$.

The training error function of the error compensator can be minimized by the genetic algorithm mentioned in Section II A. Then the error compensator could be applied to the gray predictive model. After a preliminary prediction is made with the prediction model GM (1.5), the final prediction value can

be achieved with the neural network error compensation.

III. FAULT MODEL AND SAMPLE DATA

A fault model of TBD 620 diesel is established through the AVL boost software in this paper. The AVL software is developed by AVL List GmbH for the simulation of engine performance, which can be applied for simulation analysis of an engine fault. The stimulation model [13] is shown in Fig. 1.
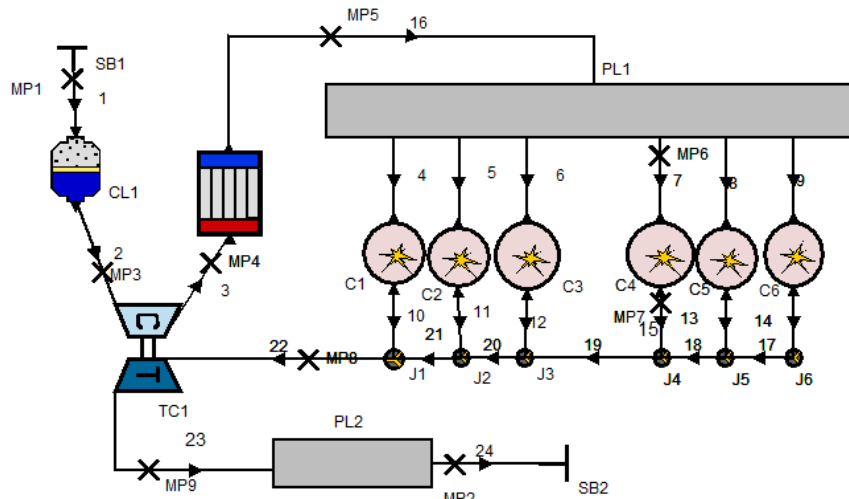

Fig. 1 TBD 620 diesel numerical calculation model

The simulation model mainly contains six cylinders (C1-C6), an air cooler (CO1), a turbocharger (TC1), two systematic boundary (SB1, SB2), an air inlet pipe (PL1), an exhaust pipe (PL2), an air filter (CL1), nine measure points (MP1-MP9) and six connecting pipes (J1-J6). The combustion mode is expressed by a single VIBE function, and the heat transfer is explained by Woschni 1978. The supercharged model chooses the simplified one [14].

TABLE I
REFERENCE FAULT VECTOR TABLE

	Normal operating	Fault 1	Fault 2	Fault 3	Fault 4	Fault 5
P_e	18.57	-1.06	15.69	14.3	17.86	18.03
g_e	186.52	0	209.27	221.62	194.58	198.25
p_z	16.12	13.32	15.13	12.66	14.07	15.38
λ_{max}	467912	479813	433581	398471	366155	412782
T_{tout}	478.12	451.82	705.24	483.63	473.03	489.43

TABLE II
FAILURE MODES LIST

Fault No.	Fault types
Fault 1	Single cylinder misfire
Fault 2	Turbocharger efficiency failure
Fault 3	Gas blowby
Fault 4	Compression ratio decrease
Fault 5	Efficiency failure of air cooler

After the calibration of the simulation model, various load testings, including 100 percent, 75 percent, 50 percent and 25

percent are carried out. And then, the work and failure state containing normal operation, single cylinder misfire, turbocharger efficiency failure, gas channeling, compression ratio decreases and air-cooling efficiency difficulties will be simulated separately. Maintaining the external conditions invariant, the characteristic parameter values of various fault conditions are measured and recorded, respectively. Moreover, the serial numbers of all kinds of faults are recorded to establish the reference vector library and the raw data of the prediction system.

TABLE III
INITIAL DATA FROM FAULT 5

sequence	P_e	g_e	p_z	λ_{max}	T_{tout}
1	17.81	203.24	15.03	396943	495.54
2	17.89	202.09	15.18	399116	493.32
3	17.93	202.26	15.21	401354	492.13
4	17.96	200.97	15.31	404985	490.63
5	18.09	198.72	15.26	409672	489.71
6	18.03	197.83	15.33	412257	489.84
7	18.12	196.48	15.38	417145	487.24
8	18.25	196.64	15.45	423206	486.56

In this paper, five thermal parameters are taken as the monitoring parameters, including mean effective pressure P_e (bar), effective fuel consumption rate g_e (g/Kw×h), peak pressure P_z (MPa), maximum pressure rise ratio λ_{max} and exhaust gas temperature behind turbine T_{out} (K)[15]. According

to status and trend analysis through the above five thermodynamic parameters, the monitoring and diagnosis of the working process of cylinders can be realized.

The reference fault data is shown in Table I, whose corresponding failure modes are explained in Table II. The raw data from the simulation model of air-cooling efficiency difficulties are chosen and expressed in Table III.

IV. SIMULATION RESULT

In this section, the prediction model raised in this paper is calculated by MATLAB 9.0a with the sample data in Table III. Two simulation experiments are used to verify the proposed method. In order to verify the accuracy of the fault prediction model, firstly three methods were used for comparison. On this basis, the diesel fault perdition is made by the gray relational degree between the prediction value and the reference data.

