Sādhanā Vol. 42, No. 9, September 2017, pp. 1473–1480
DOI 10.1007/s12046-017-0700-2

Nuclear power plant sensor fault detection using singular value decomposition-based method

SHYAMAPADA MANDAL[1,*], N SAIRAM[1], S SRIDHAR[2] and P SWAMINATHAN[3]

[1]School of Computing, SASTRA University, Thanjavur 613401, India
[2]Indira Gandhi Centre for Atomic Research, Kalpakkam 603102, India
[3]School of Engineering and School of Computing Sciences, Vels University, Chennai 600117, India
e-mail: shyamapadamandal@sastra.ac.in; sairam@cse.sastra.edu; sethu@igcar.gov.in; dean.se@velsuniv.ac.in

MS received 21 September 2015; revised 25 February 2017; accepted 2 March 2017; published 27 July 2017

Abstract. In a nuclear power plant, periodic sensor calibration is necessary to ensure the correctness of measurements. Those sensors which have gone out of calibration can lead to malfunction of the plant, possibly causing a loss in revenue or damage to equipment. Continuous sensor status monitoring is desirable to assure smooth running of the plant and reduce maintenance costs associated with unnecessary manual sensor calibrations. In this paper, a method is proposed to detect and identify any degradation of sensor performance. The validation process consists of two steps: (i) residual generation and (ii) fault detection by residual evaluation. Singular value decomposition (SVD) and Euclidean distance (ED) methods are used to generate the residual and evaluate the fault on the residual space, respectively. This paper claims that SVD-based fault detection method is better than the well-known principal component analysis-based method. The method is validated using data from fast breeder test reactor.

Keywords. Sensor fault; fast breeder test reactor; singular value decomposition; Euclidean distance.

1. Introduction

A nuclear power plant (NPP) has a large number of sensors to monitor different parameters such as temperature, pressure and flow rate of the process fluid. As the operational decisions depend on sensor signals, it is necessary that the signal produced by the sensor must be faithful. Sensors will undergo physical degradation due to ageing, which results in their readings deviating from the actual value or the calibration curve. Some early indications of sensor fault can take the form of a lagging response caused by the increase of the sensor time constant. Other early signs of a sensor fault may consist of occasional inconsistent output due to loose sensor component contacts. A conventional method of maintenance for the sensors showing signals out of the allowable range consists of off-line integrity evaluation and recalibration or replacement. This approach does not result in timely detection of sensor degradation because inspection has to wait for their scheduled process to shutdowns. The objective of online sensor monitoring is to detect early indications of a sensor fault, thus enabling predictive maintenance.

1.1 *Related work*

Several techniques have been suggested for sensor fault detection and diagnosis (FDD) in the literature. These methods can be broadly classified into two categories: model-based and data-driven methods. In model-based methods, faults are detected and isolated by residual generation and evaluation. The model describes the mathematical and systematic characteristics of the process. In these methods, fault is detected by comparing the actual output with information obtained from the model. Kalman filter (KF) model-based method was used for fault detection and isolation (FDI) in a dynamic system [1]. It has a systematic design, noise disposal and enhances sensitivity to produce the effective results. Zarei and Poshtan [2] proposed a sensor FDI scheme based on the Luenberger Observer method. Gertler [3] established a dynamic parity relation to detect and isolate the system faults systematically and integrally. Bayesian Belief Network (BBN) and multi-stage BBN models were developed for fault detection and diagnosis (FDD) in a transient or steady-state system [4]. The success of the aforementioned methods depends on the fidelity of the system or the component model expressed in a mathematical form. However, for some complex processes, it is very difficult to obtain a highly accurate mathematical model.

