E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 10, No. 1 (2023) 97–112.

Paper Type: Research Paper

A Data-Driven Design for Gas Turbine Exit Temperature Spread Condition Monitoring System

Nastaran Hajarian[1], Farzad Movahedi Sobhani[1,*], Seyed Jafar Sadjadi[1]

[1] Department of Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran; n.hajarian@srbiau.ac.ir; fmovahedi@iau.ac.ir.

[2] Department of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran; sjsadjadi@iust.ac.ir.

Citation:

Hajarian, N., & Movahedi Sobhani, F., & Sadjadi, S. J. (2023). A data-driven design for gas turbine exit temperature spread condition monitoring system. *Journal of applied research on industrial engineering, 10*(1), 97-112.

Received: 27/07/2021 Reviewed: 01/08/2021 Revised: 09/10/2021 Accepted: 01/12/2021

Abstract

One of the most complex and costly systems in the industry is the Gas Turbine (GT). Because of the complexity of these assets, various indicators have been used to monitor the health condition of different parts of the GT. Turbine Exit Temperature (TET) spread is one of the significant indicators that help monitor and detect faults such as overall engine deterioration and burner fault. The goal of this article is to use data-driven approaches to monitor TET data to detect faults early, as fault detection can have a significant impact on GT reliability and availability. In this study, the TET data of v94.2 GT is measured by six temperature transmitters to show a detailed profile. According to the statistical tests, TET data are high dimensional and time-dependent in the real world industry. Hence, three distinctive methods in the field of the GT are proposed in this study for early fault detection. Conventional Principal Component Analysis (PCA), Moving Window Principal Component Analysis (MWPCA), and Incremental Principal Component Analysis (IPCA) were implemented on TET data. According to the results, the conventional PCA model is a non-adaptive method, and the false alarm rate is high due to the incompatibility of this approach and the process. The MWPCA based on V-step-ahead and IPCA approaches overcame the non-stationary problem and reduced the false alarm rate. In fact, these approaches can distinguish between the normal time-varying and slow ramp fault processes. The results showed that IPCA could detect fault situations faster than MWPCA based on V-step-ahead in this study.

Keywords: Early fault detection, Data-driven, Gas turbine exit temperature, Time-varying, PCA model, MWPCA model, IPCA model.

1 | Introduction

As the demand for systems safety and product quality increases, the role of process monitoring in industrial procedures becomes more prominent [1]. One of the most complex and costly systems in the industry is the Gas Turbine (GT). There is increasing attention to condition monitoring of GTs in power plants operations and maintenance. The main reasons for this consideration are to enhance the reliability and availability of these valuable assets [2]. As a result, employing the monitoring strategy appears to be important to ensure that faulty conditions are detected early and that a forced plant shutdown is avoided [3]. Because of the complexity of these assets, various indicators have been used to monitor the health condition of different parts of GT, such as coast down-time, vibration, performance, maximum turbine outlet temperature at minimum fuel flow,

Corresponding Author: fmovahedi@iau.ac.ir

https://doi.org/10.22105/jarie.2021.296470.1361

Turbine Exit Temperature (TET) spread, Bearing temperature during coast down, etc. [4]. "TET spread" is one of the essential indicators that helps monitor and detect the presence of faults such as overall engine deterioration and burner fault.

There are several conceptual approaches to detecting and diagnosing failure, as well as distinct categories presented by different researchers. None of these classifications are comprehensive, and each researcher has his or her own viewpoint. Zhang et al. [5] argued that fault detection methods could be either model-based or data-based methods (*Fig. 1*). However, Chiang et al. [6] believed that three fault detecting and diagnosing methods are data-driven, analytical-based, and knowledge-based methods.

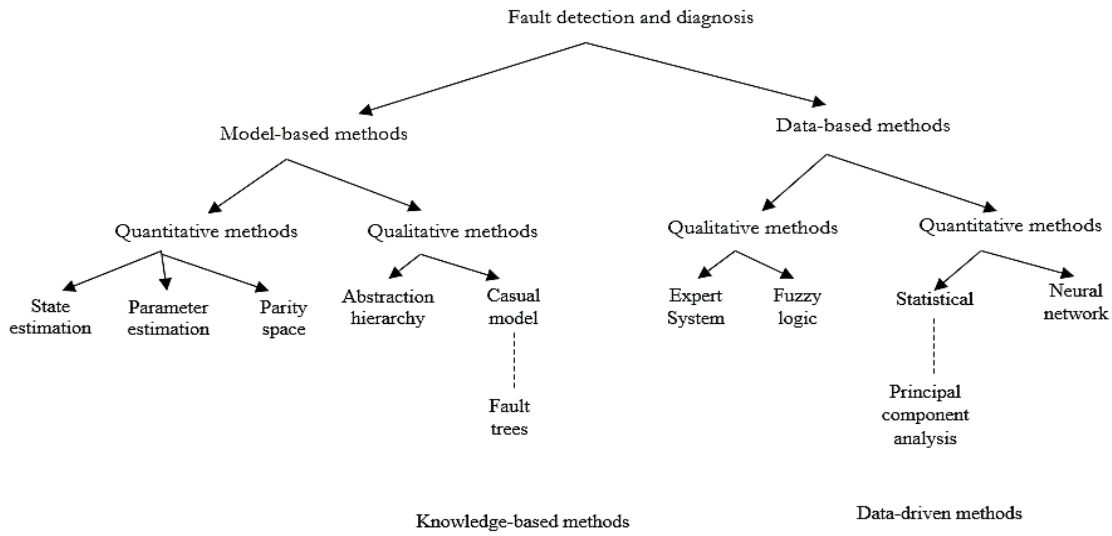

Fig. 1. Classification of fault detection and diagnosis method [5].

The analytical approach utilizes first principles to construct mathematical models of the system [7]. The analytical model cannot be applied for large-scale and complex systems. According to [8], the knowledge-based methods are mostly rule-based expert systems. The failure cases and engineers' experience are used to formulate the rules. When the detailed mathematical model cannot be reached, and when the number of inputs, outputs, and states of a system is logically small, the knowledge-based method will yield the best results [8]. The data-driven methods use product life-cycle data to detect faults and they do not depend on the first-principles models [9], [10].

Hence, data-driven methods can be used for large-scale and complex systems which are inexpensive and high accuracy [7], [11]. Upon the availability of the data of product or system, data-driven methods are chosen, but the system model is not [12]. As *Fig. 1* shows, data-driven methods consist of statistical and non-statistical methods [5]. One of the data-driven statistical tools is Principal Component Analysis (PCA) models. Instead of using base-models, we introduced PCA models for monitoring and identifying early failures in a GT using TET spread indicators in this study. Because the basic principles for constructing mathematical models of GTs are not readily available, and such a model is typically difficult to obtain due to the complexity and high dimensionality of a GT, there appears to be sufficient historical data for employing a statistical method.

Statistical Process Monitoring (SPM) developed the quality control charts to detect a system deviation from the normal behavior. When the number of variables or dimensions of a problem is beyond one value, it is necessary to use a multivariate statistical approach [13].

Monitoring, in a power plant context, means assessing the measured data in the simplest form to distinguish data of the normal operation from abnormal data. Nowadays, GTs are equipped with a lot of sensors, of which the collected data are used for monitoring purposes [14].

Hajarian et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

The idea of monitoring the TET spread monitoring has been proposed for more than two decades, and different methods have been suggested until now [4], [15]-[21]. Various research has been done in TET spread monitoring since 1992 [4], [15]-[21]. All of them mentioned the importance of TET monitoring to prevent catastrophic GT damages. In an ideal condition, the measurements of all thermocouples in exhaust must be the same. Nevertheless, in the real world, it is not. Knowles [17] described all possible reasons for this fact and suggested monitoring temperature patterns instead of temperature spread. In this method, a pattern is extracted from the healthy condition of the GT to compare the current TET pattern. Tsalavoutas et al. [21] presented a statistical method to evaluate the changes in temperature patterns. None of these researchers considered the influence of operation status and ambient conditions on their models.

