

Data-Driven Fault Prognosis for Pneumatic Valves in Train Electropneumatic Brake System



Dianzhu Gao , Jun Peng , Ning Ding , and Yingze Yang 

Abstract Pneumatic valves are key components of the train electro-pneumatic braking system. In order to obtain health indicators of pneumatic valves and provide faults early-warning, this paper proposes a fault prognosis method using principal component analysis (PCA) and support vector regression (SVR). Two health indicators (T^2 and SPE) of pneumatic valves are extracted through PCA method based on the full life cycle data set, which came from the joint simulation model. Second, a pneumatic valve fault prognosis model based on SVR is trained based on the health indicators. Combined with the working model of the train electro-pneumatic braking system, the proposed fault prognosis model can estimate the expected time of pneumatic valve fault time accurately. Results from a semi-physical simulation verification platform of DK-2 braking system indicate that the proposed method can effectively predict the occurrence of faults. This work can provide a scientific basis for the operation of braking system and maintenance strategy of pneumatic valves.

Keywords Pneumatic valve · Principal component analysis · Support vector regression · Fault prognosis

1 Introduction

As the key link of circuit-pneumatic-mechanical control loop in the train braking system, pneumatic valve plays an important role in system's safety and reliability [1]. Pneumatic valves gradually deteriorate with the accumulation of working time, thus leading to a decline in performance [2]. Due to the complex structure, the internal conditions of pneumatic valves cannot be directly observed, resulting in difficulty in predicting when the faults occur.

D. Gao · N. Ding

School of Automation, Central South University, Changsha 410075, China

J. Peng · Y. Yang (✉)

School of Computer Science and Engineering, Central South University, Changsha 410075, China

e-mail: yangyingze@csu.edu.cn

Current fault prediction methods consist of physical model-based methods, statistical model-based methods and machine learning-based methods [3]. The physical model-based methods use the thresholds generated by the actual and model values as indicators for prediction [4]. Chetan et al. [5] used a mathematical model describing the physical principles of component degradation, a Bayesian filter algorithm was used for parameter state estimation and fault prediction. The physical model-based methods require accurate mathematical models, but it is difficult to build an accurate mathematical model for pneumatic valves [6].

Instead of building a physical model, the statistical methods use historical data to establish health indicators of the degradation process [7]. Jin et al. [8] derived an auto-regressive model for filtering fault-independent signals to track the degradation process of the bearing and designed an extended Kalman filter for fault prediction. Principal component analysis is widely used in the field of fault prediction where the degradation state can be accurately measured according to the degree of deviation of the statistical indicators [9].

In recent years, machine learning-based methods for fault prediction have drawn a lot of attention, since only enough historical fault data is required to implement complex fault prediction problems [10]. For example, long and short-term memory recurrent neural networks are used to perform multi-forward voltage prediction for battery system failures prediction [11]. Hack-Eun et al. [12] introduced a support vector machine classifier to estimate the health state and to make long-term predictions of the bearing degradation state.

The purpose of this paper is to predict the degradation process in pneumatic valves and to provide early warning before a fault occurs. A support vector regression-based fault prediction method for pneumatic valves is proposed, and different fault prediction models are established for different types of faults. Firstly, the similarities in the failure characteristics of different pneumatic valves are discussed, and the features are reconstructed according to the fault types. Then a principal component analysis method is used to extract the health indicators. Finally, the support vector regression model is trained separately for each fault type to predict the time of fault occurrence.

The remainder of this paper is organized as follows. The degradation and simulation models of pneumatic valves are established in Section II. Section III describes the SVR-based fault prognosis method. Simulation results are presented in Section IV. We conclude the paper in Section V.

2 Modeling

In this section, the structure and working principle of the pneumatic valve are analyzed. The simulation models of the train/equalization control unit of the electro-air braking system based on AMESim and Matlab are built to analyze the different fault types and performance, and obtain the fault data to verify the proposed method.

