Condition monitoring of sensors with PCA method in nuclear power plants

LI Wei[1], PENG Minjun[1], LIU Yongkuo[1], and MA Zhanguo[1]

1. *College of Nuclear Science and Technology, Harbin Engineering University, 150001, Harbin, Heilongjiang, China (success870323@126.com,heupmj@163.com, lyk08@126.com,mazhanguo2013@163.com)*

Abstract: With the widespread application of digital I&C systems in Nuclear Power Plants (NPPs), more sensors are used to obtain operating information. A principal component analysis (PCA) method is applied in this paper to carry out condition monitoring for sensors in a NPP. Meanwhile to improve the model performance, a false alarm reducing method is proposed which is combined with PCA method in this paper. Sensor measurements from a real NPP are used to train and test the PCA model. Simulation results under normal operating condition indicate that the proposed false alarm reducing method really makes contribution to the model performance. Meanwhile, artificial failures with different degrees are sequentially imposed to test the functionality of the proposed PCA model, and the simulation results show that the proposed PCA model which is combined with a false alarm reducing method is effective on the condition monitoring of sensors no matter with major or small failures.

Keyword: NPP; sensor failures; PCA; condition monitoring; false alarm reducing

1 Introduction

With the wide application of digital I&C systems in nuclear power industry, more sensors are installed. On one hand, it is beneficial to get more operating information of the NPP. On the other hand, with the large application of sensors in the NPP, the number of sensors that may fail is significantly increasing[1].

Sensors usually work in high temperature, high pressure, high humidity or high radiation environment in a NPP, therefore various failures may appear, such as decreasing precision, drift or complete failure. These may result in the running deviated from the optimal condition which affects the economy of the NPP. In severe cases even safety and reliability of the NPP will be affected[2].

According to statistics in various industrial processes, more than 60% of process failures are caused by sensors directly or indirectly[3]. Since safety is primary of importance in nuclear industry, and safety related sensors are almost redundant in a NPP. Thus the process malfunctions resulted from sensor failures present a much lower level in nuclear industries compared to other process industries. Only about 10% of the abnormal behaviors or accidents are caused by sensor failures in a NPP according to the statistics of NPP running events by the world association of nuclear operators (WANO). As we all know, the fault detection and identification (FDI) system for a industrial process is based on the process measurements of the industry. As a result, if a NPP is configured with a FDI system, then the incorrect measurements from faulty sensors will convey to the FDI system of the NPP. Further, wrong decisions may be made by the FDI system based on the incorrect sensor measurements, and a corrective action will be promptly initiated automatically or manually by operators which actually leads to the running deviated from the predetermined condition. Then the subsequent huge economic loss, unnecessary downtime or some severe accidents may be caused, that is safety of the NPP is endangered[4-6].

Therefore, it is necessary to implement condition monitoring for sensors in order to improve the economy and reliability of the NPP[7]. The general idea for sensor condition monitoring is to compare the residuals between the actual sensor measurements and the values by sensor estimation models. And redundancy is the only way to generate residuals. According to the generating way of redundancy, sensor condition monitoring methods can be divided into two categories: physical redundancy methods and analytical redundancy methods. The physical redundancy methods usually obtain the residuals by in-

Received date: May 8, 2017
(Revised date: May 23, 2017)

creasing the number of sensors, and usually three or more sensors are required to measure the same variable. In this way, not only is the equipment cost greatly increased, but also the number of sensors required to be monitored, calibrated and maintained is significantly increased. The analytical redundancy methods apply analytical models to generate estimation values of sensors. Usually, it is not necessary to increase the number of sensors. The analytical redundancy methods can be further divided into two classes according to the type of analytical models: the mathematical model and the data-driven model. The mathematical model generates the estimation values by the precise mathematical model of the sensors. However it is difficult to establish the precise mathematical models of sensors in practice. Thus the data-driven models which are based on operating data are adopted in many industries due to the convenience and usability of the data acquisition system in modern industries.

A lot of work has been done for sensor condition monitoring with data-driven methods in many industries. In the literature, the most used data-driven methods are multivariate state estimation technique (MSET)[8], artificial neural network (ANN)[9-10], independent component analysis (ICA)[11], support vector machine (SVM)[12-13] and PCA[14-18,4].

A comprehensive investigation of the condition monitoring techniques with relevance to NPPs was detailed presented in a review paper by Ma[19]. Xu *et al*. adopted a neural network technique for sensor validation in a power generation process[20]. An ANN was applied for sensor FDI in a distillation process by Perla *et al*. To get better performance, the time-delay effects in the distillation process were considered during the training stage[21]. Andrew and Song proposed a sensor condition monitoring method for industrial combustion process which was based on clustering algorithms. Current data points from the process were compared with the clusters to identify sensor faults. The method showed its better performance on sensor condition monitoring with the simulation test in an industrial boiler process[22]. Liu *et al*. proposed a distributed fault diagnosis system for a NPP based on fuzzy neural network architecture. In the diagnosis system, local and multi-source information were combined to allow for an advanced type of global fault diagnosis[23].

The traditional method for sensor condition monitoring is physical redundancy method in the NPP. The major problem for this method is the cost and some inherent limitations. The inherent limitations existed in physical redundancy method are explained as follows:

(1) If the residuals between redundant sensors are within the confidence limit, and then the failure will not be detected by physical redundancy method. However the running of the NPP has deviated from the optimal condition.

(2) If the residuals between redundant sensors are beyond the confidence limits. In this context, the failure could be detected by physical redundancy method. However, if the failure occurs on the majority or all sensors in the redundant group simultaneously, then the physical redundancy method will produce an incorrect fusion result. The foregoing assumptions are reasonable, since the redundant sensors are usually manufactured by the same factory, and exposed in the same environment [24].

A study conducted by Hines *et al.* concluded that the simplicity of analytical redundancy techniques based on data-driven models and the tractability of their uncertainty calculations could favor them for acceptance by regulatory bodies[25]. Meanwhile with the wide application of digital I&C systems in a NPP, more operating information (namely sensor measurements) is available. And the advanced computer technology also contributes to application of PCA method. Again PCA model also can implement condition monitoring for sensors with or without redundant configuration. Then the reliability and economy of the plant is greatly improved. Because of these reasons, PCA model is adopted for sensor condition monitoring based on its relative simplicity and individual strong points in this paper.