Some researchers applied TET to diagnose a specific failure. Medina et al. [19] focused on combustion chambers failure detection with the aid of TET monitoring. They developed a TET model based on the basic principles of a GT, which estimates the TET, then compared the actual temperature measurement with the TET estimation to detect each combustor chamber failure. Korczewski [18] delivered a method to detect the failures of the automatic engine control system according to the alteration in TET during transient conditions such as start-up and acceleration processes. Besides, he presented diagnostic tolerances based on the statistical quality control methods. Jinfu et al. [16] presented an early fault detection method of the hot component. They introduced an indicator to detect the early faults of the hot component in the GT. Kenyon et al. [22] proposed an anomaly detection in exhaust gas temperature system by data mining algorithm to monitor the related parameters when anomalies are identified. Navi et al. [23] proposed partial kernel PCA for sensor fault detection of an industrial GT. Palmé et al. [20] used Auto Associative Kernel Regression method to simulate the TET pattern. In this method, the value of each thermocouple measurement would be estimated according to the previous records of exhaust temperature. Literature review shows that previous studies were often in the field of fault detection of the GT with an analytical approach, and there was no special attention to the important point that the trend of TET measured data is high dimensional and non-stationary. They are affected by operation status and ambient conditions.

According to our findings, no research has been conducted on early fault detection by monitoring TET spread using a statistical approach with non-stationary and high-dimensional data assumptions. This study aims to find an appropriate approach for early fault detection of turbine gas using TET spread, with a focus on data that will change over time, namely non-stationary data. In this paper, a low and straightforward calculation cost method will be proposed. It will help us better understand the condition of the TET pattern during the encounters of any fault generation and propagation. This study presents two approaches to cope with high-dimensional and time-dependent features. The Moving Window Principal Component Analysis (MWPCA) and Incremental Principal Component Analysis (IPCA) are two new approaches that can solve the dimensionality and time-dependent problems.

This paper also looks into the basic PCA, MWPCA, and IPCA in Section 2. In addition, Section 3 presents the details of the results of implementation and fault detection methods based on the conventional PCA, MWPCA, and IPCA. Finally, the conclusion is given in Section 4.

2 | The Multivariate Statistical Approach based on PCA, MWPCA, IPCA

Researchers believe that among the monitoring of multivariate processes, PCA, partial least squares, canonical correlation analysis and factor analysis are the suitable monitoring approaches [24].
The mentioned approaches and their expansion indicate the capability of conducting a high-dimensional data process. All turn the high-dimensional process into a lower-dimensional subspace and control the process behavior accordingly [25]. In the PCA, the original space variables, usually correlated, are transformed linearly into a new space of variables; these new variables are uncorrelated or orthogonal to each other [26]. If $X \in R^{n\times m}$ is the standardized data matrix with zero mean and unit variance with the scale parameter vectors x and S as the mean and variance vectors, respectively, where n denotes the sample

number and m represents the variable number. Defining the covariance matrix R is the first step to obtain PCA:

$$R = \frac{1}{n-1} X^T X. \tag{1}$$

Then accomplishing Singular-Value Decomposition (SVD) decomposition on R is as follows:

$$R = V \Lambda V^T, \tag{2}$$

where Λ represents a diagonal matrix consisting of the non-negative real eigenvalues as decreasing ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$). Columns of matrix V are the eigenvectors of R. Based on r principal eigenvalues, the transformation matrix $P \in R^{m \times r}$ is generated by choosing r eigenvectors or columns of V. The space of the measured variables is turned into the reduced dimension space by matrix P:

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

The columns of P are called loading and items of T are called scores. Scores are the values of the original measured variables that have been converted into the reduced dimension space [27]. According to *Eq. (3)*, the scores can also be turned into the original space as follows

$$\widehat{X} = TP^T. \tag{4}$$

The residual matrix E is calculated as:

$$E = X - \widehat{X}. \tag{5}$$

Finally, the original data space can be computed as

$$X = TP^T + E. \tag{6}$$

The important part of the implementation of PCA is choosing the number of principal components, which should be done carefully because TP^T illustrates the main source of variability and E represents the variability known as noise [28]. Several methods are proposed for choosing the number of principal components. SCREE procedure and Cumulative Percent Variance (CPV) approach are the most popular methods.

A plot of the eigenvalues is built in descending order by the SCREE procedure [28], which is a graphical approach that detects the knee in the curve. The number of knees on the plot indicates the count of principal components. CPV is the other method [29], [30] which determines the percent variance ($CPV(r \geq \%90)$ calculated by the first r principal components as follows [31]:

$$CPV\ r) = \frac{\sum_{i=1}^{r} \lambda_i}{trace(R)}. \tag{7}$$

2.1 | Multivariate Statistical Process Control based on PCA

After constructing a PCA model based on the historical data collected, it is necessary to have an instrument that controls variation. It is possible to plot the multivariate control charts using the Hotelling T^2 and Square Prediction Error (SPE) or Q to detect the fault. Determining two orthogonal subspaces of the original space can decrease the monitoring of these two variables (T^2 and Q).

The significant variation and the random noise in the data can be controlled by T^2 and Q, respectively. The T^2 statistic can be calculated for each new observation x by:

$$T^2 = x^T P \Lambda_r^{-1} P^T x, \tag{8}$$

where Λ_r is the squared matrix constructed by the first r rows and columns of Λ and, as previously mentioned, P is r eigenvectors or columns of V. The upper confidence limit for T^2 is acquired using the F-distribution:

$$T_\alpha^2 = \frac{r(n-1)}{n-r} F_{\alpha,r,n-r}, \tag{9}$$

where r is the number of the principal components and n is the number of samples in the data and α is the level of significance. A violation of the threshold would mean that variations of the system are out of control. Another statistic is the SPE or Q that can monitor the portion of the measurement space related to the lowest m-r eigenvalues. In fact, the Q statistic is calculated as the sum of squares of residuals.

$$Q = x^T\left(I - PP^T\right)x, \tag{10}$$

where *I* is the identity matrix.

The upper confidence limit for the Q can be calculated from its approximate distribution:

$$Q_\alpha = \theta_1[\frac{C_\alpha\sqrt{2\,\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0\; h_0 - 1)}{\theta_1^2} + 1]^{1/h_0}, \quad h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}, \theta_i = \sum_{j=r+1}^{m} \lambda_j^i, \tag{11}$$

where C_α is the value of the normal distribution with the α level significance. A violation of the threshold would indicate that an unusual event has occurred that had changed the covariance structure of the model [28].

Monitoring and fault detection based on the PCA model considers two steps:

1. Off-line: Acquire training data, collected under normal operation, this matrix must be normalized to zero mean and unit variance with the scale parameter vectors x and S as the mean and variance vectors, respectively. Then, we should implement PCA algorithm, find eigenvalue and eigenvector and determine the number of principal components and finally the upper control limits for T^2 and Q statistics.
2. Online:

   - *Get a new instance and scale it using the scale parameter vectors x and s.*
   - *Calculate $T^2$and Q statistics using the result of PCA model.*
   - *When the value of statistics is compared to thresholds, the violation is interpreted as an alarm.*
   - *Repeat from step A.*

The monitoring process with PCA is indicated in *Fig. 2*.

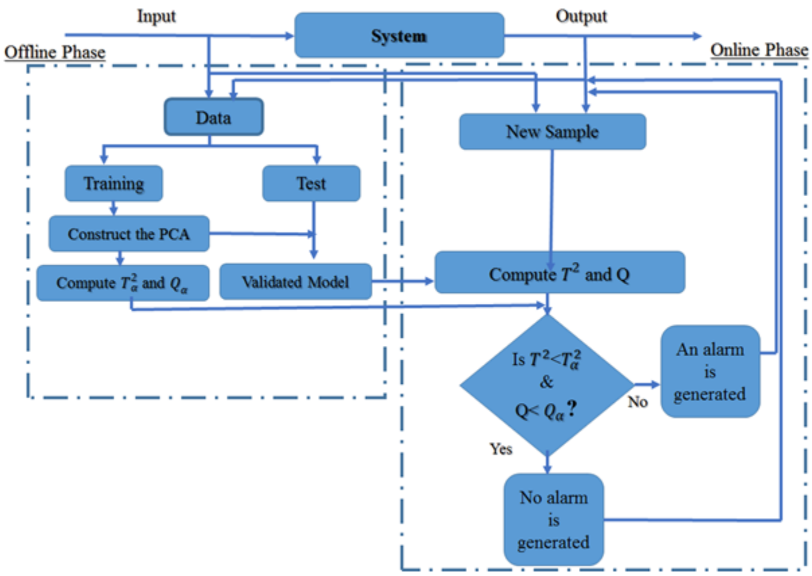

Fig. 2. Monitoring process with PCA method diagram [32].