Pneumatic valve is a mechanical valve, which does not rely on electromagnetic drive. In the braking system, pneumatic valves mainly include relay valves, and other valves. Relay valve, as one of the core parts of braking system, is very representative in different types of pneumatic valves. Therefore, this paper takes the relay valve as the research object to verify the effectiveness of the method. The brake control unit is built with Simulink toolbox in MATLAB.

The simulation model of train/equalization control unit is realized by AMESim and Simulink toolbox in MATLAB. The model includes two parts: air channel and brake control unit. The air channel is simulated by the Pnuematic Kit and the Pnuematic Component Design Kit in AMESim.

The AMESim simulation model of the relay valve consists of three parts: the air supply valve, the air evacuation valve and the main piston. The function of the relay valve is to control the pressure of the train pipe according to the pressure of the equalizing reservoir, and to realize the functions of air inflation, exhaust and maintaining pressure.

The inflation function is to increase the pressure in the train tube so that the train can run normally. The exhaust function is to reduce the pressure in the train tube to slow down or stop the train. The pressure maintaining function is to keep the train in the braking state.

As shown in Fig. 1, this paper proposes a fault prognosis method for pneumatic valve based on health indicators and support vector regression. Different fault prognosis models are trained with the data generated by different fault types. The importance of the degradation features was evaluated, and a group of features with the highest weight for each fault type was selected as sample data. Principal component analysis is used to calculate statistics T^2 and SPE as health indicators. The prognosis model is established based on the support vector regression algorithm.

It mainly consists of two parts, features extraction using principal component analysis and model establishment based on support vector regression.

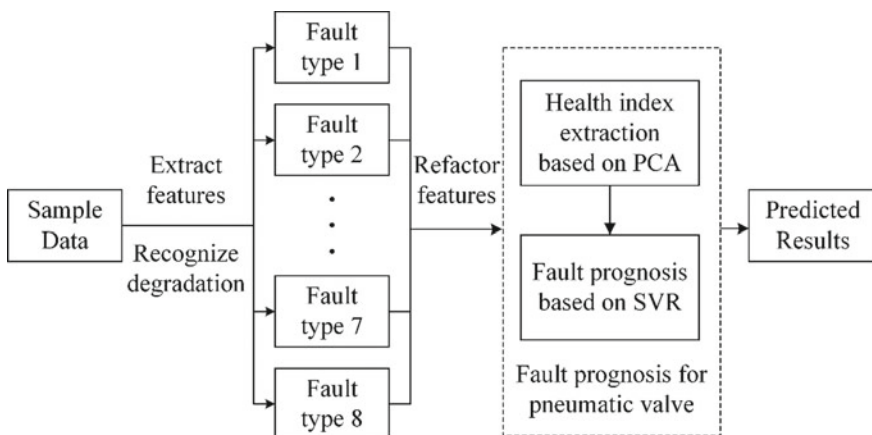


Fig. 1 The proposed fault prognosis framework for pneumatic valve

3 Fault Prognosis Using Health Index Extraction and Support Vector Regression

Pneumatic valves are key components of train braking system. Effective fault prediction method of pneumatic valves can avoid the failure of pneumatic valve and ensure the safety and reliability of train operation. Therefore, a fault prediction model of pneumatic valve based on health index and support vector regression is proposed in this paper.

3.1 Health Index Extraction Based on Principal Component Analysis

After the sample data reconstruction based on the importance of fault features, the health indicators corresponding to the data samples should be extracted according to the fault symptom types of the samples. Aiming at the problem of pneumatic valve health indicators are difficult to extract, this section on the basis of the characteristics of the reconstruction, this paper proposes a pneumatic valve health indicators on the basis of the principal component analysis model, mainly through the orthogonal transformation, the original data space decomposition is given priority to yuon space and residual space, while reducing the data dimension reserves the main information, and then extracted respectively from two Spaces health indicators.