The previous successful applications of data-driven methods on sensor condition monitoring in a NPP are as follows. One is the PEANO system developed by institute for energiteknikk (IFE) in Norway which is on the basis of Halden project. The system is based on the ANN and fuzzy logic methods, which is verified to be effective on the detection of sensors with severe drifts. However sometimes it may fail to identify the detailed faulty sensor[36]. AFAL analysis is proposed for the calibration of sensors in the technical report by EPRI in 2004. It really contributes to the extension of the calibration cycle of sensors in a NPP, however on-line monitoring for the NPP sensors cannot be realized with this AFAL analysis method[7]. Whereas all these issues can be solved by a PCA model which will be described in detail in the following sections.

On the other hand, as is also often the case in the previous studies with PCA models, false alarms are inevitable during the condition monitoring processes by PCA models. Thus a false alarm reducing method is combined with the proposed PCA model to reduce the false alarms existed in the model. That is, the proposed PCA model can be more effective on the detection and identification of faulty sensors in the monitored sensor group with the application of this false alarm reducing method. In this way, the accuracy and reliability of the PCA model are greatly improved in this paper.

The paper is organized as follow: Section 1 describes the necessity of sensor condition monitoring and a comprehensive PCA model is proposed. Section 2 outlines the PCA methodology. In section 3, a false alarm reducing method is demonstrated to show its significant contribution on the performance improvement of PCA model. The proposed PCA model is tested and evaluated with measurements acquired from a real NPP in section 4. Conclusions and future work are given in the last section.

2 The methodology of sensor condition monitoring with PCA model

PCA is a kind of multivariate statistical analysis methods which is widely used in data analysis. It transforms a set of correlated variables into a smaller set of new variables (principal component, PC) that are uncorrelated and retains most of the information in the original data[16-17]. The reduced PCs obtained from the uncorrelated variables are used to detect the abnormalities of the operating process in a robust way[26]. The basic theories of PCA method and the fault detection procedures will be briefly depicted in this paper, the specific derivation processes can refer to Li, He or Camacho[27-29]. Meanwhile a fault identification methodology based on contributions of various variables to the total variance in PC and residual space is proposed, and it will be described in detail.

2.1 Basic theories of PCA method

In general, an original data matrix (n samples and m variables) can be given as:

$$X_0 = [x(1), x(2), \ldots x(i), \ldots, x(n)]^T \tag{1}$$

where $x(i)$ is a sample vector in X_0, it can be described in detail as:

$$x(i) = [x_1(i), x_2(i), \ldots, x_m(i)] \tag{2}$$

The original data matrix X_0 should be normalized first to eliminate the influence caused by different magnitudes of variables in X_0, and X_0 is then scaled to a new data matrix X with zero mean value and one unit variance. After that, the new data matrix X is projected onto a new space ordinate system by making use of a linear transformation P:

$$T = XP \tag{3}$$

where T is the score matrix and P is the loading matrix respectively. And loading matrix P can be derived from the covariance matrix of X. In this case, T and P can be expressed as the following form:

$$\begin{aligned} T &= [t_1, t_2, \ldots, t_m] \\ P &= [p_1, p_2, \ldots, p_m] \end{aligned} \tag{4}$$

The vectors t_i in matrix T are orthonormal, they are the linear combination of the data matrix X and represent that how the samples are related to each other. Meanwhile the vectors p_i in matrix P are also orthonormal, and they are the eigenvectors of covariance matrix of X. These eigenvectors indicate that how the variables are related to each other. Each orthonormal vector p_i is associated with the eigenvalue λ_i of covariance matrix of X, that is:

$$C = P \Lambda P^T \tag{5}$$

where C is the covariance matrix of X, and $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_m)$. Here diag() is a function that used to generate a diagonal matrix with the diagonal elements being $\lambda_1, \lambda_2, \ldots, \lambda_m$.

Then the PCs can be determined according to the eigenvalues. There are different criteria to select the number of PCs[30]. In this paper, the commonly used cumulative percent variance (CPV) criterion is adopted. CPV represents the variation of the selected PCs account for all the variation of X. As a result, to determine the number of PCs (namely k), CPV can be defined as:

$$CPV = \frac{\sum_{i=1}^{i=k} \lambda_i}{\sum_{i=1}^{i=m} \lambda_i} \times 100\% \tag{6}$$

Based on the selected PCs, the data matrix X can be decomposed into the sum of a PC matrix $\hat{X}$ and a residual matrix E. The PC matrix contains information of the system variation. Whilst the residual matrix mainly contains information of the noise or model error[31].

$$X = \hat{X} + E = T_k P_k + E \tag{7}$$

where $P_k = [p_1, p_2, ..., p_k]$, and $T_k = [t_1, t_2, ..., t_k]$. Then the following fault detection and identification processes will be carried out in the PC and residual matrixes respectively[32].

2.2 Fault detection with PCA method

There are two commonly used statistics to implement fault detection in a PCA model: Q statistic and Hotteling's T^2 statistic. They are defined to measure the variation in residual and PC space respectively. If Q or T^2 statistic exceeds the effective region in PC space or a significant residual is observed in residual space, a special event, either due to disturbance or due to changes in the relationship between variables, can be detected[32].

Q statistic is the squared prediction error between the testing vector and the model. It quantifies the distance a testing vector falls from the PC model. For a testing vector x, $x = [x_1, x_2, ..., x_m]$, Q statistic can be defined as:

$$Q = x(I - P_k P_k^T) x^T \le Q_\alpha \tag{8}$$

Meanwhile the Hotteling's T^2 statistic measures the variation within the PCA model. It can be defined as:

$$T^2 = t_i \Lambda^{-1} t_i^T = x P_k \Lambda^{-1} P_k^T x^T \le T_\alpha^2 \tag{9}$$

Q_α and T_α^2 are the corresponding confidence limits for Q and T^2 statistics respectively. The specific calculation of the confidence limits can refer to the doctoral thesis by Li[32].

From Equation(9), it can be seen that T^2 statistic is based on the top k eigenvalues; and it is the remaining (m-k) eigenvalues rather than the top k eigenvalues that are applied in the calculation of Q statistic. Thus Q statistic and T^2 statistic are applied simultaneously to implement fault detection in this paper. If any statistic exceeds its corresponding limit, it indicates that some unknown failures have been occurred in the monitored sensor group.