*For correspondence

On the other hand, the data-driven methods are developed using historical data. Well-known data-driven methods for sensor malfunction detection are principal component analysis (PCA) and its extensions. Wang and Cui [5] have developed a method for sensor FDD using PCA. They utilised Q-statistic and Q-contribution plot to detect the fault. A non-linear approach of PCA was proposed by Huang *et al* [6] for sensor and actuator fault in non-linear systems. Wang and Xiao [7] have invented a technique for air handling unit (AHU) system sensor data validation using PCA, where Q-statistic or squared prediction error was used to set the threshold and the fault is fixed by Q-contribution plot. However, PCA is not aware of the pattern. It finds the relation between the variables, and that may produce false alarms. Several ANN approaches were developed to design sensor FDD in different industrial control systems. The ANN was used to predict the correlated variable in NPPs for sensor fault detection and plant monitoring [8]. Auto-associative neural network (AANN) was proposed for engine control system's sensor fault diagnosis and reconstruction [9]. NN-based classification scheme was used to recognise the sensor drift in Gas Turbine systems [10]. A dual NN strategy was designed in AHUs for sensor fault detection and its efficiency analysis [11]. Cascade NN (CNN) was used for sensor fault detection and identification [12]. The NN and KF schemes for sensor validation were proposed for flight control systems [13]. The FDI technique was developed by integrating two successful data-driven methods, PCA and NN by Zhou *et al* [14]. The combination of wavelet and fractal analysis with NN scheme was designed by Zhu *et al* [15] in the AHU system sensor FDD. A hybrid of data-driven soft measurement and modeling method was proposed for power plant sensor condition monitoring [16]. This method includes generalised regression neural network (GRNN), mean impact value (MIV), partial least squares regression (PLSR) and B-Spline transformation techniques. Hou *et al* [17] designed a method utilising data mining principle for sensor validation in air-conditioner systems using rough set (RS) and ANN. Training the network using the neural network and its combination method was time-consuming.

1.2 *Motivation and contribution of this paper*

In data-driven methods, sensor fault detection is based on calculating metrics that are used to detect the fault. Numerous metrics are used to identify the fault along with the PCA method that includes the Hotelling T^2-statistic [18] and the Q-statistic [19]. The T^2-statistic is used to estimate the maximum variation captured by the principal components, whereas the Q-statistic is used to calculate the deviation of the residual that are not captured by the principal components. Wang and Chen [20] noticed that the T^2-statistic value exceeds the threshold in nominal process condition and produces false alarms. As pointed out by Romagnoli and Palazoglu [19], the T^2-statistic will not detect the actual fault because the compact space is sometimes unable to capture small variation in the data. The Q-statistic will detect the fault by computing the square of the error between the actual data and its redundancy [21]. So it is more sensitive than T^2-statistic. Again, the performance of Q-statistic depends on the redundancy of the data. The data redundancy or reconstruction is based on the choice of the number of principal components. If the principal components are under- or over-estimated, the Q-statistic creates false alarms. The Q-statistic is seriously affected if the eigenvalues of the covariance matrix are roughly equal. It produces an effective result if the first eigenvalue is large and the remaining are very small when the eigenvalues are sorted in decreasing order. Thus, the Q-statistic will not always produce good results.

Recently, Jha and Yadav [22] reported that the SVD is more effective than the PCA for denoising the signals. In this work, the SVD method is used to overcome the PCA reconstruction problem. The SVD reconstruction is not affected by the type of eigenvalues, because in the SVD, the first singular value is always very large (for sensor data) and others are very small. The fault detection metric is applied to the difference between the training and the testing residual, but not directly to the residual. The fault is detected if the distance between the mean of the training and the testing residual crosses the threshold value. The Euclidean distance and the Mahalanobis distance are applied to estimate the distance between two vectors or matrices. If the variable variation is taken into account, the Mahalanobis distance is better than the Euclidean distance. Each sensor variation is approximately same in the fast breeder test reactor (FBTR). So, the Euclidean distance is proposed as a fault detection metric.

This paper is organised into five sections including the present one. The next section presents the proposed SVD-based method for sensor fault detection. The brief description about the FBTR is given in section 3. Section 4 explains the PCA-based fault detection method using Hotelling T^2-statistic and Q-statistic. Results obtained with the proposed method are compared with the existing PCA-based method and elaborated in section 5. The final conclusion of the paper and recommendations for future research work are given in section 6.

2. Proposed method

The SVD is an important tool in signal processing and statistical analysis. An important feature of SVD is its feature extraction capability. In this paper, extraction property of SVD is applied to detect a sensor fault in NPP. The proposed method consists of two steps: (i) residual generation and (ii) residual evaluation. Residual is generated using the SVD tool by projecting the original data into a new base set. The deviation of approximate of data from the original is the residual. The fault is detected by computing the distance between training and test residual. The Euclidean distance is imposed to calculate the distance between two matrices. The block diagram of the proposed method is given in figure 1.