A. Experiment Results

GM (1.5) is used to make a preliminary prediction with the former seven sets of raw data, so the crude value and error can be obtained. The genetic algorithm is employed to optimize the network to gain the best connection weight and threshold value. The former seven sets of raw data are taken as inputs, and the corresponding error as expected output of neural network. Then the optimized BP neural network is introduced for error compensator building. The sum of the preliminary prediction and error compensation value are treated as the final prediction value.

The optimization course and training iteration result are shown in Figs. 2 and 3, separately. According to the figures, genetic algorithm gains the best fitness 0.0186, and the training error of network is achieving 9.97919e-005 when the network is converged after 1569 iterations.

The computation results are achieved, and the prediction compensator error vector is [0.24 3.01 0.28 8710 8.03]. Table IV shows the comparison of the final prediction value and relative error.

Table IV shows that the combination prediction model GBPGA attains more accurate result than the others. The error decreases noticeably, and the prediction result is more closely to the actual one. Both of the optimization of the genetic algorithm and the error compensator of the neural network

improve the accuracy of the prediction model. Therefore, the combination model is effective and more accurate in diesel fault prediction.
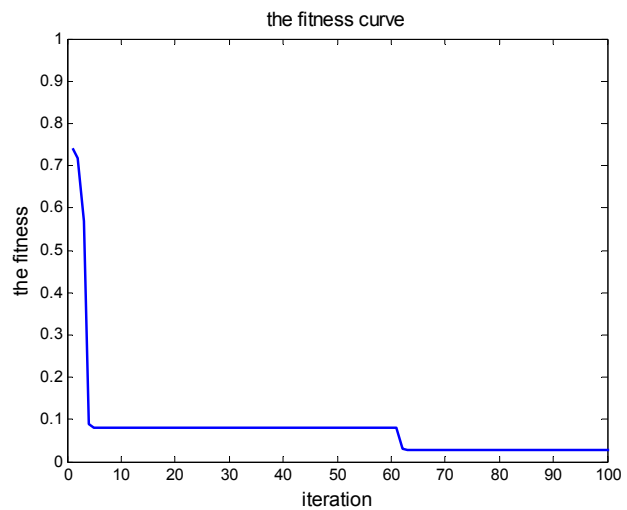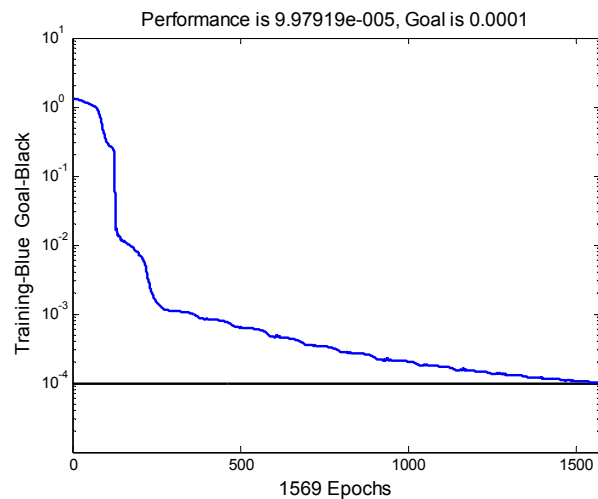

Fig. 2 Best fitness

Fig. 3 Training result

TABLE IV
THE PREDICTION VALUE AND RELATIVE ERROR

	Raw data	Prediction value GM	error（%）GM	Prediction value NN	error（%）NN	Prediction value GBPGA	error（%）GBPGA
P_e	18.25	17.61	3.52	17.77	2.61	18.01	1.34
g_e	196.64	191.11	2.81	191.70	2.51	193.63	1.53
p_z	15.45	14.79	4.26	14.96	3.14	15.17	1.81
λ_{max}	423206	406786	3.88	408944	3.37	414488	2.06
T_{tout}	486.56	470.80	3.24	472.64	2.86	478.53	1.65

B. Fault Prediction

According to the gray correlative analysis [16], the gray relational degree between the tested data and reference fault data can be calculated. The gray relational matrix can then be

obtained, whose rows stand for fault types, and the columns explain the tested data. In accordance with the judging criteria of the maximum gray relational degree [17], the maximum value in each row of the matrix is corresponding with the fault

type of the tested data.

Calculate the gray relational degree of the prediction value with the reference fault data, and the result is expressed by the gray matrix as:

$$M_1 = \begin{bmatrix} 0.8273 & 0.6779 & 0.7654 & 0.6797 & 0.8077 & \underline{0.9319} \end{bmatrix}$$

The first column of the gray matrix M_i is expressed normal status, and the next columns are corresponding to the status, Fault 1 to Fault 5, respectively.

The fault model corresponding to the maximum of the row in the matrix is the diagnosis result (The bottom line has been marked). The maximum value appears in the last column, which represents a failure of the air cooler. It shows the same fault diagnosis result as the raw data description. Therefore, the results give a persuasive demonstration for the combination prediction model, which makes an effective and accurate prediction on the fault of the diesel engine.