2.3 Fault identification with PCA method

When Q and T^2 statistics exceed the corresponding confidence limits, abnormality is detected. Then fault identification is required to locate the faulty sensor. The contributions of various variables to the total variance in PC space and residual space are used together to locate the faulty sensor. Just as analyzed above, T^2 statistic represents the total variance in PC space, and Q statistic measures the total variance in residual space[33]. The calculating processes of the fault identification indexes are shown as follows.

Supposed that a testing vector x is expressed as $x=[x_1, x_2,\dots,x_m]$, m is the number of variables in x. The contribution of variable x_i to the total variation in residual space is defined as:

$$Q_{x_i} = \frac{\left\| x_i (I - PP^T) \right\|}{\left\| x(I - PP^T) \right\|} = \frac{e_i^2}{e^2} \times 100\% = \frac{e_i^2}{e_1^2 + e_2^2 + ... + e_m^2} \times 100\% \tag{10}$$

The sum of the contributions of all m variables to the total variation in residual space is exactly equal to the Q statistic of x. A large contribution on x_i usually means a faulty state on sensor i.

The contribution of variable x_i to the total variation in PC space can be calculated as the following steps.

(1)Calculating the contribution of x_i to score vector t_j :

$$CR_{j,x_i} = \frac{t_j p_{j,i}}{\lambda_j} x_i \qquad (i = 1,2,...,m) \tag{11}$$

where $p_{j,i}$ is the i$^{\text{th}}$ element of eigenvector p_j .

(2)Calculating the contribution of x_i to T^2 statistic:

$$T_{x_i}^2 = \sum_{j=1}^{k} CR_{j,x_i} = \sum_{j=1}^{k} \left(\frac{t_j p_{j,i}}{\lambda_j} x_i \right) \qquad (i = 1,2,...,m) \tag{12}$$

The sum of the contributions of all m variables in PC space is equal to the T^2 statistic of x. Similarly, a large contribution on x_i usually means a faulty state on sensor i as well.

Usually if failures occur in the monitored sensor group, both or any of T^2 and Q statistics will exceed their confidence limits, then Q_{x_i} and $T_{x_i}^2$ can be directly applied to locate the faulty sensor. However, if the failure occurs on sensors is just a small glitch which may not be detected by T^2 and Q statistics until the fault reaches a severe level. In this context, before T^2 and Q statistics detect the small drift, a relatively evident increasing trend can be seen in Q_{x_i} and $T_{x_i}^2$ of the drift sensor. That is, these two indexes can also make contribution to the timely fault detection for sensors with small drifts.

Small drifts in sensors may not result in severe accidents, but if the drift sensor participates in important control processes of the NPP, it can lead to the operation deviated from the optimal condition which may be involved in potential decline of the plant economy. On the other hand, if sensors with small drifts which do not participate in important control processes and just served as monitoring purposes, these two indexes can also make contribution to the realization of condition-based maintenance (CBM) strategy in a NPP. Sensors can be calibrated, maintained or repaired when they are required based on the operating conditions determined by these two indexes. Thus, these two indexes provide guidance on the maintenance schedule of a NPP, and excessive calibration and maintenance activities for sensors can be avoided to some extent.

2.4 The condition monitoring framework for sensors with PCA method

The flow chart for sensor condition monitoring with PCA method in this paper is illustrated in Fig.1. Especially, contents in the ellipses is another main part in this paper: false alarm reducing method. It is used to improve the performance of PCA model. And it will be explained in the following section in detail.

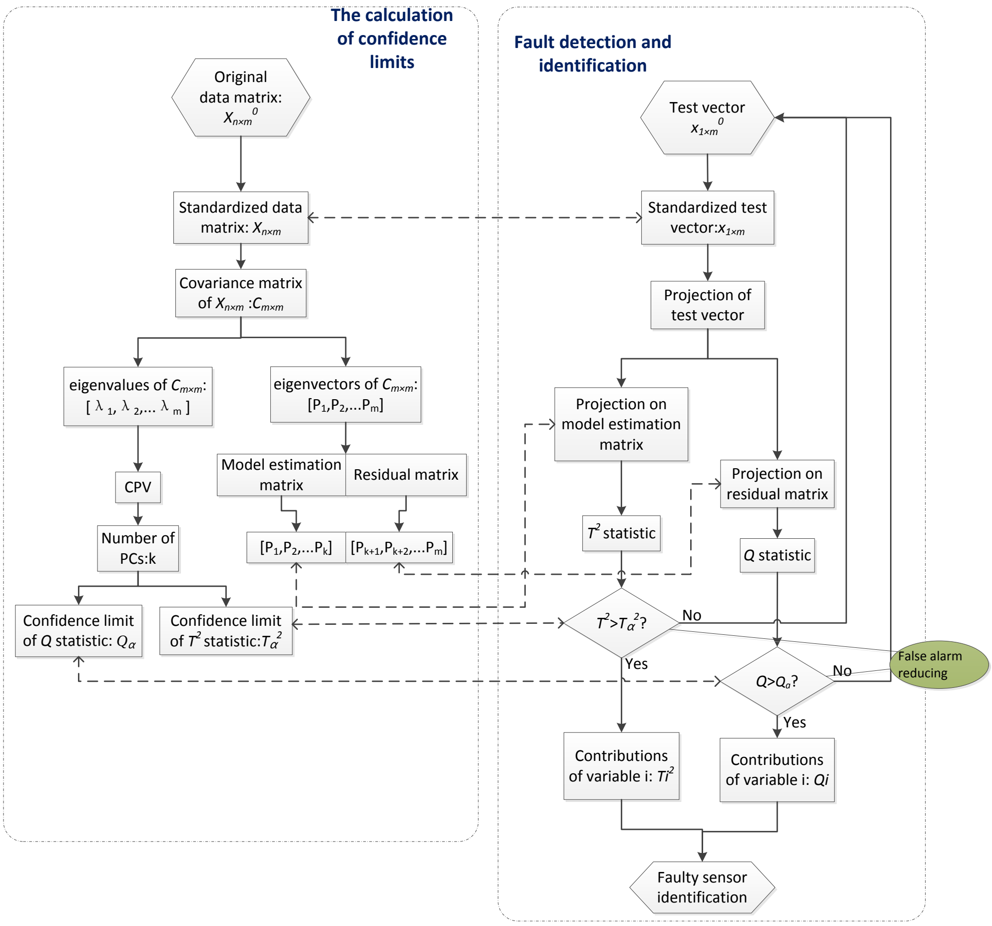

Fig.1 The flow chart for sensor condition monitoring with PCA method.