The reconstructed degradation characteristics of pneumatic valves were used as the input data of the health index extraction method. Given a two-dimensional data matrix $X \in \mathbb{R}^{n \times m}$, n is the number of observations, m is the number of process variables included in the principal component analysis.

$$X = \hat{X} + E = TP + E = \sum_{i=1}^m t_i p_i^T + E \quad (1)$$

where \hat{X} is the cross product sum of subspace, t_i is the principal component vector; p_i is the projection direction of the principal component by transforming the basis vector; $E \in \mathbb{R}^{n \times m}$ is the residual matrix of the prediction error of the principal component model. E is measures the weak trend in the change of uncertainty and degradation process. According to the change trend, two statistics (T^2 and SPE) are obtained to detect the system state, as shown in Eqs. (2) and (3):

$$T_k^2 = \sum_{i=1}^A \frac{t_{ki}^2}{\sigma_i} \quad (2)$$

$$Q_k = \sum_{j=A+1}^m t_{kj}^2 \tag{3}$$

where T_k^2 and Q_k are the sample of T^2 and SPE at time k ; t_{ki}^2 and t_{kj}^2 are the component scores at time k ; $\sigma_i \in \mathbb{R}$ is estimated variance of the score variable. For different fault types, the data of pneumatic valve in normal working state are collected to establish the principal component model, and different principal component models are established for each fault type. When the pneumatic valve begins to degenerate, the health index T^2 and SPE gradually deviate from the principal component model, and when the deviation is too large, it indicates the failure of the pneumatic valve. The threshold equation of health index T^2 can be represented as:

$$T_\alpha^2 = \frac{A(n-1)}{n-A} F_{A,n-A,\alpha} \tag{4}$$

where n is the number of the samples, A is the number of the principal components. α represents significance level, which value is set at 0.01. A and $n-A$ are the confidence of the F distribution.

The threshold equation of health index SPE can be represented as:

$$Q_\alpha = \theta_1 \left(\frac{C_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right)^{\frac{1}{h_0}} \tag{5}$$

where, C_α is the value of the normal distribution at α the significance level, θ_i and h_0 can be calculated by Eqs. (6) and (7), which represents the eigenvalue of covariance matrix.

$$\theta_i = \sum_{j=A+1}^m \lambda_j^i (i = 1, 2, 3) \tag{6}$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^3} \tag{7}$$

Through the study in this section, the health indicators and threshold values corresponding to the degradation symptom types of each sample can be obtained, which the health indicator will be used to evaluate the current degradation state of the pneumatic valve and make a prediction, and the threshold value of the indicator will be used to judge whether the predicted health indicator has a failure.

According to the T^2 and SPE, the error conditions can be divided into four cases:

1. T^2 and SPE within their threshold limits;
2. T^2 within the threshold limits, but SPE is large than the threshold limits;
3. SPE within the threshold limits, but T^2 is large than the threshold limits;
4. T^2 and SPE are both large than the threshold limits;

Generally, (2), (3) and (4) are failure conditions, while (1) is a normal condition.

3.2 Error Prediction Based on Support Vector Regression

In this section, according to the health indicators of different fault types of pneumatic valves obtained, support vector regression models are built respectively for numerical prediction of health indicators. The health indexes of pneumatic valves were constructed as training samples

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$$

$$f(x) = \omega^T x + b \tag{8}$$

where $x \in \mathbb{R}^n$, $y_i \in \mathbb{R}$. $f(x)$ is the estimated output of the model. ω and b are respectively the weight and deviation corresponding to the input. But when in the case the distance between $f(x)$ and y is greater than ε , this case can be considered to be an error. This means that ω is as flat as possible.

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m \ell_\varepsilon(f(\mathbf{x}_i) - y_i) \tag{9}$$

where C is the regularization constant, and ℓ_ε is ε -insensitive loss function, which can be formulated as:

$$\ell_\varepsilon(z) = \begin{cases} 0 & , \text{if } |z| \leq \varepsilon \\ |z| - \varepsilon & , \text{if } |z| > \varepsilon \end{cases} \tag{10}$$

where $z = y_i - \omega x + b$. According to Eq. (9), the ξ_i and $\hat{\xi}_i$ are used as slack variable. Based on Eqs. (10), (11) can be formulated as:

$$\min_{\omega, b, \xi_i, \hat{\xi}_i} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^m (\xi_i + \hat{\xi}_i) \tag{11}$$