2.2 | Multivariate Statistical Process Control Based on MWPCA

The process parameters, such as the mean or covariance, change with time, making the process non-stationary [33]. Several complementary MSPM methods have been introduced to tackle the time-varying issue. Three classes of approaches of Recursive Principal Component Analysis (RPCA), MWPCA and IPCA were applied to develop PCA methods to address non-stationary data [34].

Ketelaere et al. [35] investigated PCA-based statistical process-monitoring methods in terms of time-dependent and high-dimensional data, including MWPCA and RPCA [35]. RPCA techniques update the model for ever-increasing data consisting of new samples without discarding the old ones. Although RPCA is theoretically simple, it has been successfully employed for process monitoring. However, its implementation might not be easy for two main reasons: the ever-growing data set on which the model is updated, which eventually slows down the speed of adaptation as the data size increases. RPCA also consists of older data that are unrepresentative of the time-varying process. The forgetting factor cannot be easily selected without a priori knowledge of likely fault conditions when given to down-weight older samples [36]. The MWPCA method can tackle some of the limitations mentioned above by gathering a sufficient number of data points in the time-window to help build an adaptive process. Specifically, MWPCA removes older samples to choose the new samples representing the current operation process. Hence, for window size H, the data matrix at time t is $X_t=(x_{t-H+1}, x_{t-H+2}, \ldots, x_t)'$ and, at time t + 1, it is $X_{t+1}=(x_{t-H+2}, x_{t-H+3}, \ldots, x_{t+1})'$. The observations in the new window can be used to obtain the updated $\bar{x}_{t+1}$ and s_{t+1} [37]. The window includes a number of samples to cover enough process variation for modeling and monitoring purposes. Thus, window size is important. If a high number of samples select for the window, the MWPCA computation speed reduces drastically. If the result data used to enhance the computational efficiency is in a smaller window size, the relationship between the process variables will be important. If the model's adaptability to the process changes rapidly, it will be difficult to notice abnormal behavior, and the narrow window will be risky. Chiang et al. [6] determined the window size needed to estimate the T^2-statistic accurately. This was done according to the convergence of the T^2 distribution to the F distribution suggesting that minimum window sizes should be greater than approximately 10 times the number of variables. The monitoring process with MWPCA is indicated in *Fig. 3*. It is worth noting that over-fitting might be observed in MWPCA and a slow ramp could not be detected. As a result of the introduction of V-step-ahead prediction, MWPCA is based on the application delay. This approach is implemented using a model estimated at time t to predict the system behavior at time t + V and observe the likely faults. This step is used to ensure that the model does not over-adapt to the data and that it can detect faults that develop over time and are identified as regular observations at each time point [36], [38].

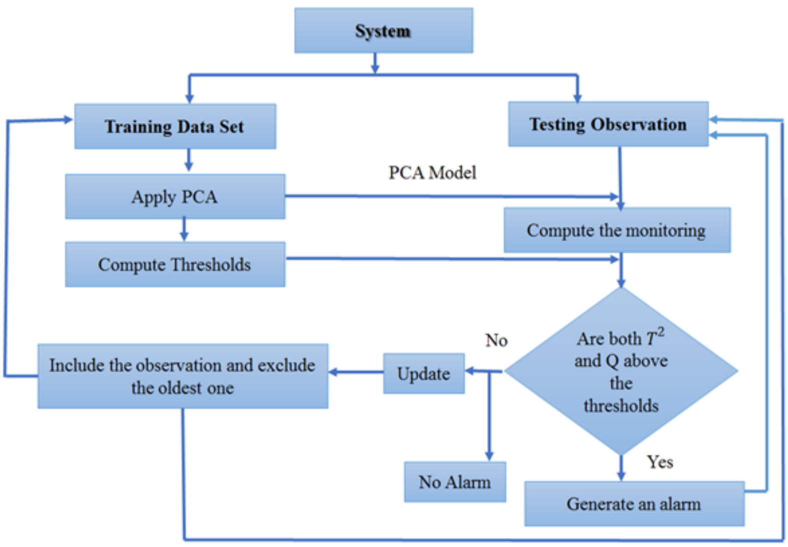

Fig. 3. Monitoring process with MWPCA method diagram [32].

2.3 | Multivariate Statistical Process Control based on IPCA

In adaptive methods, the monitoring models are updated by new sample and the normal time-varying information is added to the monitoring model to recognize between the normal time-varying and slow ramp fault processes. Another approach is IPCA.

Unlike the adaptive method, this method presents a novel approach in which the monitoring model is constant. When the PCA model implements time-varying process data, the PCs are also time-varying. This method introduces a new parameter as Incremental Principal Component (IPC). IPC calculate as follows:

$$\mathrm{IPC_i} = \mathrm{mean}\big(\mathrm{PC_i}\ \mathrm{k} - \mathrm{L}{:}\,\mathrm{k})\big) - \mathrm{PC_i}\ \mathrm{k} - \mathrm{W} - 2\mathrm{L}{:}\,\mathrm{k} - \mathrm{W} - \mathrm{L}). \tag{12}$$

The IPC is proposed to define the variations of the PCs. IPCs explain time-varying information. L indicates the number of PCs used to calculate the mean subtracted, and W indicates the interval of the two moments ($W > L$). A good introduction to IPCA can be found in [39]. IPCA model contains two steps as follows:

1. Offline:

   I. Acquire training data set A, which is collected under normal operation. This matrix must be normalized to zero mean and unit variance with the scale parameter vectors $\bar{x}$ and S as the mean and variance vectors, respectively.

   II. Construct the conventional PCA model with training data set, using the eigenvalue decomposition algorithm without reducing the dimension, and keep the loading matrix P and PCs.

   III. Compute the IPCs according to *Eq. (12)* in dataset A, and then compute the corresponding $IT^2$of each sample with the IPCs according to *Eq. (13)*.

$$\mathrm{IT_i^2} = \mathrm{IPC_i^T}\Lambda_{\mathrm{IPC}}^{-1}\mathrm{IPC_i}. \tag{13}$$

   IV. Use the Kernel Density Estimation (KDE) algorithm to define the control limit CL_j and CL of IT^2 as well as the threshold CL_j of IPC_j with the 99% confidence level.

2. Online step:

   I. Collect N_1 normal online samples, where $N_1 > W + 2 * L$. Normalize the samples through the means and variances of the training dataset A, and compute the PCs with the loading matrix P.

   II. Collect a new online sample. Normalize the sample with the means and variances of the training dataset A and calculate the PCs with the loading matrix P of the PCA model.

   III. Compute IPC_i of the i^{th} sample with the PCs of the forward W +2*L samples through *Eq. (12)*.

   IV. Compute the statistic $IT_i^2$of the i^{th}sample with the IPC_i of the i^{th} sample through *Eq. (13)*.

   V. If $IT_i^2 > CL$, then a fault may be present in the process. Otherwise, the sample is a normal one. In addition, when $IPC_{i,j}$>CL_j, the corresponding $PC_{i,j}$ is replaced by $PC_{i-i,j}$.

   VI. Repeat (II) – (III).

The monitoring process with IPCA is indicated in *Fig. 4*.