3 False alarm reducing method

If T^2 or Q statistic is beyond T_α^2 or Q_α under normal operating condition, then it will be called as a alarm condition. T^2 and Q statistics should be within their confidence limits under normal operating condition in theory. In practice, the external environment factors and internal modeling error factors have combined to result in false alarms of T^2 and Q statistics. That is, false alarms refer to a condition which is normal in fact, however it is determined as faulty by T^2 and Q statistics of the PCA model due to the foregoing mentioned factors. As a result, another confidence limit is proposed to reduce the false alarms for T^2 and Q statistics to a lower level which can be easily accepted by NPP operators in this paper. If Q_α and T_α^2 are the first confidence limits, this confidence limit is called the second confidence limit.

Supposed that the probability of false alarms for T^2 and Q statistics is α . In accordance with the statistic experience in process industry, the commonly used experience value for α is between 0 and 0.05[34]. And this experience value is also adopted as the false alarm probability for T^2 and Q statistics in this paper. In other words, if α =0.05, T^2 or Q statistics will still exceed the corresponding T_α^2 or Q_α with a probability of 0.05 under normal operating condition. It results from the random fluctuations in the testing data which is inevitable. However we can try our best to reduce the false alarms to a lower level which can be accepted by the operators in the NPP.

Taking n as a basic observation unit (that is a sequence of testing vectors within a time window of size n), the probability distribution of T^2 and Q statistics in each unit can be approximately expressed as:

$$P(m;n) = C_n^m \alpha^m (1-\alpha)^{n-m} \tag{13}$$

where m is the number of alarms in n. It can be seen that it is referred to the binominal distribution if all testing vectors in each unit are independent with each other. Then the second confidence limit can be derived from the following formula:

$$F(m;n) = P(i;n) = \sum_{i=0}^{m} C_n^i \alpha^i (1-\alpha)^{n-i} < \beta \tag{14}$$

where β is also an experience value which is determined by the model precision. Usually it is set between 0.95 and 0.99 according to the statistics in process industries. Then the largest allowable m in n can be derived from Equation (14). And the allowable m is just the second confidence limit for T^2 or Q statistic in this paper. If the number of alarms for T^2 or Q statistic is more than m in a observation unit n, then the current testing vector x_j is regarded as a true alarm. Otherwise it will be treated as a false alarm and ignored.

Since the model performance is directly related to the accuracy of the PCA model. And the accuracy represents how accurately the PCA model can detect the faulty condition and identify the specific faulty sensors in the monitored sensor group. With the application of this false alarm reducing method, it is obvious that the accuracy of the PCA model is greatly increased due to the reducing of false alarms of T^2 and Q statistics in the PCA model. As a consequence, the conclusion can be drawn that the performance of the proposed PCA model can be significantly improved with the application of the false alarm reducing method.

Considering the sensitivity of fault detection, a large value for n is inadvisable; meanwhile considering the effectiveness of false alarms reducing, a too small value for n is also inadvisable. Based on the foregoing analysis, as well as the simulation tests with different lengths of observation unit, n=8 is adopted in this paper as the reasonable length of observation unit. With various experience values of α and β , the corresponding second confidence limits m are summarized in Tab.1.

Table 1 Second confidence limit *m* with various reference values of α and β for T^2 or Q statistic.

α =	0.01	0.02	0.03	0.05
β =0.99	2	3	3	3
β =0.98	2	2	3	3
β =0.95	2	2	2	3

4 Simulations and results

Since thousands of sensors are applied in a NPP, it is impossible to put all the sensors into a single PCA model. As a result, a distributed condition monitoring framework is proposed, and there may be several PCA models running in parallel. How to separate the sensors into various PCA models are not the focus in this paper. However it is still expected that most suitable sensors can be grouped together to get better model performance. Considering these, a specific sensor grouping method rather than a random grouping way is applied in this paper for a PCA model. As it is known that PCA is a linear analysis method in previous analysis[4]. Naturally it is advantageous to group the linear dependent variables (namely various sensors) into a single set. While All the operating measurements would present more or less nonlinear relationship in a real NPP, and it is impossible that the measurements of various sensors are absolutely linear dependent in practice. Correlation coefficient is just the metric of the linear relationship for various variables[35]. A larger correlation coefficient indicates a higher linear relationship between various variables. Thus faced with this issue, what we can do is to put the variables with higher correlation coefficients into a single PCA model.

Supposed that the reactor coolant temperature sensor in some channel is selected as an example to show the procedures. And this sensor is exactly marked NO.1 in the original database. According to the foregoing grouping criterion, the correlation coefficients between this sensor and all the other sensors in the original database are calculated. In this paper, 15 sensors are applied in a PCA model. Then another 14 sensors with top correlation coefficients with NO.1 sensor are picked out. The detailed correlation coefficients with NO.1 sensor are as follows which are arranged from large to small order: [1 0.9419 0.9177 0.9158 0.9154 0.9153 0.9025 0.8982 0.8940 0.8935 0.8917 0.8902 0.8852 0.8823 0.8821]. Meanwhile the corresponding numbers for the selected 15 sensors in original database are [1 40 39 33 21 10 34 9 45 24 5 20 11 22 13]. And depending on the original database, it can be seen that all the selected 15 sensors are used to measure the reactor coolant temperature in various channels.

4.1 Simulations with normal sensor measurements

To test the functionality of the proposed PCA model, sensor measurements are acquired under normal operating condition from a real NPP with full power. 1000 samples (namely 1000 training vectors) are acquired to train the PCA model and another 1000 samples (namely 1000 testing vectors) are acquired for the simulation test. The results are given in Fig.2. The red dotted lines are the corresponding confidence limit for T^2 and Q statistics. It can be seen that Q statistics present quite a few alarms under normal operating condition. For T^2 statistics, the situation is much better, and only several alarms appear during the whole test. Since as we all know the simulation is carried out with normal data, then these alarms can be seen as false alarms before any further processing.