2.1 *Residual generation by SVD*

SVD is an effective tool for feature extraction and compression. It factorises the given matrix into a singular value and singular vector matrices. Any matrix, $\boldsymbol{X} \in R^{m\times n}$, where $m > n$, can be decomposed as

$$\boldsymbol{X} = \boldsymbol{U\Sigma V}^T \tag{1}$$

where U is an $m \times m$ left singular vector, $\boldsymbol{V}$ is an $n \times n$ right singular vector and $\boldsymbol{\Sigma}$ is an $m \times n$ diagonal matrix with singular values in descending order i.e. $\lambda_{11}\text{-}> \lambda_{22} > \cdots > \lambda_{nn}$. The left and right singular vectors are the eigenvectors of $\boldsymbol{XX}^T$ and $\boldsymbol{X}^T\boldsymbol{X}$, respectively. The singular vectors are orthogonal. If the $\boldsymbol{\Sigma}$ is written as

$$\begin{aligned}\boldsymbol{\Sigma} &= \begin{bmatrix} \lambda_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \lambda_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_{nn} \end{bmatrix}\end{aligned} \tag{2}$$

then Eq. (1) can be re-written as

$$\boldsymbol{X} = \sum_{i=1}^{n} u_{\mathrm{m}i}\lambda_{ii}v_{ni}^T. \tag{3}$$

The SVD approach operates by projecting the original data onto a new basis, which captures the original features. By Eq. (3), the data can be reconstructed by selecting important singular values. The goodness of the SVD-based fault detection depends on an accurate selection of principal singular values. Over- and underestimation of the number of singular values can initiate noise that disguises the important features and omits important variations in the data, which degrades the reconstruction by SVD. So, it is important to choose appropriate principal singular values. Like PCA, important singular values can be selected by the cumulative percent variance [23], parallel analysis, sequential tests [24], etc. In this paper, principal singular values are selected by the cumulative percent variance that captured over 90% of the cumulative sum of the eigenvalues.

If the number of principal singular value k is determined, the dimension of the singular vectors and singular value matrices are truncated into k dimension as $\boldsymbol{U}_{m\times k} \in R^{m\times k}, \boldsymbol{V}_{n\times k} \in R^{n\times k}$ and $\boldsymbol{\Sigma}_{k\times k} \in R^{k\times k}$.

The data can be reconstructed by multiplication of the truncated matrices:

$$\bar{\boldsymbol{X}}_{m\times n} = \boldsymbol{U}_{m\times k} \times \boldsymbol{\Sigma}_{k\times k} \times \boldsymbol{V}^T_{n\times k}. \tag{4}$$

Residual or error is generated by the difference of the original data X and reconstructed data $\bar{\boldsymbol{X}}$, as

$$\boldsymbol{E} = |\boldsymbol{X} - \bar{\boldsymbol{X}}|. \tag{5}$$

2.2 *Residual evaluations*

The main objective of data reconstruction is to remove noise or outlier and approximate to a normal data. A fault-free data set is used to train the model, therefore, the elements of the training error matrix are very small (near to zero). If the test data matrix has any anomaly (anomalies), then there is a deviation between actual and reconstructed data, and the elements of the test residual matrix are high compared to the training residual matrix. The faults or anomalies are detected by calculating element-wise distance between the test residual and the mean of the training residual. If the distance exceeds the threshold value, then there exists a fault. Statistical techniques are applied to calculate the element-wise distance between two matrices. The Euclidean distance [25] between the points $x = (x_1, \ldots, x_p)^T$ and $y = (y_1, \ldots, y_p)^T$ in the p-dimensional space is defined as

$$\begin{aligned} d_{ED}(x,y) &= \sqrt{(x_1 - y_1)^2 + \cdots + (x_p - y_p)^2} \\ &= \sqrt{(x-y)^T(x-y)} \end{aligned} \tag{6}$$

The Mahalanobis distance [25] between two points $x = (x_1, \ldots, x_p)^T$ and $y = (y_1, \ldots, y_p)^T$ in the p-dimensional space is defined as

$$d_{MD}(x,y) = \sqrt{(x-y)^T S^{-1}(x-y)} \tag{7}$$

where S is the covariance matrix of X.

Both the methods are efficient for calculating the distance between two points, but the Mahalanobis distance employs covariance. If the covariance is one, then both are identical, and if the covariance or variance varies from variable to variable, then the Mahalanobis distance will be effective. In FBTR, each sensor variation is approximately the same, and hence the Euclidean distance is applied to compute the distance between the test error matrix and the mean of the training error matrix.