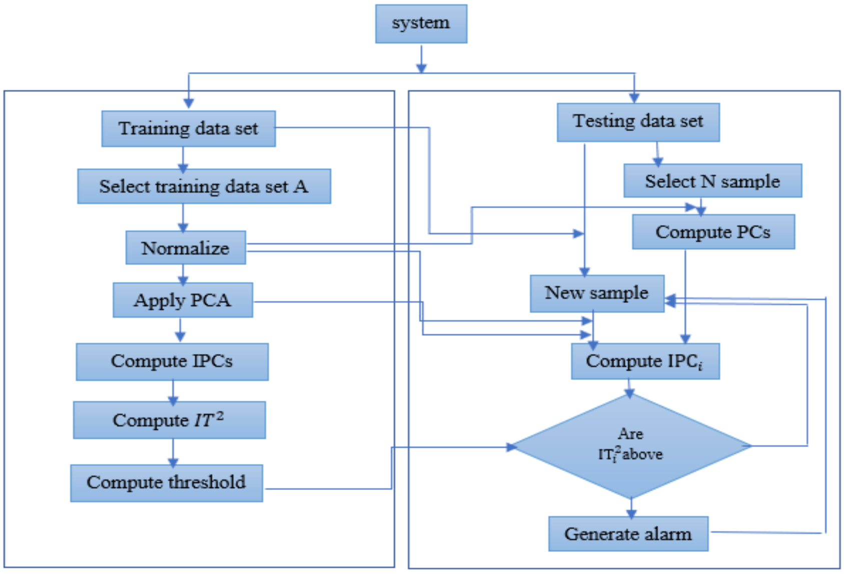

Fig. 4. Monitoring process IPCA method diagram.

3 | Static PCA, MWPCA Based on V-Step-Ahead and IPCA Applied to TET Spread

GTs are designed for many different purposes. In the industry, they are commonly used to drive compressors to transport gas through pipelines and generators that produce electrical power [40], [41]. In the past, the use of GTs has traditionally been limited to generating power during times of peak demand. Still, nowadays, they are being used in combined cycle power plants for baseload production [42]. Consequently, their availability, as well as reliability, play a significant role in these machines. The development of the GT in recent years has been facilitated most considerably by three factors:

I. Metallurgical developments that can be used to apply high temperatures in the combustor and turbine components.
II. Increased underlying knowledge of aerodynamics and thermodynamics.
III. Designing and simulating turbine airfoils and combustor and turbine blade cooling configurations by computer software.

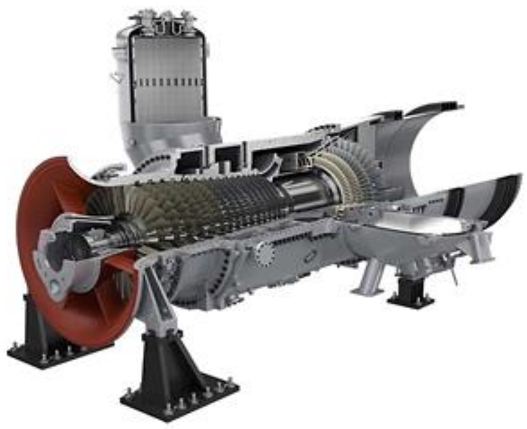

Fig. 5. Siemens V94.2 GT, from left to right: compressor, combustion chamber, turbine [43].

After compressing the air in the compressor, the fuel will be injected into it and combustion will increase the temperature of the gas. Turbine Inlet Temperature (TIT) is the average temperature of the flue gas that will face the first stage of turbine blades. During the expansion of the flue gas in the turbine, the pressure and temperature will decrease and the flue gas will leave the turbine with TET. The overview of a V94.2 GT manufactured by Siemens is shown in *Fig. 5*.

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

If TIT could be increased, GT's efficiency and specific power would improve. Nonetheless, there is a technological limit to designing and producing turbines that can withstand larger TIT. Since TIT is too hot to be measured directly, it is usually calculated by measuring TET. Both TET and TIT have a profile on their section due to the flue gas stream's rotation and turbulence. In v94.2 GTs, TET is being measured by six temperature transmitters to show a detailed profile. GT manufacturers use various methods to calculate the TIT regarding the measured TET. The operator keeps the GT in a protected condition by monitoring the TET.

The TET of an Iranian GT company is used to show the behavior of the methods throughout this study. As shown in *Fig. 6*, the data consists of six sensors, with each sensor representing the TET of V94-2 GT measured approximately with 1500-minute intervals.

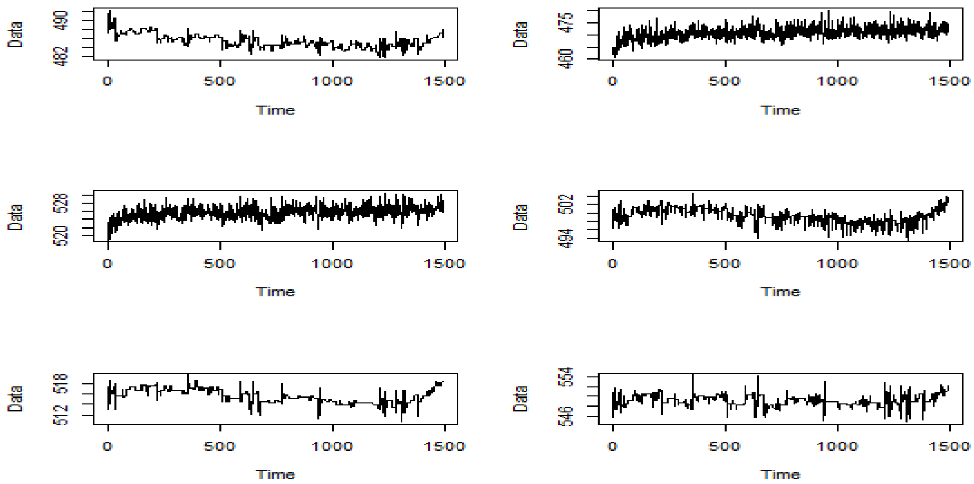

Fig. 6. Trends of TET data for normal process.

These data were collected when all parameters were under control. Statistical tests were implemented to detect the behavior of the data. These data are non-stationary following the test of Kwiatkowski–Phillips–Schmidt–Shin (KPSS). In this subsection, a static PCA model was applied on TET data for controlling the behavior of GT and early detecting fault.

PCA model was introduced into that data along with the results of the observations. Since all data was present, missing data methods were not used. Preprocessing is normally done in various fields. The type of preprocessing depends on the type of process. In the case of the TET data, no special preprocessing is necessary; standardizing the data was the only necessary procedure. After static PCA was applied to the TET data, three components were retained following the CPV criterion. The T^2- and Q-statistics were plotted. The first 400 observations were used to fit the underlying model. No considerable change in the vibration intensity of any of the sensors during this period was observed. The estimated model was evaluated against the well-behaved data observed before $t = 400$, and the confidence limit was set to 99%. Therefore, a distinction is made between phase I, which occurs when $t > 400$, and phase II. As shown in *Fig. 7*, conventional PCA cannot be a good instrument for TET data since the mean of data changes over time, and static PCA created a model with the first bunch of data and was unable to update the model over time. As a result, the model generates a false alarm, despite the fact that these changes are an essential element of the system. The first chart from the left is T^2 control chart and the second chart indicate monitoring by Q.

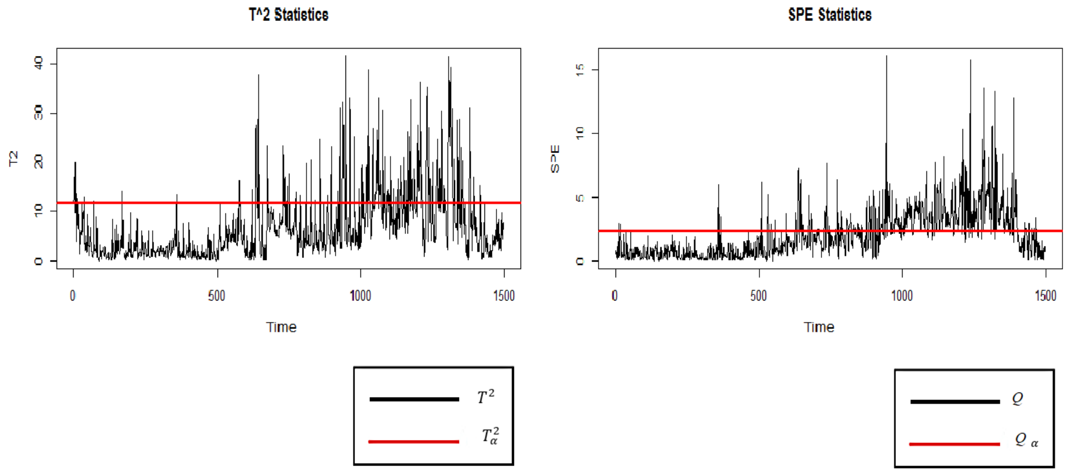

Fig. 7. Monitoring performance of conventional PCA for normal process.