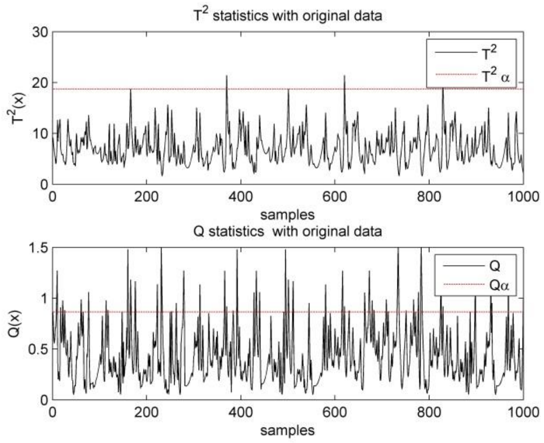

Fig.2 T^2 and Q statistics of PCA model with normal testing measurements.

To reduce the false alarms existed in T^2 and Q statistics, a false alarm reducing method is proposed. Figure 3 displays the results. The red circles in Fig.3 represent the remaining alarms after the application of false alarm reducing method. Obviously, false alarms have been greatly reduced with the application of this method. And the existence of the remaining alarms are still false alarms since the simulation is under normal operating condition, some other effective methods are further required to reduce these false alarms in the future.

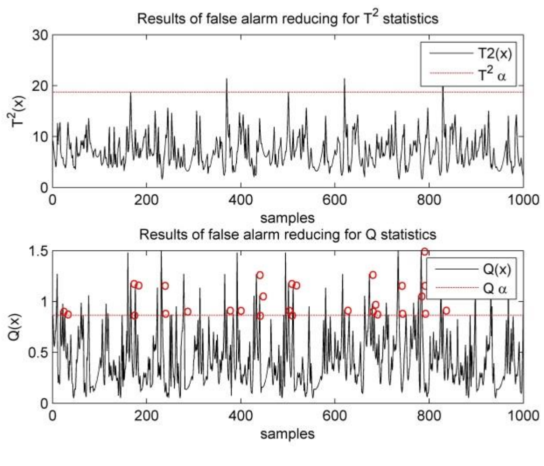

Fig.3 Results of false alarm reducing for T^2 and Q statistics with normal testing measurements.

To directly show the effectiveness of the proposed false alarm reducing method, the data results are summarized in Tab.2. As it can be seen, the probability of false alarms for T^2 statistics can be reduced to zero with the application of this technique. Similarly, the probability of false alarms for Q statistics can be reduced from 6.9% to 1.9%. And 1.9% has been a very low level that can be accepted in practice.

Table 2 False alarm probability for T^2 and Q statistics in PCA model with or without false alarm reducing method.

	With original data	With false alarm reducing
T^2	0.5%	0
Q	6.9%	1.9%

Under normal operating condition, the contributions of variables to T^2 and Q statistics are shown in Fig.4. It can be seen that the contributions of the 15 variables either to T^2 or to Q statistics are not absolutely equal due to the influence of model precision or other unknown uncertainties. Thus it is not reliable to infer a faulty sensor only depending on the contributions at a single testing point. As a result, two testing points are selected (namely the 600th and 1000th testing points) as a contrast.

Q statistics are taken as an example for detailed explanation. At 600th testing point, the contribution of NO.1 sensor to Q statistics is about 11%, meanwhile the contribution of NO.24 sensor to Q statistics is about 5%. It is clear that there is a large difference. However at 1000th testing point, the contribution of NO.1 sensor to Q statistics is still around 11%, and also that of NO.24 sensor Q statistics is still around 5%. That is, the contribution of NO.1 sensor to Q statistics is different from that of NO.24 sensor at 600th or 1000th testing points, which is also true at any other testing points. Meanwhile it also can be seen that the contribution of NO.1 sensor to Q statistics at 600th testing points is almost equal to that at 1000th testing points, which is also true at any other testing points. The case is the same for NO.24 sensor as well as all the other sensors in the monitored sensor group. In other words, there are differences on the contributions of variables at different testing points, there are no obvious contribution change for each variable between different testing points. Then the conclusion can be drawn that no faulty sensor occurs.

4.2 Simulations with abnormal sensor measurements

To verify the FDI ability of the PCA model, artificial failures are imposed to the testing data. In the selected 15 sensors, NO.1 and NO.10 sensors both are used to measure the coolant outlet temperature. Thus two artificial drifts are imposed to NO.1 and/or NO.10 sensor at 400th testing sample point respectively. One drift simulates a common problem that affects process sensors and may result from aging. The simulated drift is a ramp that grows to 0.4 ℃ (maximum value) for NO.1 or NO.10 sensor measurements. It corresponds to a 0.13% change which is imperceptible in a time profile. Another drift is bigger which represents a common issue that may affect process sensors and may result from mechanical failures. This simulated drift is also a ramp that grows to 3.2 ℃ (maximum value) for NO.1 or NO.10 sensor measurements. And this drift is equivalent to a 1.05% change of the sensor measurements which can be seen in the time profile. Both of the foregoing drift failures on NO.1 and NO.10 sensors will not arouse the proportion integration differentiation (PID) controller, that is to say, the other sensor measurements in the monitored sensor group will not be affected by the measurement changing of NO.1 and NO.10 sensors.

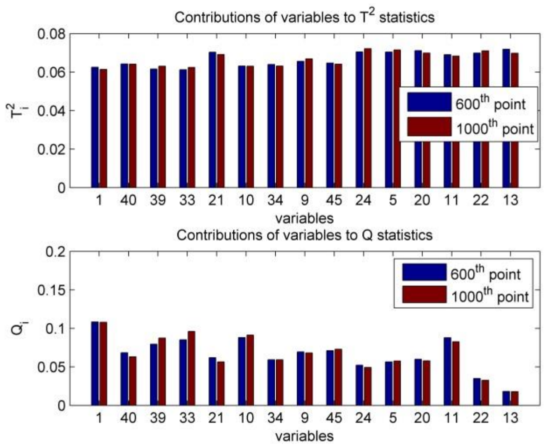

Fig.4 The contributions of variables to T^2 and Q statistics with normal testing measurements.