where $-\varepsilon - \xi_i \leq z \leq \varepsilon + \hat{\xi}_i$, $\xi_i \geq 0$, $\hat{\xi}_i \geq 0 (i = 1, 2, \dots, m)$. According to the Lagrange multiplier and optimal constraints, the kernel method of support vector regression model can be expressed as Eq. (12):

$$f(x) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(x, x_i) + b \quad (12)$$

where $\kappa(x, x_i) = \phi(x_i)^T \phi(x_j)$ is the kernel function. The commonly used kernel functions of support vector regression algorithms include linear kernel function polynomial kernel function radial basis kernel function and Sigmoid kernel function. The radial basis kernel function has a good performance when dealing with regression problems. The radial basis kernel function is selected as the kernel function of the support vector regression prediction model.

As the degradation process of pneumatic valves generally takes a long time, it is very difficult to obtain a large number of complete life cycle data. Therefore, in order to fully apply the training samples and improve the accuracy of the model, the idea of time sliding window is used to construct the prediction target artificially. The degradation process data of a pneumatic valve is slide-wise intercepted into multiple sample sets in the form of window, and each sample set includes training data and prediction indexes. In order to better train the model, different data sample sets are intercepted from the training set according to time, and each sample set includes training number samples and test samples. The support vector regression algorithm is used for multiple training, which can effectively improve the accuracy of the model.

4 Simulation Results and Discussions

In this section, the simulation indicators are introduced, then the joint simulation model fault acquisition method is described, and finally, it is revealed the time for fault occurrence by predicting health indicators based on degradation data.

4.1 Failure Prediction Evaluation Index

The prediction of failure is to predict future health indicators based on existing health indicators. In this paper, Mean Squared Error (MSE) and R Squared (R^2) are selected as the evaluation indexes of the pneumatic valve failure prediction model, which are calculated as shown in Eqs. (13) and (14).

$$MSE = \frac{1}{m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2 \quad (13)$$

$$\begin{aligned}
 R^2 &= 1 - \frac{\sum_i (\hat{y}^{(i)} - y^{(i)})^2}{\sum_i (\bar{y} - y^{(i)})^2} = 1 - \frac{\sum_i (\hat{y}^{(i)} - y^{(i)})^2/m}{\sum_i (\bar{y} - y^{(i)})^2/m} \\
 &= 1 - \frac{MSE(\hat{y}, y)}{Var(y)} \tag{14}
 \end{aligned}$$

According to Eqs. (13) and (14), the size of regression evaluation index MSE is related to the dimension of the model. MSE increases with larger dimensions. With consistent dimensions, the decrease in MSE will result in better model performance. The regression evaluation index R^2 removes the dimensionality of the model and returns an accuracy between 0 and 1. Training models with a higher R^2 will have better accuracy.

4.2 Preparation for Fault Prediction Data

In order to obtain the process data with fault occurrence, a joint simulation model of AMESim and Matlab was built, and Simulink was used to dynamically adjust the real-time parameters of a part of the relay valve to obtain the fault process data by artificially giving the degradation process curve of the part. When the real-time parameters are given, the model parameters are the same for each cycle, and the model parameters are different between cycles.

In Fig. 2, three degradation curves represent three different degradation processes. We model the different degradation processes for the rubber ring wear in relay

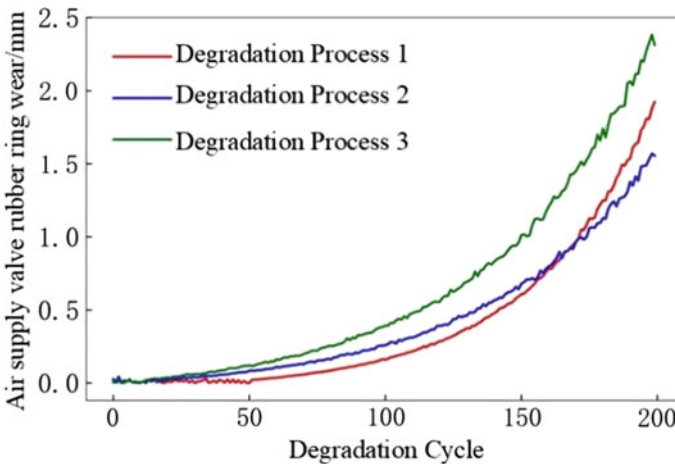


Fig. 2 The degradation process relay valve supply air valve rubber ring wear failure parameters change curve

valve supply air valves and demonstrate how the proposed failure prediction method accurately predicts the time of failure when the form of degradation is uncertain.