As explained in the previous section, the MWPCA based on V-step-ahead prediction and IPCA were applied to the TET data to solve this problem. The T^2 and Q-statistics regarding MWPCA based on V-step-ahead were plotted in *Fig. 8*. Considering the points mentioned in the previous section, the opinion expert window size and delay size were 400 and 40, respectively. The confidence limit was set to 99%. The first chart from the left is T^2 control chart and the second chart indicate monitoring by Q.

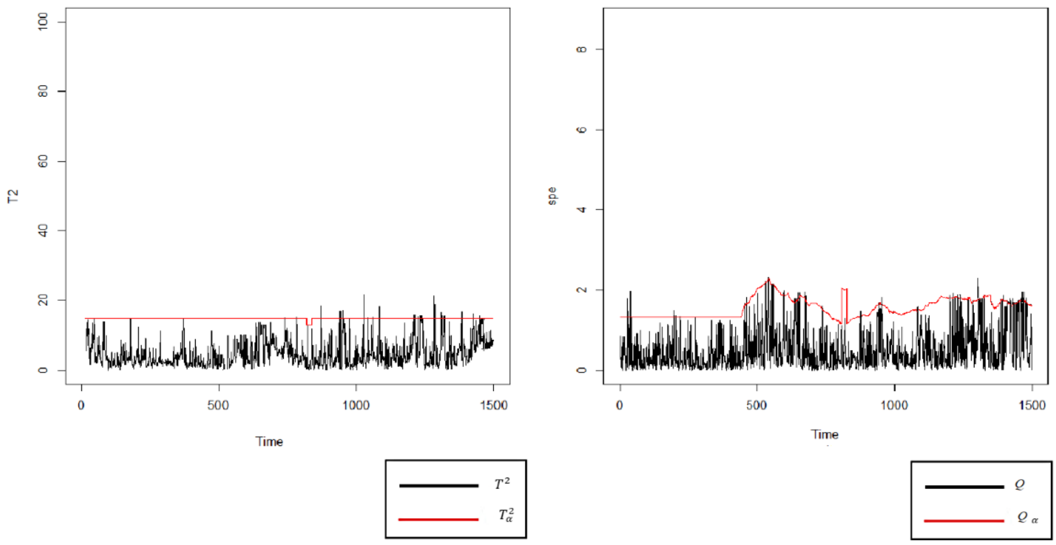

Fig. 8. Monitoring performance of MWPCA based on V-step-ahead for normal process.

Also, the IPCA method was applied to TET data. As shown in *Fig. 9*, the PCs are also non-stationary because the TET data is time-varying. The red lines represent the means of all PCs, and as can be seen, the PCs gradually deviate from the means. IPCs are computed through *Eq. (12)*, and the parameters L and W are set as 6 and 40, respectively. The results of IPCs were plotted in *Fig. 10*. This figure shows IPCs remained around mean and a statistic made by the IPCs will not increase slowly for the normal time-varying process.

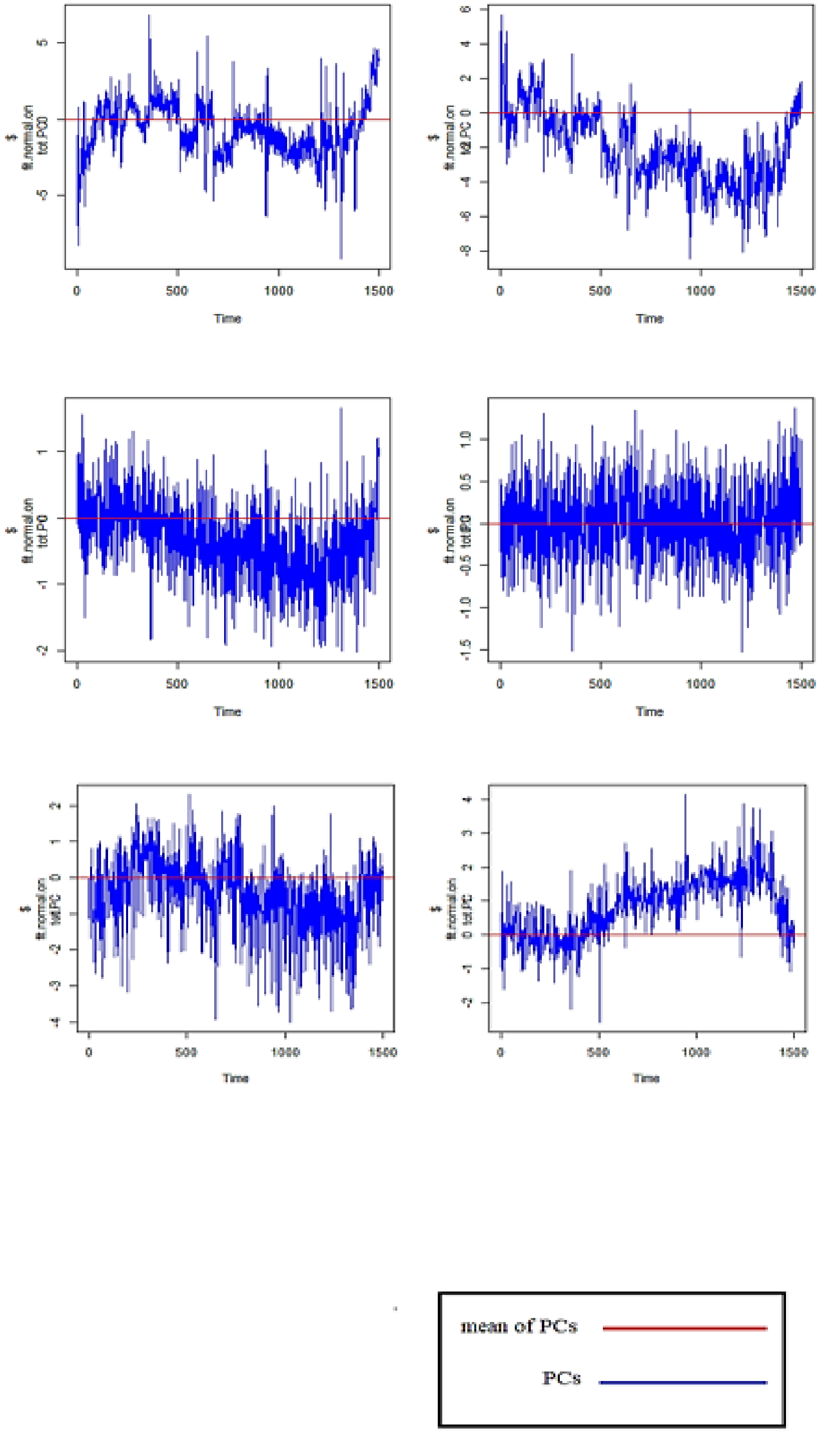

Fig. 9. Trends of PCs for normal process.