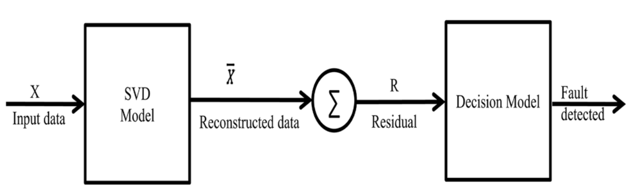

Figure 1. SVD model for sensor fault detection.

3. Fast breeder test reactor

The FBTR uses plutonium-uranium mixed carbide as fuel and liquid sodium as a coolant. The entire system is broadly grouped into three systems: primary sodium, secondary sodium, and steam and water circuit. The important components of the primary sodium system are reactor assembly, two intermediate heat exchangers (IHX), two sodium pumps and interconnecting piping. The secondary system includes sodium pumps, re-heaters, surge tanks, steam generator and connecting piping. Heat generated in the fuel sub-assemblies is removed by circulating liquid sodium through the reactor core. Two centrifugal pumps are used to pump sodium through the fuel sub-assemblies in the reactor core. Three thermocouples are used to measure the sodium temperature at the inlet of the reactor core. The central fuel sub-assembly contains four thermocouples (Tna000X, Tna000Y, Tna000Z and Tna000W) at the outlet and the rest of each 84 fuel sub-assemblies contain two thermocouples (Tna00nX, Tna00nY, for $n = 1$ to 84) at the outlet. Cromel-Alumel-type thermocouples are used to measure the temperature of sodium at the inlet of the reactor core and at the outlet of the fuel sub-assemblies. Eddy current flow meters are used to measure the primary sodium flow. The schematic description of fast breeder reactor is given in figure 2.

4. Fault detection using PCA-based method

The PCA is a widely used statistical tool for fault detection [19, 20]. In PCA method, the fault is detected by selecting the principal components that capture maximum variations of data or by computing the intensity of the error between data and its approximation. Let $\boldsymbol{X}$ be a data set matrix with dimension $m \times n$, where m is the number of observation and n is the number of variables. The PCA transforms the data matrix $\boldsymbol{X}$ into optimal vector space that captures the maximum variation of the data as follows:

$$\boldsymbol{S} = \boldsymbol{X}\hat{\mathbf{V}} \tag{8}$$

where $\boldsymbol{S} = [s_1, s_2, \ldots, s_l] \in R^{m \times l}$ is called the score vector or principal component vector and $\hat{\boldsymbol{V}} = [v_1, v_2, \ldots, v_l] \in R^{n \times l}$ are the eigenvectors (called the loading vectors) corresponding to l largest eigenvalues computed by singular value decomposition from the covariance matrix $\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, of the data matrix after scaling to zero mean and unit variance. Here $\boldsymbol{U}$ and $\boldsymbol{V}$ are the ortho-normal and $\boldsymbol{\Sigma}$ is the diagonal matrix of eigenvalues in decreasing order i.e. $\lambda_{11} > \lambda_{22} > \cdots > \lambda_{nn}$.

The data X can be expressed by the PCA approximation method as

$$\boldsymbol{X} = \bar{\boldsymbol{X}} + \boldsymbol{E} = \boldsymbol{S}\hat{\boldsymbol{V}}^T + \boldsymbol{X}\left(I - \hat{\boldsymbol{V}}\hat{\boldsymbol{V}}^T\right) \tag{9}$$

where $\bar{\boldsymbol{X}}$ is the approximation of $\boldsymbol{X}$ and $\boldsymbol{E}$ is the error. The elements of the residual matrix of the fault-free data are very small, but in the case of faulty data it is significantly larger. An important task in PCA method is to select l largest eigenvalues where the respective eigenvectors or loading vectors find the significant direction of the variation of the variables.

4.1 *Fault detection metrics*

In PCA-based fault detection method, Hotelling T^2 statistic and Q-statistic are applied as fault detection metrics. The PCA model is trained by fault-free data and the fault detection metric (threshold) is estimated. A brief description of Hotelling T^2-statistics and Q-statistics are given below.