4.3 *Extracting Health Indicators by Principal Component Analysis*

The principal component models are established for the degradation process of rubber ring wear fault of relay valve air supply valve to extract health indicators. Since the establishment of the principal component model needs to test whether the original spatial data obey the normal distribution, K-S test is used to prove that the original spatial data obey the normal distribution, and the principal component analysis method is used to extract the health indicators T^2 and SPE from the original space.

Fig. 3 Health index T^2 of relay valve

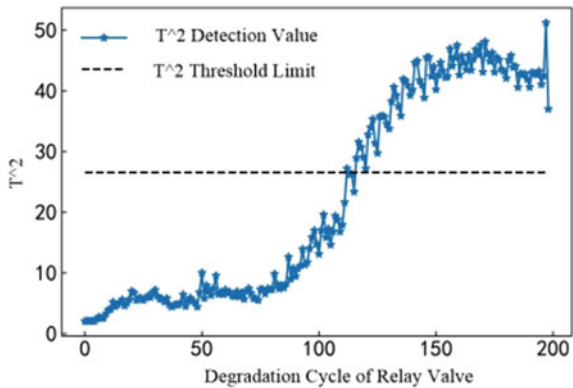


Fig. 4 Health index SPE of relay valve

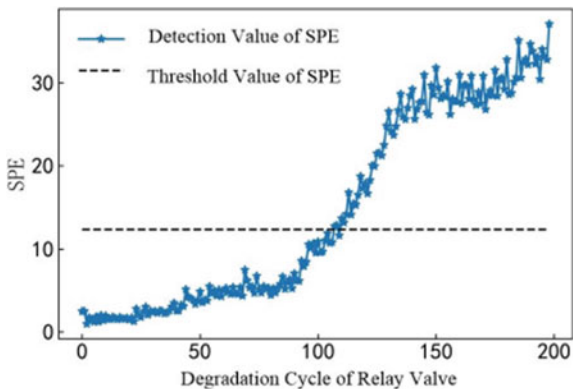
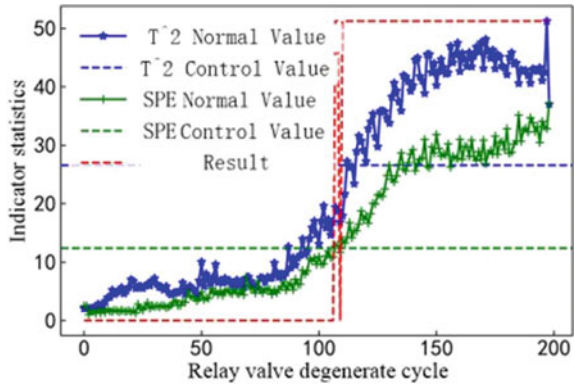


Fig. 5 Fault judgment results



In Figs. 3 and 4, the health indicators T^2 and SPE are represented by the blue curve, and represents the control limits of the two indicators are represented by the black dotted line. The fault judgment results are shown in Fig. 4.

In Fig. 5, the index T^2 is represented by the blue curve, the control limit of T^2 is represented by the blue dotted line, the index SPE is represented by the green curve, the control limit of SPE is represented by the green dotted line, and the fault detection results are represented by the red dotted line. When the red dotted line is equal to 0, it means that the relay valve has no fault. When the red dotted line is not equal to 0, it means that the relay valve has fault. It can be seen from the figure that the relay valve fails after the 105th cycle.

4.4 Result Analysis

The grid search method is used to select the super parameter value of support vector regression model. The value range and step size of super parameter are selected according to experience, and then the grid search method is used to search each super parameter combination within the value range according to step size.