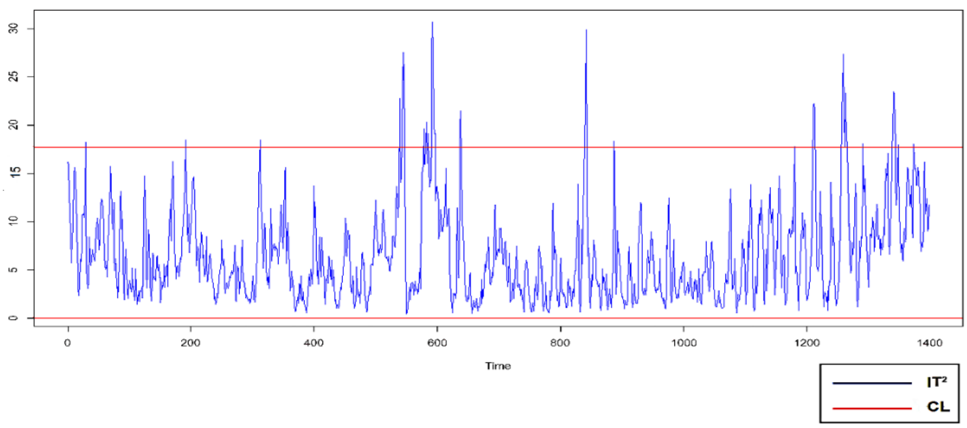

Fig. 10. Trends of PCs for normal process.

Then IT^2 was plotted in *Fig. 11*.

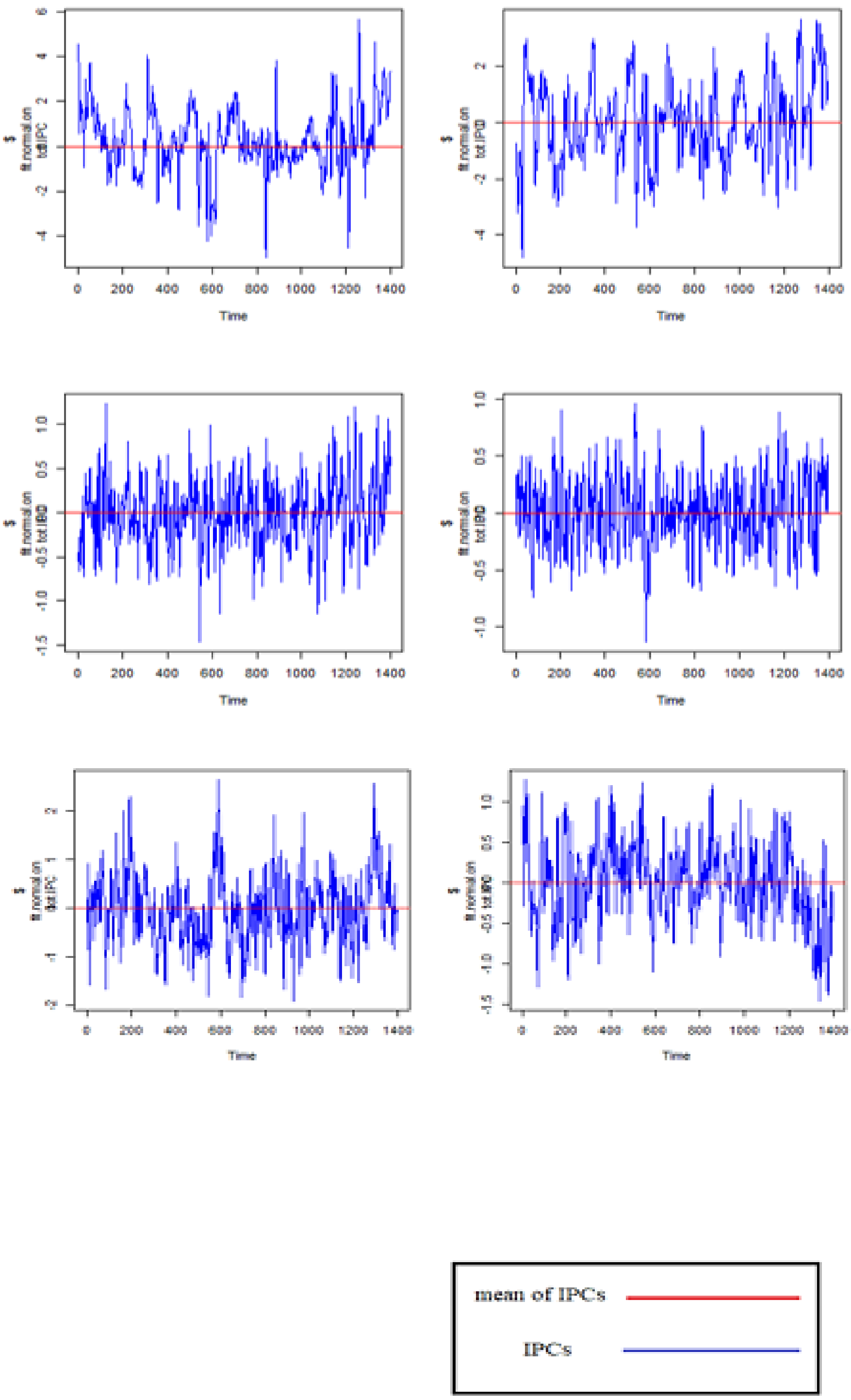

Fig. 11. Monitoring performance of IPCA process with fault 1.

A good fault detection technique should have robust against the training data set and react quickly in fault detection. The robustness is measured by computing the False Alarm Rate (FAR) upon fault free testing data set. Promptness fault detection is quantified by calculating Delay Time Detection (DTD) upon faulty testing data set. According to *Eq. (14)*, a FAR was calculated [44].

$$FAR = \frac{\text{Number of normal samples above the limits}}{\text{total number of normal sample}} * 100. \tag{14}$$

The results show that both MWPCA based on V-step-ahead and IPCA methods overcome non-stationary in this case. *Table 1* shows the percentage of false alarms rate and shows that IPCA performs better than MWPCA based on V-step-ahead.

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

Table 1. FAR percentage.

	MWPCA based on V-step-ahead	IPCA
T^2	5.3%	3.5%
Q	4%	

Two types of faults were added deliberately to check the power of MWPCA based on V-step-ahead prediction and IPCA in finding fault early.

Fault 1. Step change of second TET is introduced at the 1001st sample.

Fault 2. Linear ramp with 0.3 increments of second TET again is introduced beginning from the 1001st. As it was explained in the previous section, the T^2 statistic describes the useful information about system variation, and Q statistic represents model error or noise information. In case of a problem, the covariance structure of the model will be altered and the statistics can represent it. The results of MWPCA based on V-step-ahead are shown in *Figs. 12* and *13*.

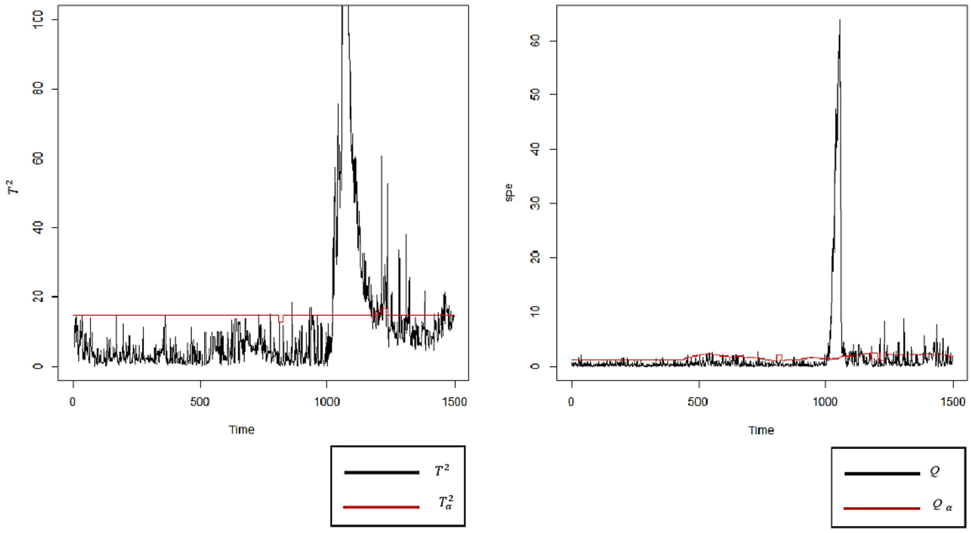

Fig. 12. Monitoring performance of MWPCA based on V-step-ahead for process with fault 1.