4.1a *Hotelling's T^2 statistics*: The T^2-statistic measures the variations in the principal components of different observations in the time space as

$$T^2 = x^T \boldsymbol{V}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{V}^T x \tag{10}$$

where $\hat{\Sigma}$ is the diagonal matrix with eigenvalues. For each new test data, the T^2-statistic is calculated and a threshold, T_α is calculated as

$$T_\alpha = \frac{l(M-1)}{M-l} F(l, M-l, \alpha), \tag{11}$$

where $F(l, M-l, \alpha)$ is the F-distribution with $(l, M-l)$ degrees of freedom with significance level α, M is the number of observations of the training data and l is the number of selected principal components. If the T^2-statistic exceeds T_α, then the fault is considered as detected.

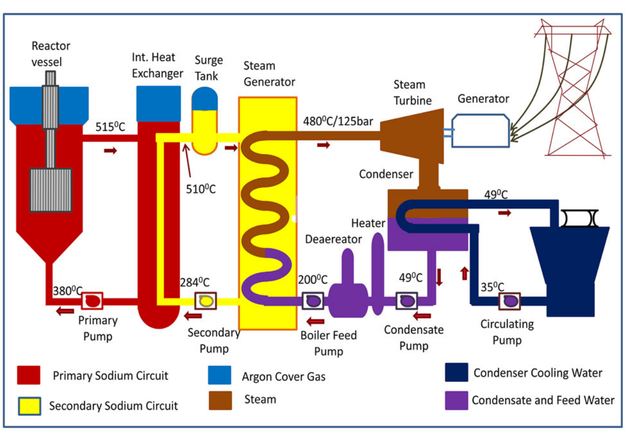

Figure 2. Schematic flow diagram of the main heat transport system in FBTR.

Table 1. Different sensors in primary sodium circuit.

Sensor application region	Sensor name	No. of sensors
Inlet sodium temperature	Cromel-Alumel thermocouple	3
Outlet sodium temperature	Cromel-Alumel thermocouple	172
Primary sodium flow	Eddy current flow meter	6

4.1b *Q-statistics*: The Q-statistic measures the square of error not captured by principal components in approximation. It is defined as

$$Q = \|E\|^2 = \left|\left|x\left(\boldsymbol{I} - \hat{\boldsymbol{V}}\hat{\boldsymbol{V}}^T\right)\right|\right|^2 \tag{12}$$

Whenever a new test data is available, the Q-statistic value is estimated and it is compared to a threshold Q_α defined as

$$Q_\alpha = \theta_1 \left[\frac{c_\alpha h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \tag{13}$$

where c_α is the standard normal distribution with α significance level, θ_i and h_0 are defined as follows:

$$\theta_i = \sum_{j=l+1}^{N} \lambda_j^i, \quad h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$$

where N is the number of variables. If the Q-statistic exceeds Q_α, the fault is detected.

5. Result and discussion

This section presents the results for sensor fault detection in the primary sodium circuit system of the nuclear reactor using the proposed SVD method and the existing PCA method. In the primary sodium system, 181 sensors are used and the details are given in table 1. Chromel-Alumel Thermocouples are used for measurement of inlet and outlet sodium temperature and Eddy flow meter is used for measurement of primary sodium flow into the reactor. There are three inlet thermocouples that are located in east, west and middle. The outlet centre fuel sub-assembly has four thermocouples and the remaining 84 sub-assemblies each have two thermocouples.

The results of the proposed method are compared with those of the existing PCA method using the FBTR data samples given in table 2. The sample data have only 14 rows and 5 columns. The actual data may be having 929 rows and 181 columns.

The SVD data approximation is closer to the original than the PCA, because in the PCA, the selected principal components, i.e. the principal eigenvectors, are unable to capture all the important features of the data matrix. The eigenvector corresponding to the smaller eigenvalues carries some important features of the data. The approximation is good if the first eigenvalue is large and the remaining are very small when the eigenvalues are sorted in decreasing order. The approximation is not closer to the original if the eigenvalues are roughly equal. The eigenvalues of the primary sodium circuit's sensor data are of the same order of magnitude, as shown in figure 3. For the SVD method, the first singular value is always very large for the sensor data, and the remaining are negligible. So, the first singular vector captured all the important features of the data matrix. The SVD singular values of the primary sodium circuit's sensor data are shown in figure 4. The reconstruction and residues generated by the SVD and the PCA are shown in figures 5–8.

The Euclidean distance metric is effective for primary sodium circuit sensor fault detection, because the operator has the freedom to set the threshold value by the characteristic of the sensors. In Q-statistic, the fault propagates through the model error, the threshold value Q_α is

Table 2. Sample data of the primary sodium system sensors in FBTR.