4.2.1 Sensor measurements with a small drift

Firstly the small drift is only imposed on NO.1 sensor measurements, and the results for this case are shown in Fig.5. This small drift simulates the inherent limitation (1) in physical redundancy method which is mentioned in section 1. From Fig.5 we can see that T^2

statistics cannot detect the drift on NO.1 sensor. However Q statistics gradually present an increasing trend at the later phase of the test. Thus, based on Q statistics, faulty sensor still can be detected. Then contributions of variables to T^2 and Q statistics are applied to identify the faulty sensor. The corresponding results are given in Fig.6.

For the contributions of variables to T^2 statistics in Fig.6, only a small contribution increasing can be seen on NO.1 sensor between 600th and 1000th testing points. This can be explained in Fig.5, since T^2 statistics either have no significant change during the whole test. For the contributions of variables to Q statistics, the situation is completely not the same. At 600th testing point, the contribution of NO.1 sensor to Q statistics is about 12%, while it almost reaches to 38% at 1000th testing point. A significant increasing occurs on the contribution of NO.1 sensor between 600th and 1000th testing points. In contrast, declining contributions are present for all the other 14 sensors between 600th and 1000th testing points. For example, the contribution of NO.10 sensor to Q statistics is about 8% at 600th testing point, and it goes down to about 4% at 1000th testing point. All this phenomena imply that there is something wrong with NO.1 sensor.

As a consequence, it is entirely within the capacity of the proposed PCA model to identify the faulty sensor. Even if the failure is only a small drift with a maximum 0.15% change on NO.1 sensor which cannot be detected by physical redundancy method placed in practice.

The false alarm reducing results under this case are shown in Fig.7. It can be seen that more and more true alarms appear with the development of the failure. Then a faulty state is determined finally.

Then the small drift is imposed to NO.1 and NO.10 sensor measurements simultaneously to demonstrate the effective fault detection and identification ability of the PCA model. It simulates the inherent limitation (2) in physical redundancy method which is mentioned in section 1. The results for this case is shown in Fig.8.

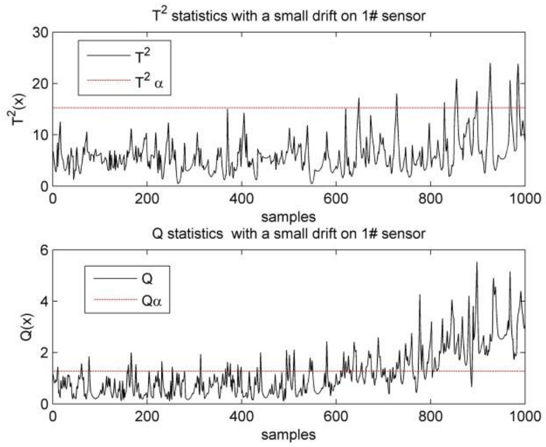

Fig.5 T^2 and Q statistics of PCA model with a small drift on NO.1 sensor.

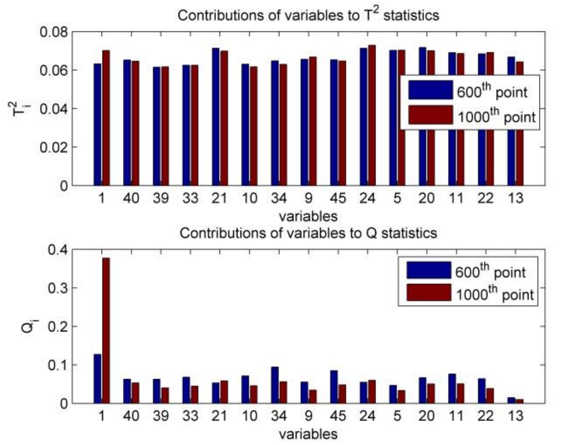

Fig.6 The contributions of variables to T^2 and Q statistics with small drift on NO.1 sensor.

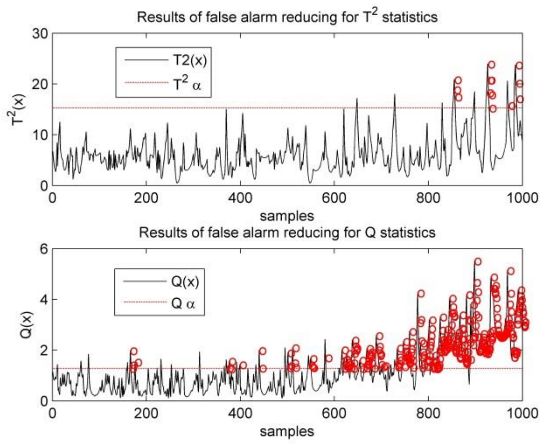

Fig.7 The results of false alarm reducing for T^2 and Q statistics with small drift on NO.1 sensor.

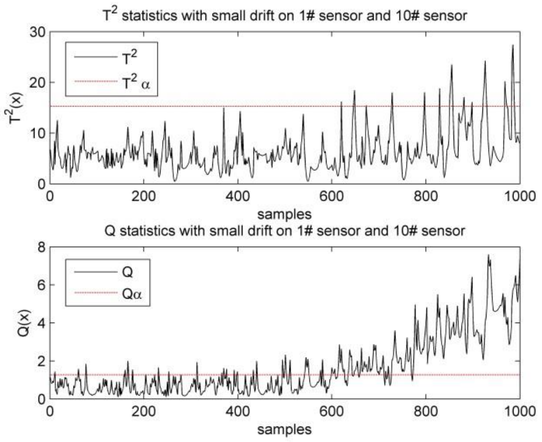

Fig.8 T^2 and Q statistics of PCA model with small drift on NO.1 and NO.10 sensor.