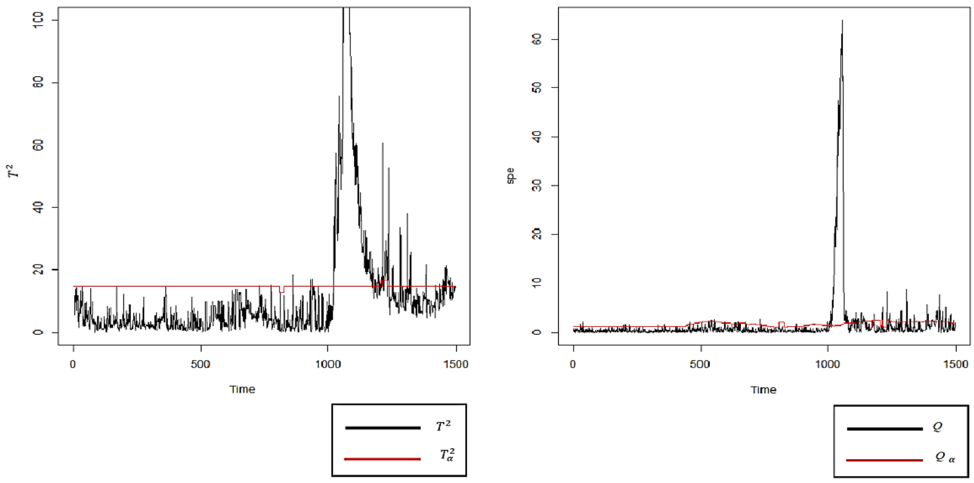

Fig. 13. Monitoring performance of MWPCA based on V-step-ahead process with fault 2.

The results of IPCA are presented in *Figs. 14* and *15*. The DTD index was used to compare these two methods:

$$\text{DTD} = \text{fault detection time} - \text{fault accorance time}. \tag{15}$$

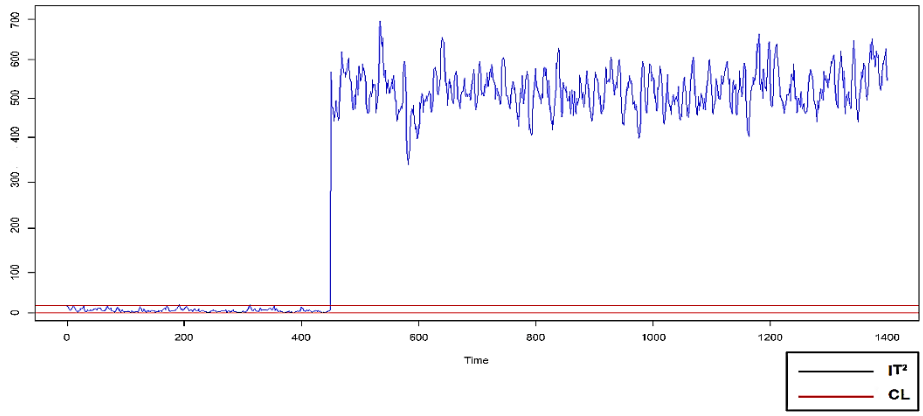

Fig. 14. Monitoring performance of IPCA process with fault 1.

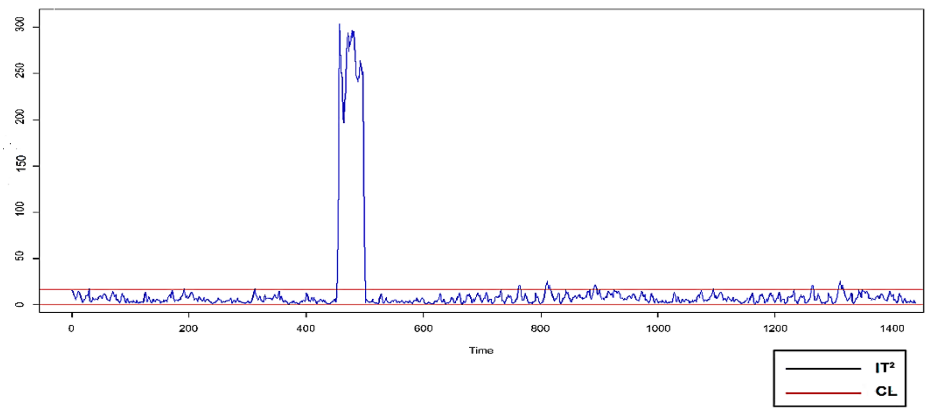

Fig. 15. Monitoring performance of IPCA process with fault 2.

As shown in *Table 2*, both of the methods easily detected the step-change process in time. Also, these charts can distinguish the normal time-varying and slow ramp fault processes.

But according to *Table 2*, slow ramp fault can be detected more quickly by the IPCA model in this case based on the TET data.

Table 2. Detection time delay.

	MWPCA based on V-Step-Ahead		IPCA
	T^2	Q	IT^2
Fault 1	0	0	0
Fault 2	18	11	9

4 | Conclusion

Early fault detection can play an important role in the reliability and availability of GTs. Therefore, providing a monitoring approach is necessary to guarantee early detection of faulty conditions before they lead to a forced plant shut down. The efficiency and specific power of GT would be improved if TIT could be increased. Since the TIT is too hot to be measured directly, it is usually calculated by measuring TET. This study used an appropriate data-driven approach for early fault detection of GTs. To detect faults, data-driven approaches rely on product life-cycle data rather than first-principles models. Hence, data-driven methods can be used for large-scale and complex systems; they are also cheap and inexpensive. PCA model is one of the data-driven methods. Still, unlike simulated data, actual data has ambient and system characteristics, and these parameters do not provide stationary data over

Hajarian et al. | J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112

time. Finding an approach for early detection is important. In this study, the PCA model is implemented on six TET sensors, but it is not a suitable approach due to the non-stationary data. For this purpose, MWPCA based on V-step-ahead and IPCA are implemented on data. MWPCA based on V-step-ahead is a novel monitoring approach for non-stationary TET data. MWPCA based on V-step-ahead is an adaptable approach since it can update the monitoring model and control limit when the newly monitored sample is detected as a normal one. On the other hand, in the IPCA approach, the monitoring model remains unchanged and uses the new statistic called IT^2 to monitor data.

The results reveal that these approaches are data compatible and can detect a step-change fault in real time. However, when it comes to incremental ramp faults, the IPCA approach performs better. According to TET data behavior which is non-stationary and changes over time, a suitable approach has been recommended for early fault detection of turbine gas. In the future, there'll be more work to be done about how monitor high-dimensional, non-stationary, and autocorrelation data from GTs. It is also suggested to find the root of the faults by new fault isolation method and compare their results with classical methods.

References

[1] Sunadi, S., Purba, H. H., & Saroso, D. S. (2020). Statistical process control (SPC) method to improve the capability process of drop impact resistance: a case study at aluminum cans manufacturing industry in Indonesia. *Journal of applied research on industrial engineering*, *7*(1), 92-108.

[2] Molaei, S., & Cyrus, K. M. (2014). Robust design of maintenance scheduling considering engineering insurance using genetic algorithm. *International journal of research in industrial engineering*, *3*(1), 39-48.

[3] Liu, J., Liu, J., Yu, D., Kang, M., Yan, W., Wang, Z., & Pecht, M. G. (2018). Fault detection for GT hot components based on a convolutional neural network. *Energies*, *11*(8), 2149. https://doi.org/10.3390/en11082149

[4] Meher-Homji, C. B., & Bhargava, R. (1992). Condition monitoring and diagnostic aspects of gas turbine transient response. In *turbo expo: power for land, sea, and air* (78965, V004T11A006). American Society of Mechanical Engineers. https://doi.org/10.1115/92-GT-100

[5] Zhang, Y., & Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual reviews in control*, *32*(2), 229-252.

[6] Chiang, L. H., Russell, E. L., & Braatz, R. D. (2000). *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media.