Date and time	TNA000W	TNA000X	TNA000Y	TNA000Z	TNA001X
05-06-2012#09:53:41	185.67	184.92	184.04	185.04	184.04
05-06-2012#09:53:43	185.78	184.66	183.91	185.03	183.91
05-06-2012#09:53:45	185.81	184.69	184.19	185.19	183.94
05-06-2012#09:53:47	185.42	184.67	184.04	185.17	183.92
05-06-2012#09:53:49	185.67	184.42	183.92	185.04	183.92
05-06-2012#09:53:51	184.92	184.79	184.04	185.04	183.92
05-06-2012#09:53:53	185.79	184.17	184.04	185.17	183.92
05-06-2012#09:53:55	185.66	184.66	184.16	184.91	183.78
05-06-2012#09:53:57	185.79	184.79	184.04	185.17	183.92
05-06-2012#09:53:59	185.78	184.53	183.78	185.16	183.91
05-06-2012#09:54:01	185.53	184.66	184.03	185.16	183.91
05-06-2012#09:54:03	185.54	184.67	184.04	185.04	183.79
05-06-2012#09:54:05	185.79	184.67	183.42	185.17	183.92
05-06-2012#09:54:07	185.28	184.78	183.91	184.91	183.78

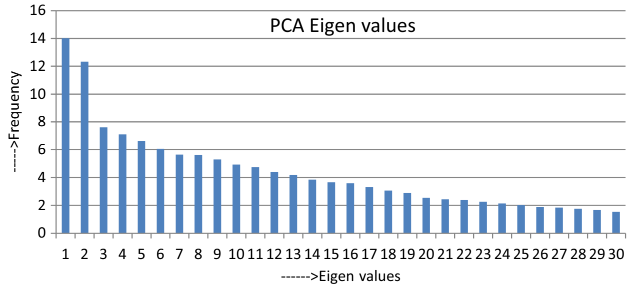

Figure 3. Eigenvalues of the primary sodium circuit sensor data.

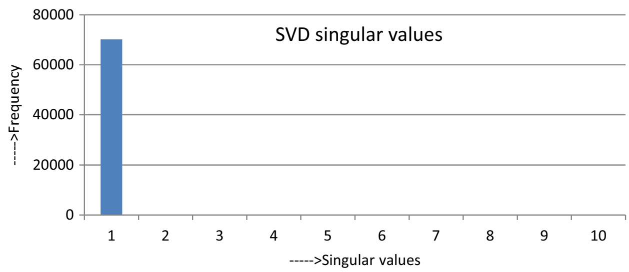

Figure 4. Singular values of the primary sodium circuit sensor data.

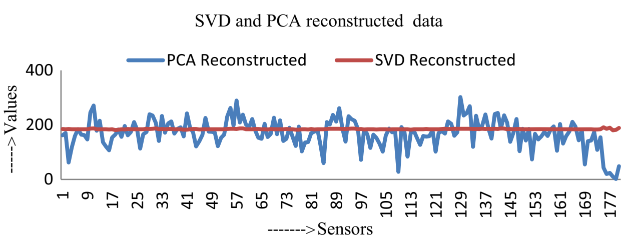

Figure 5. Comparison of reconstruction by SVD and PCA.

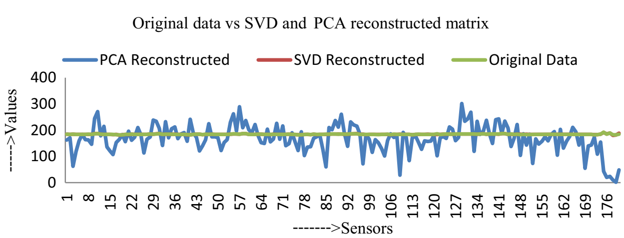

Figure 6. Comparison of reconstruction (by SVD and PCA) and original signals.

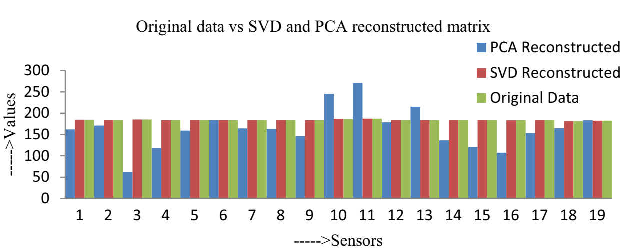

Figure 7. Histogram of reconstructed data by the SVD and the PCA along with original data.