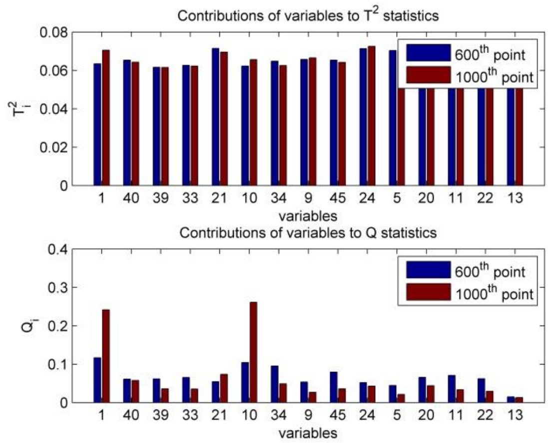

Fig.9 The contributions of variables to T^2 and Q statistics with small drift on NO.1 and NO.10 sensor.

From Fig.8 we can see that both T^2 and Q statistics present a faster speed on fault detection compared to the previous situation in Fig.5. Based on traditional physical redundancy method, this drift will not arouse alarms. Since the residuals among redundant sensors are too small to exceed the alarming threshold. On the other hand, even the imposed failures are enough to arouse the alarms, since the measurements from two drift sensors are corresponding, then the residuals between the two drift sensors will be within the alarming threshold, while the residuals between any drift sensor and the third normal sensor will be beyond the alarming threshold. As a consequence, the third redundant sensor (which is normal in fact) will be treated as the faulty sensor based on the 2 of 3 logic. The final fusion coolant outlet temperature will be the average of two drift sensor measurements. Obviously, this fusion result is incorrect.

Meanwhile the contributions of the 15 sensors to T^2 and Q statistics are given in Fig.9. From the contributions of variables to Q statistics, it can be inferred that some failures appear on NO.1 and NO.10 sensors. Since only the contributions of NO.1 and NO.10 sensors present increasing trend from 600th to 1000th testing points, while the contributions of the other 13 sensors almost present decreasing trend from 600th to 1000th testing points. According to the simulation test, it can be seen that these identification indexes are also beneficial to help identify multiple failures in a physical redundant sensor group.

4.2.2 Sensor measurements with a major failure

In contrast, the results with a larger drift imposed to NO.1 sensor measurements are shown in Fig.10. The figure indicates that the PCA model can detect the faulty sensor with a much faster speed compared to the foregoing simulations. It is understandable that a PCA model is certainly more sensitive to a major failure. The contributions of the 15 sensors to T^2 and Q statistics are shown in Fig.11.

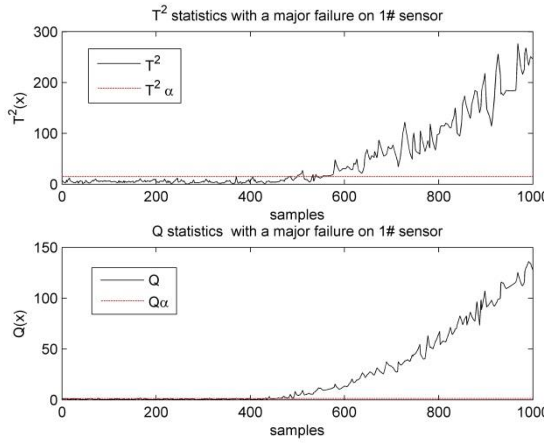

Fig.10 T^2 and Q statistics of PCA model with major failure on NO.1 sensor.

Figure 11 illustrates the results of fault identification. Compare to the above cases, the contribution of NO.1 sensor is obviously larger in this case. The contribution of NO.1 sensor to Q statistics reaches almost 75% at 1000th testing point. Meanwhile the contributions of all the other 14 sensors to Q statistics are almost below 5%. Even for T^2 statistics which is not very sensitive to small failures, the contribution of NO.1 sensor also come up to 35% at 1000th testing point which is significantly bigger than that of other sensors. As a result, the conclusion can be drawn that a major failure is occurred on NO.1 sensor.

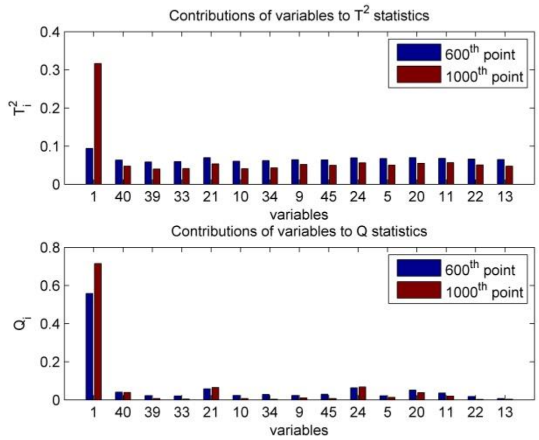

Fig.11 The contributions of variables to T^2 and Q statistics with major failure on NO.1 sensor.

5 Conclusions and perspectives

In this paper, a PCA model is applied to implement condition monitoring for sensors in a NPP. Two fault identification indexes are proposed to identify the faulty sensors in this paper. Meanwhile to improve the performance of the proposed PCA model, a false alarm reducing method is combined with the PCA model.

The efficiency of the proposed PCA model is evaluated by the sensor measurements acquired from a real NPP. The simulation results with normal and abnormal sensor measurements show that the proposed PCA model can successfully detect and identify the single or multiple faulty sensors no matter with small or major failures. Meanwhile the results also indicate that the statistics-based false alarm reducing method can greatly reduce the false alarms of T^2 and Q statistics, and the performance of the PCA model is significantly improved with the application of this method. In this way, the reliability of the proposed PCA model is also further improved.

Although these valuable achievements have been attained in this paper, there is still some work to do in the future. The reconstruction of the faulty sensor measurements is another important task. How to best reconstruct the faulty sensor measurements when the faulty sensor is located will also be discussed in detail in another paper.

Acknowledgement

The authors acknowledge the financial support of the national project of “Research on on-line monitoring and operation support techniques in a nuclear power plant” for the research present.

References

[1] HUASONG, F., XINMING, J., JIANWEN, L., and FENGSHENG, X.: Preliminary study on aging management of sensors in nuclear power plant. Instrument and Sensor, 2009, 0422(3):422-425.

[2] JIANHONG, C., HONGKUN, L., DEREN, S., and WEI, L.: A hybrid data-driven modeling method on sensor condition monitoring and fault diagnosis for power plants. Electrical Power and Energy Systems, 2015, 71:274-284.

[3] YANG, J. C-Y, and CLARKE, D. W.: A self-validating thermocouple. IEEE trans. On control systems technology, 1997, 5(2):239-253.