[7] Chiang, L. H., Russell, E. L., & Braatz, R. D. (2012). *Fault detection and diagnosis in industrial systems*. Springer London. https://doi.org/10.1007/978-1-4471-0347-9

[8] Fan, C. M., & Lu, Y. P. (2008). A bayesian framework to integrate knowledge-based and data-driven inference tools for reliable yield diagnoses. *2008 winter simulation conference* (pp. 2323-2329). IEEE.

[9] Alzghoul, A., Backe, B., Löfstrand, M., Byström, A., & Liljedahl, B. (2014). Comparing a knowledge-based and a data-driven method in querying data streams for system fault detection: a hydraulic drive system application. *Computers in industry*, *65*(8), 1126-1135.

[10] Yin, S. (2012). *Data-driven design of fault diagnosis systems* (Doctoral Dissertation, Duisburg, Essen, Universität Duisburg-Essen). https://core.ac.uk/download/pdf/33798712.pdf

[11] Meslameni, W., & Kamoun, T. (2021). Detection of an imbalance fault by vibration monitoring: case of a screw compressor. *Journal of applied research on industrial engineering*, *8*(1), 27-39.

[12] Sankavaram, C., Pattipati, B., Kodali, A., Pattipati, K., Azam, M., Kumar, S., & Pecht, M. (2009). Model-based and data-driven prognosis of automotive and electronic systems. *2009 IEEE international conference on automation science and engineering* (pp. 96-101). IEEE.

[13] Gol-Ahmadi, N., & Raissi, S. (2018). Residual lifetime prediction for multi-state system using control charts to monitor affecting noise factor on deterioration process. *Journal of applied research on industrial engineering*, *5*(1), 27-38.

[14] Losi, E., Venturini, M., Manservigi, L., Ceschini, G. F., & Bechini, G. (2019). Anomaly detection in gas turbine time series by means of bayesian hierarchical models. *Journal of engineering for gas turbines and power*, *141*(11), 111019. https://doi.org/10.1115/1.4044781

[15] Dundas, R. E., Sullivan, D. A., & Abegg, F. (1992). Performance monitoring of gas turbines for failure prevention. In *turbo expo: power for land, sea, and air* (78965, V004T10A013). American Society of Mechanical Engineers.

[16] Jinfu, L., Jiao, L., Jie, W., Zhongqi, W., & Daren, Y. (2017). Early fault detection of hot components in gas turbines. *Journal of engineering for gas turbines and power*, *139*(2). https://doi.org/10.1115/1.4034153

[17] Knowles, M. (1994). Gas turbine Temperature spread monitoring detection of combustion system deterioration. In *turbo expo: power for land, sea,* and air (78873, V005T17A002). American Society of Mechanical Engineers.

[18] Korczewski, Z. (2011). Exhaust gas temperature measurements in diagnostic examination of naval gas turbine engines. *Polish maritime research*, *18*(4), 49-53.

[19] Medina, P., Saez, D., & Roman, R. (2008). On line fault detection and isolation in gas turbine combustion chambers. In *turbo expo: power for land, sea, and air* (43123, pp. 315-324). https://doi.org/10.1115/GT2008-51316

[20] Palmé, T., Liard, F., & Therkorn, D. (2013). Similarity based modeling for turbine exit temperature spread monitoring on gas turbines. In *turbo expo: Power for land, sea, and air* (*55188*, p. V004T06A020). American Society of Mechanical Engineers.

[21] Tsalavoutas, A., Mathioudakis, K., & Smith, M. K. (1996). Processing of circumferential temperature distributions for the detection of gas turbine burner malfunctions. In *turbo expo: power for land, sea, and air* (78767, V005T15A010). American Society of Mechanical Engineers.

[22] Kenyon, A. D., Catterson, V. M., & McArthur, S. D. J. (2010). Development of an intelligent system for detection of exhaust gas temperature anomalies in gas turbines. *Insight-non-destructive testing and condition monitoring*, *52*(8), 419-423.

[23] Navi, M., Davoodi, M. R., & Meskin, N. (2015). Sensor fault detection and isolation of an industrial gas turbine using partial kernel PCA. *IFAC-papers online*, *48*(21), 1389-1396.

[24] Li, W., Peng, M., Liu, Y., Cheng, S., Jiang, N., & Wang, H. (2018). Condition monitoring of sensors in a NPP using optimized PCA. *Science and technology of nuclear installations*, *2018*, 1-16.

[25] Navi, M., Meskin, N., & Davoodi, M. (2018). Sensor fault detection and isolation of an industrial gas turbine using partial adaptive KPCA. *Journal of process control*, *64*, 37-48.

[26] Jolliffe, I. T. (2002). *Principal component analysis for special types of data* (pp. 338-372). Springer New York.

[27] Halligan, G. R., & Jagannathan, S. (2011). PCA-based fault isolation and prognosis with application to pump. *The international journal of advanced manufacturing technology*, *55*, 699-707.

[28] Jackson, J. E. (2005). *A user's guide to principal components*. John Wiley & Sons.

[29] Valle, S., Li, W., & Qin, S. J. (1999). Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. *Industrial & engineering chemistry research*, *38*(11), 4389-4401.

[30] Jaffel, I., Taouali, O., Harkat, M. F., & Messaoud, H. (2015). Online process monitoring using a new PCMD index. *The international journal of advanced manufacturing technology*, *80*(5-8)), 947-957.

[31] Said, M., Fazai, R., Abdellafou, K. B., & Taouali, O. (2018). Decentralized fault detection and isolation using bond graph and PCA methods. *The international journal of advanced manufacturing technology*, *99*(1-4), 517-529.

[32] Li, W., Peng, M., & Wang, Q. (2018). False alarm reducing in PCA method for sensor fault detection in a nuclear power plant. *Annals of nuclear energy*, *118*, 131-139.

[33] Rato, T., Reis, M., Schmitt, E., Hubert, M., & De Ketelaere, B. (2016). A systematic comparison of PCA-based statistical process monitoring methods for high-dimensional, time-dependent processes. *AIChE journal*, *62*(5), 1478-1493.

[34] Gao, Y., Wang, X., Wang, Z., & Zhao, L. (2016). Fault detection in time-varying chemical process through incremental principal component analysis. *Chemometrics and intelligent laboratory systems*, *158*, 102-116.

[35] De Ketelaere, B., Hubert, M., & Schmitt, E. (2015). Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data. *Journal of quality technology*, *47*(4), 318-335.

[36] Kruger, U., & Xie, L. (2012). *Statistical monitoring of complex multivatiate processes: with applications in industrial process control*. John Wiley & Sons.

[37] De Ketelaere, B., Rato, T., Schmitt, E., & Hubert, M. (2016). Statistical process monitoring of time-dependent data. *Quality engineering*, *28*(1), 127-142.

[38] Wang, X., Kruger, U., & Irwin, G. W. (2005). Process monitoring approach using fast moving window PCA. *Industrial & engineering chemistry research*, *44*(15), 5691-5702.

[39] Ding, S. X. (2014). *Data-driven design of fault diagnosis and fault-tolerant control systems*. Springer London.

[40] Langston, L. S., Opdyke, G., & Dykewood, E. (1997). Introduction to gas turbines for non-engineers. *Global gas turbine news*, *37*(2), 1-9.

[41] Ikpe, A. E., Iluobe, I. C., & Imonitie, D. I. E. (2020). Modelling and simulation of high pressure fogging air intake cooling unit of Omotosho phase II gas turbine power plant. *Journal of applied research on industrial engineering*, *7*(2), 121-136.

[42] International Energy Agency. (2006). *Energy technology perspectives*. https://www.iea.org/reports/energy-technology-perspectives-2006

[43] Siemens Energy. (2020). *Ingenious design to the core*. https://www.siemens-energy.com/global/en/offerings/power-generation/gas-turbines/sgt5-2000e.html

[44] Ammiche, M., Kouadri, A., & Bensmail, A. (2018). A modified moving window dynamic PCA with fuzzy logic filter and application to fault detection. *Chemometrics and intelligent laboratory systems*, *177*, 100-113.

Hajarian et al.| J. Appl. Res. Ind. Eng. 10(1) (2023) 97-112