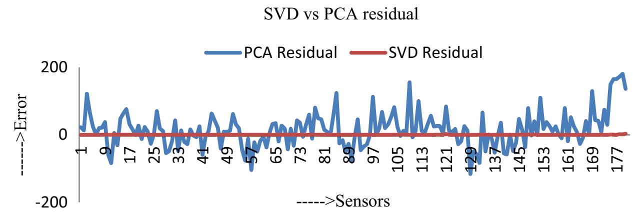

Figure 8. Comparison of the residual of a signal by the SVD and the PCA.

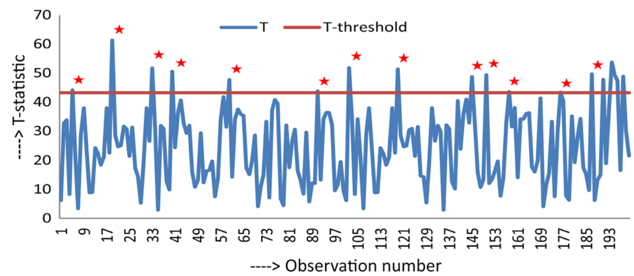

Figure 9. Fault detection using the PCA and T^2-statistic.

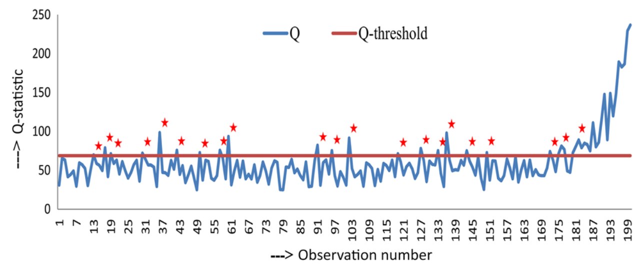

Figure 10. Fault detection using the PCA and Q-statistic.

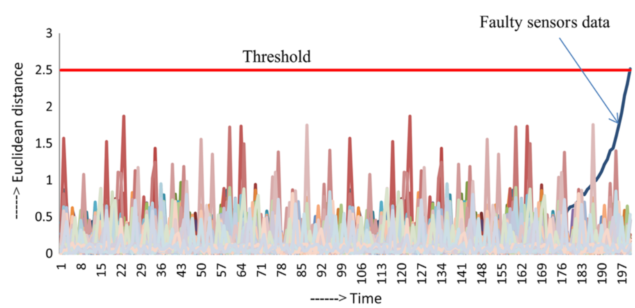

Figure 11. Fault detection using the SVD and Euclidean distance.

dependent on the number of principal eigenvalues and hence the operator has no freedom to set the threshold value. The variation of T-statistic (Hotelling T^2-statistic), Q-statistic and the SVD-based Euclidean distance for the same data are shown in figures 9–11. The T^2-statistic is unable to detect the actual fault, but raises some false alarms. This is indicated by the red colour star symbol in figure 9. The Q-statistic is able to detect the actual fault along with some false alarms as shown by the red star symbol in figure 10. The SVD method along with the Euclidean distance detects the actual fault without false alarm.

6. Conclusion and future work

Online monitoring of the sensor's physical condition can avoid many problems associated with manual calibration of sensors. The SVD model is developed for the detection of sensor faults in nuclear power plants. SVD is a simple linear algebraic factorisation method. SVD is used to generate the residual matrix by selecting few largest singular values. The reconstruction matrix is closer to the original data. The Euclidean distance is employed in residual difference space to detect the error length. If the error length exceeds the threshold value, then a fault is detected. In a dynamic process, it is very flexible to change the threshold value. The proposed method will detect and identify the faulty sensor sooner. The accuracy of the fault detection is higher than all the other existing methods and will not produce any false/spurious alarm.

There are some limitations in the reconstruction of the data using the PCA and the SVD methods, and fault detection using these methods. The following are recommended for future work:

- The type of data for which the PCA eigenvalues are roughly equal and the type of the data for which the first eigenvalue is large and remaining are small should be obtained.
- A method to set the threshold value for fault detection from several variables with different variances using the Mahalanobis distance needs to evolve.