[4] KECHANG, F.: Structure optimized PCA and its application. Zhejiang University: Control Science and Control Engineering, 2007.

[5] On-line Testing of Calibration of Process Instrumentation Channels in Nuclear Power Plants. US Nuclear Regulatory Commission, NUREG/CR-6343, 1995.

[6] Guidance on Cross Calibration of Protection System Resistance Temperature Detectors. UNITED STATES NUCLEAR REGULATORY COMMISSION, Washington DC., NUREG-08001997.

[7] Guidelines for Instrument Calibration Extension/Reduction-Revision 2. EPRI 2014 Technical Report, 2014.

[8] HINES, J. W., GARVEY, D., SEIBERT, R., and USYNIN, A.: Technical review of on-line monitoring techniques for performance assessment. (NUREG/CR-6895). Nuclear Regulatory Commission.2007.

[9] MESSAI, A., MELLIT, A., ABDELLANI, I., and MASSI, P. A.: On-line fault detection of a fuel rod temperature measurement sensor in a nuclear reactor core using ANNs. Progress in Nuclear Energy.2015, 79:8-21.

[10] DESHPANDE, P., WARKE, N., KHANDARE, P., and DESHPANDE, V.: Thermal power plant analysis using artificial neural network. Proceedings of 20th ICONE, 2012.

[11] AJAMI, A., and DANESHVAR, M.: Data driven approach for fault detection and diagnosis of turbine in thermal power plant using Independent Component Analysis(ICA). Electrical Power and Energy Systems. 2012, 43:728-735.

[12] SALAHSHOOR, K., KORDENSTANI, M., and KHOSHORO, M. S.: Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. Energy.2010, 35:5472-5482.

[13] KAIYING, C., LONGSHENG, C., MUCHEN, C., and LEE, C. L.: Using SVM based method for equipment fault detection in a thermal power plant. Computer in Industry. 2011, 62:42-50.

[14] FARHAN, J., MUHANMMAD, A., INAMUL, H., ABDUL, Q. K., and MASOOD, I.: Fault diagnosis of Pakistan Research Reactor-2 with data-driven techniques. Annals of Nuclear Energy, 2016, 90:433-440.

[15] DANESHVAR, M., and RAD, B. F.: Data driven approch for fault detection and diagnosis of boiler system in coal fired power plant using principal component analysis. International Review of Automatic Control. 2010, 3(2):198-208.

[16] ROSANI, M. L. P., and HINES, J. W.: Using principal component analysis modeling to monitor temperature sensors in a nuclear research reactor [www.engr.utk.edu].

[17] SHUN, L., and JIN, W.: A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform. Energy and Buildings. 2014, 68:63-71.

[18] YUNPENG, H., GUANNAN L., HUANXIN C., HAORONG L., and JIANYAN, L.: Sensitivity analysis for PCA-based chiller sensor fault detection. International Journal of Refrigeration. 2016, 63:133-143.

[19] JIANPING, M., and JIN, J.: Applications of fault detection and diagnosis methods in nuclear power plants: A review. Progress in Nuclear Energy. 2011,(53): 255-266.

[20] XU, X., HINES, J. W., and UHRIG, R. E.: Sensor validation and fault detection using neural networks. Proceedings of Maintenance and Reliability Conference.(MARCON), University of Tennessee, 1999.

[21] PERLA, R., MUKHOPADHYAY, S., SAMANTA, A. N.: Sensor fault detection and isolation using neural networks. Proceedings of TENCO 2004 IEEE Rdgion10 Conference. 2004, D:676-679.

[22] ANDREW, K., and ZHE, S.: Sensor fault detection in power plants. Journal of Energy Engineering. 2009,135(4):127-137.

[23] LIU Y. K., PENG, M. J., XIE, C. L., and YAXIN, D.: Research and design of distributed fault diagnosis system in nuclear power plant. Progress in Nuclear Energy. 2013, 68:97:110.

[24] JIANG, L.: Sensor fault detection and isolation using system dynamics identification techniques. The University of Michigan. PhD thesis, 2011.

[25] HINES, J. W., and SEIBERT, R.: Technical review of on-line monitoring techniques for performance assessment: State-of-the-Art. Nuclear Regulatory Commission. 2006, NUREG/CR-6895, Volume 1.

[26] LI, F., UPADHYAYA B. R., and SERGIO R. P. .: Fault diagnosis of helical coil steam generator systems of an integral pressurized water reactor using optimal sensor selection. IEEE Transactions on Nuclear Science. 2012, 59(2):403-410.

[27] LI, R. Y.: Research on statistical process monitoring based on PCA. Zhejiang University: Control Theory and Control Engineering, 2007.

[28] HE, X. Q.: Research of principal component analysis and its application in data reconstruction of fault sensor. Northeastern University: Control Theory and Control Engineering, 2010.

[29] JOSE, C., ALEJANDRO, PEREZ, V., PEDRO, G. T., and GABRIEL, M. F.: PCA-based multivariate statistical network monitoring for anomaly detection. Computer & Security. 2016, 59:118-137.

[30] VALLE, S., WEIHUA, L., and QIN, S. J.: Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. Industrial & Engineering Chemistry Research. 1999, 38:4389-4401.

[31] CHUDONG, T., TING, L., and XUHUA, S.: Fault detection and diagnosis of dynamic processes using weighted dynamic decentralized PCA approach. Chemometrics and Intelligent Laboratory Systems. 2017, 161:34-42.

[32] HESSAM, M., and TAREK, E.: Nonlinear reduction of combustion composition space with kernel principal component analysis. Combustion and Flame. 2014,161:118-126.

[33] LI, F.: Dynamic modeling, sensor placement design, and fault diagnosis of nuclear desalination systems. The University of Tennessee, 2011.

[34] TAO, C, MARTIN, E., and GARY, M.: Robust probabilistic PCA with missing data and contribution analysis for outlier detection. Computational Statistics and Data Analysis. 2009, 53:3706-3716.

[35] CHEN, X.: Research on data preprocess method for thermal parameters. North China Electric Power University: School of Control and Computer Engineering, 2013.

[36] FANTONI, P. F.: Experiences and application of PEANO for online monitoring in power plants. Progress in Nuclear Energy. 2005, 46(2):206-225.