Another raw data set is taken to analyze in the same way as above, which is produced from the normal operation state model. The raw data of the normal operation states is shown in Table V. The final prediction value and relative error by the combination prediction model are obtained, and the results are also expressed by the gray matrix M_2.

TABLE V
INITIAL DATA FROM NORMAL OPERATION

sequence	P_e	g_e	p_z	λ_{max}	T_{tout}
1	18.36	193.53	15.71	465168	481.23
2	18.43	192.34	15.76	466835	481.15
3	18.49	191.15	15.82	467438	480.93
4	18.58	189.98	15.94	468992	480.62
5	18.51	188.82	16.03	468815	479.10
6	18.62	189.47	16.11	469135	478.23
7	18.76	187.61	16.15	470305	479.26
8	18.83	185.73	16.24	470916	476.52

$$M_2 = \begin{bmatrix} \underline{0.9802} & 0.7806 & 0.7451 & 0.6310 & 0.7143 & 0.7944 \end{bmatrix}$$

The gray matrix shows that the prediction status through the combination prediction model is a normal operational state, which is the same as the actual simulation model. The combination model makes an effective and accurate prediction on the fault status of the diesel engine. Both of the above examples strongly demonstrate that combination prediction can be effectively and accurately used in the condition analysis and fault prediction for a diesel engine.

V. CONCLUSION

An error compensator is established by a neural network, and genetic algorithm is introduced to optimize the neural network for higher accuracy for the prediction model of the gray system, which comprises of the combination prediction model GBPGA. Simulation experiments show that the combination prediction model is accurate and reliable, and can be effectively used in condition analysis and fault prediction for diesel engines.

REFERENCES

[1] Hountalas. Dimitrios. T. Prediction of Marine Diesel Engine Performance under Fault Conditions. Applied Thermal Engineering, 2000, 20(18): 1753-1783P.
[2] Mustafa Canakci, Ahmet Necati Ozsezen, Erol Arcakioglu, et al. Prediction of Performance and Exhaust Emissions of Diesel Engine Fueled with Bio-diesel Produced from Wase Frying Palm Oil(J). Expert Systems Applications, 2009, 36(5): 9268-9280.
[3] S. F. Lu and N. M. Xie, *Grey System Theory and Application* (M). Beijing: Science Press, 2008, pp. 48-50.
[4] Z. W. Wang, T. Yu at al. The Application of GM (1.1) Model for Performance Prediction of Marine Diesel Engine(C). International Conference on Maritime Technology. 2012.
[5] X. B. Ye. Fault Forecast and Diagnosis of Sequential Turbocharging System Based on Gray Theory and Neural Network (J). Transactions of CSICE, 2008, 26(6): 543-549.
[6] X. Y. Li , "Optimized Grey RBF Prediction Model Based on Genetic Algorithm." CSSE2008, pp. 74-77.
[7] M. Zhou, and S. D. Sun, *The Principle and Application of Genetic Algorithm*. Beijing: National Defence Industry Press, 1999, pp.15-22.
[8] J. L. Yuan, L. Zhong. The Dynamic Gray Radial Basis Function Prediction Model and Its Application(C). Proceeding of IEEE, ICICIC06, 2006: 582-585.
[9] Z. J. Yi and X. B. Ye, "Fault Prediction Model Based on Multiparameter Grey Error Neural Network." Application Research of Computers, 2012.
[10] G. Y. Li, "Gray Neural Network Algorithm Improved by Genetic Algorithm." Control Engineering, 2013.
[11] Z. Chen. Optimization Study on Ultra-short Term Wind Speed Forecasting of Wind Farms Based on BP Neural Network and Genetic Algorithm(J). Renewable Energy Resources, 2012, 30(2); 32-36.
[12] Benkaci M, Hoblos G, Langlois N, et al. Sensor Fault Detection and Isolation in Diesel Air Path Using Fuzzy-ARTMAP Neural Network. IEEE. ICNSC(C). Etienne DU, France, 2013, 728-733.
[13] Maass B, Stobart R. Control Oriented Real Time Model of Marine Power-station Diesel Engine Based on Neural Network (J). Computer Research and Development, 2011, 17(4): 386-390.
[14] Ghaffarpour, M, Esfahanian, V, Javaheri. A. Thermal Analysis of An SI Engine Piston Using Different Combustion Boundary Condition Treatments. Applied Thermal Engineering, 2006: 277-287P.
[15] Maja Krcum, Gojmir Radica, Zdeslav Juric. Application of The Expert Diagnostic System for Marine Diesel Engine Condition (C). The IASTED International Conference on Database and Applications. P169-173.
[16] Yuan Jingling, Zhong luo, Li X. Y. Intelligent Simulation and Prediction Methods of Structure State, 2008 Chinese Control and Decision Conference(2008CCDC), 1253-1255.
[17] W. Q. Zhao, D. X. Niu. A Mid-long Term Load Forecasting Model Based on Improved Gray Theory. 2010 2nd IEEE International Conference on Information Management and Engineering, 2010, (1):633-635.