Acknowledgement

This research work is financially supported by the Board of Research in Nuclear Science (BRNS), India, under the research project bearing the Number: 2013/36/42-BRNS/2502. We would also like to thank the Indira Gandhi Centre for Atomic Research (IGCAR), Kalpakkam, India, for supplying the data.

References

[1] Huang S, Tan K K and Lee T H 2012 Fault diagnosis and fault-tolerant control in linear drives using the Kalman filter. *IEEE Trans. Ind. Electron*. 59: 4285–4292
[2] Zarei J and Poshtan J 2011 Sensor fault detection and diagnosis of a process using unknown input observer. *Math. Comput. Appl.* 16: 31–42
[3] Gertler J J 1997 Fault detection and isolation using parity relations. *Contr. Eng. Pract.* 5: 653–661
[4] Mehranbod N, Soroush M and Panjapornpon C 2005 A method of sensor fault detection and identification. *J. Process Control* 15: 321–339
[5] Wang S and Cui J 2005 Sensor-fault detection, diagnosis and estimation for centrifugal chiller systems using principal-component analysis method. *Appl. Energy* 82: 197–213
[6] Huang Y, Gertler J and McAvoy T J 2000 Sensor and actuator fault isolation by structured partial PCA with nonlinear extensions. *J. Process Control* 10: 459–469
[7] Wang S and Xiao F 2004 AHU sensor fault diagnosis using principal component analysis method. *Energy Build.* 36: 147–160
[8] Upadhyaya B L and Eryurek E 1992 Application of neural network for sensor validation and plant monitoring. *Nuclear Technol.* 97: 170–176
[9] Huang X 2004 Sensor fault diagnosis and reconstruction of engine control system based on autoassociative neural network. *Chinese J. Aeronaut.* 17: 23–27
[10] Palme T, Fast M and Thern M 2011 Gas turbine sensor validation through classification with artificial neural networks. *Appl. Energy* 88: 3898–3904
[11] Du Z, Fan B, Chi J and Jin X 2014 Sensor fault detection and its efficiency analysis in air handling unit using the combined neural networks. *Energy Build.* 72: 157–166
[12] Hussain S, Mokhtar M and Howe J M 2015 Sensor failure detection, identification, and accommodation using fully connected cascade neural network. *IEEE Trans. Ind. Electron*. 62: 1683–1692
[13] Napolitano M R, Windon D A, Casanova J L, Innocenti M and Silvestri G 1988 Kalman filters and neural-network schemes for sensor validation in flight control systems. *IEEE Trans. Control Syst. Technol*. 6: 596–611
[14] Zhou J, Guob A, Cellerc B and Su S 2014 Fault detection and identification spanning multiple processes by integrating PCA with neural network. *Appl. Soft Comput.* 14: 4–11
[15] Zhu Y, Jin X and Du Z 2012 Fault diagnosis for sensors in air handling unit based on neural network pre-processed by wavelet and fractal. *Energy Build.* 44: 7–16
[16] Chen J, Li H, Sheng D and Li W 2015 A hybrid data-driven modeling method on sensor condition monitoring and fault diagnosis for power plants. *Electr. Power Energy Syst.* 71: 274–284
[17] Hou Z, Lian Z, Yao Y and Yuan X 2006 Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Convers. Manag.* 47: 2479–2490
[18] Hotelling H 1933 Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* 24: 417–441
[19] Romagnoli J A and Palazoglu A 2006 *Introduction to process control*, 2nd edn. Boca Raton, FL: CRC Press
[20] Wang S and Chen Y 2004 Sensor validation and reconstruction for building central chilling systems based on principal component analysis. *Energy Convers. Manag.* 45: 673–695
[21] Harkat M, Mourot G and Ragot J 2006 An improved PCA scheme for sensor FDI: application to an air quality monitoring network. *J. Process Control* 16: 625–634

[22] Jha S K and Yadav R D S 2011 Denoising by singular value decomposition and its application to electronic nose data processing. *IEEE Sensor J.* 11: 35–44
[23] Diana G and Tommasi C 2002 Cross-validation methods in principal component analysis: a comparison. *Stat. Methods Appl.* 11: 71–82
[24] Jolliffe I 2002 *Principal component analysis*, 2nd edn. Berlin: Springer
[25] Deza M M and Deza E 2009 *Encyclopedia of distances*. Heidelberg, London, New York: Springer, pp. 94–95