It can be seen from Table 1 that the R-squared values of the first two training samples are relatively large, and the prediction results are relatively good, while the

Table 1 SVR Evaluation of fault prediction results

Evaluating indicator	R-Squared		MSE	
	T^2	SPE	T^2	SPE
Sample1	0.93357006	0.96067862	16.08351596	6.15154239
Sample2	0.48019985	0.91665023	44.06930541	23.42579304
Sample3	0.74403404	0.90005641	43.07602770	76.77761936
Sample4	0.48877394	0.88889388	20.67797548	64.13940336

R-squared values of T^2 health indicators of the last two training samples are relatively small, indicating that the prediction results are relatively poor. On the whole, the method proposed in this paper has good performance in both fault prediction results and fault prediction accuracy, and can effectively realize the fault prediction of pneumatic valve.

5 Conclusion

This paper proposes a novel data driven fault prognosis method for pneumatic valves in train braking system. Two health indicators, i.e., T^2 and SPE, are extracted through PCA method, which are further used to train a fault prognosis model based on SVR. The proposed method is validated on a semi-physical simulation verification platform of DK-2 braking system, results show that the proposed fault prognosis model can estimate the expected time of pneumatic valve fault time accurately.

References

1. Soo-Ho, J., Boseong, S., Hyunseok, O., Byeng, D.Y., Dongki, L.: Model-based fault detection method for coil burnout in solenoid valves subjected to dynamic thermal loading. *IEEE Access* **8**, 70387–70400 (2020)
2. Jianyong, Z., Jingxian, D., Shun, P., Guo, H., Tiefeng, Z., Jialiang, L.: Fault feature extraction of relay valve leakage based on test for standard EMU brake system. *DEStech Trans Eng Technol Res* (2017)
3. Yuan, X., Ying, L., Qunxiong, Z.: Multivariate time delay analysis based local KPCA fault prognosis approach for nonlinear processes. *Chin. J. Chem. Eng.* **24**(10), 1413–1422 (2016)
4. Daigle, M.: Real-time prognostics of a rotary valve actuator. In: Annual Conference of the Prognostics and Health Management Society 2015, San Diego, CA, United States (2015)
5. Chetan, S.K., Matthew, D., Kai, G.: Implementation of prognostic methodologies to cryogenic propellant loading testbed. In: IEEE International Automatic Testing Conference 2013, vol. 2013, pp. 1–7. IEEE, Schaumburg, IL, United States (2013)
6. Qing, L., Steven, Y.L.: Degradation trend prognostics for rolling bearing using improved R/S statistic model and fractional Brownian motion approach. *IEEE Access* **6**, 21103–21114 (2017)
7. Mechri, W., Hai-Canh, V., Phuc, D., Klingelschmidt, T., Peysson, F., Theilliol, D.: A study on health diagnosis and prognosis of an industrial diesel motor: hidden Markov models and particle filter approach. *Adv. Solut. Diagnos. Fault Tolerant Control* **635**, 380–389 (2017)
8. Xiaohang, J., Yi, S., Zijun, Q., Yu, W., Tommy, W.S.C.: Anomaly detection and fault prognosis for bearings. *IEEE Trans. Instrum. Measure.* **65**(9), 2046–2054 (2016)
9. Min, H., Jinbing, L., Bing, H., Kai, Z.: Fault subspace decomposition and reconstruction theory based Mlnline fault prognosis. *Control. Eng. Pract.* **85**, 121–131 (2019)
10. Lotfi, S., Jaouher, B.A., Eric, B., Mohamed, B.: Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR. *Appl. Acoust.* **120**, 1–8 (2017)
11. Jichao, H., Zhenpo, W., Yongtao, Y.: Fault prognosis of battery system based on accurate voltage abnormality prognosis using long short-term memory neural networks. *Appl. Energy* **251** (2019)
12. Hack-Eun, K., Andy, C.C.T., Joseph, M., Byeong-Keun, C.: Bearing fault prognosis based on health state probability estimation. *Expert Syst. Appl.* **39**(5), 5200–5213 (2012)