Proceedings of the Pakistan Academy of Sciences: A Physical and Computational Sciences 62(1): 31-40 (2025)

ISSN (Print): 2518-4245; ISSN (Online): 2518-4253
http://doi.org/10.53560/PPASA(62-1)875

Pakistan Academy of Sciences

Research Article

Design and Development of Intelligent Visual Simulator for Fault Detection, Identification and Diagnosis in PWR Nuclear Power Plant

Arshad Habib Malik[1*], Feroza Arshad[2], Aftab Ahmad Memon[3], and Raheela Laghari[4]

[1]Faculty of Engineering, Information and Technology, Sindh Institute of Management and Technology, Karachi, Pakistan
[2]Department of Information System Division, Karachi Nuclear Power Generating Station, Pakistan Atomic Energy Commission, Karachi, Pakistan
[3]Faculty of Electrical, Electronic and Computer Engineering, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan
[4]Faculty of Architecture, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan

Abstract: In this research, the AP600 Pressurized Water Reactor (PWR)-type Nuclear Power Plant (NPP) is studied due to its large number of components and complex, diversified systems. Operating a reliable and economical PWR NPP without malfunctions is desirable, with maximum safety as the primary goal. A Personal Computer Transient Analyzer (PCTRAN) is used as a data-driven source for AP600 PWR NPP, enabling simulations of both normal and abnormal operations. A state-of-the-art, fully automated, intelligent fault detection, identification, and diagnosis software (AI-FDID-PCTRAN) is designed and developed in Visual Basic to address various safety concerns and enhance the reliability and availability of AP600 PWR NPP systems. AI-FDID-PCTRAN is formulated, programmed, and configured based on unsupervised machine learning using Principal Component Analysis (PCA), a fully Automated Multivariate Statistical Process Control Technique (AMSPCT). The proposed PCA-based technique is a purely software-driven, systematically structured, and fully automated approach, developed specifically for the AP600 PWR nuclear industry. This specialized software offers capabilities not found in highly expensive, commercially available alternatives. FDD-PCTRAN has been tested against benchmark normal and abnormal transients available in AP600 PCTRAN and has proven to be highly reliable and accurate in fault detection and diagnosis.

Keywords: Automated System, Fault Detection and Diagnosis, Unsupervised Machine Learning, AP600, Abnormal Operation, PWR.

1. INTRODUCTION

Nuclear power plants are inherently designed with high level of engineering technology and safety. Safety is the prime concern in nuclear industry because the energy source in nuclear power plants is nuclear fission i.e. the production of radiation and ultimate release of nuclear energy. Pressurized Water Reactor (PWR) type nuclear power plants are far safer than other types of nuclear power plants. The PWR type nuclear power plants are basically built on the concept of negative temperature coefficients of reactivity. Nuclear power plants operate around its set-points and sufficient design control bands are provided for transient operation. Plants can cope small, large and even severe transient. However, fault conditions are designed associated with different systems of nuclear power plant and are required to detect and identify timely and accordingly the corrective actions are designed as mitigated actions after proper diagnosis of faults. A detailed literature review is conducted to

Received: September 2024; Revised: February 2025; Accepted: March 2025
* Corresponding Author: Arshad Habib Malik <engr.dr.arshad.habib.malik@gmail.com>

study various fault detection, identification and diagnosis techniques in nuclear power plants. A detailed technical review is conducted to study different fault detection and diagnosis methods used in nuclear power plants by Ma and Jiang [1]. Another detailed review is conducted to study different supervised and unsupervised data driven machine learning techniques used for condition monitoring in nuclear power plants by Hu *et al.* [2]. Fault detection in PWR type nuclear power plant components is addressed by Maio *et al.* [3] using statistical techniques. An online fault detection and isolation technique is established for PHWR type nuclear power plant using PCA technique by Yellapu *et al.* [4]. The research is explored for detection and identification of faults in PWR NPP instruments using Principal Component Method (PCA) by Ma and Jiang [5] while similar research is conducted for multiple faults in PWR sensors by Yu *et al.* [6]. A multi physics informed neural network (PIN) and deep neural network (DNN) based fault detection in regulating valves of nuclear power plants is explained by Lai *et al.* [7]. An entirely different approach is identified by Gallo and Capozzi [8] for pattern classification and feature selection using nonlinear PCA. Similar research is conducted for fault detection in dynamic time varying systems by Kazemi *et al.* [9] using neural network based PCA. A Simulink based reactor power control system is designed for PWR type nuclear power plant using Westinghouse Personal Computer Transient Analyzer (PCTRAN) by Ihrayz [10]. PCTRAN is used as virtual simulator of PWR dynamics. A fault diagnosis system is developed for CPR1000 3-loop PWR type nuclear power plant using deep neural network by Liu *et al.* [11]. Research is further explored for CPR1000 3-loop PWR type nuclear power plant by Ren *et al.* [12] in which Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) based model is constructed for fault diagnosis in normal and abnormal operating conditions. These scenarios are designed in specific CPR1000 PCTRAN. This research is further extended by Zhang *et al.* [13] in which fault diagnosis for CPR1000 PCTRAN is accomplished using sparrow search algorithm optimized CNN LSTM neural network. A PCA and Support Vector Machine (SVM) based fault diagnosis system is established for Main Steam Line Break (MSLB) accidental condition in CPR1000 PCTRAN by Xin *et al.* [14]. An investigation is performed for prediction of time series of nuclear power plant parameters using Backpropagation (BP) artificial neural network (ANN) by Liu *et al.* [15]. This research is carried out for Advanced Pressurized Water Reactor of 600 MWe rating (AP600) nuclear power plant. Another similar investigation is performed by Yong-kuo *et al.* [16] in which signed directed graph (SDG), PCA and ANN based cascaded fault diagnostic technique is developed for similar AP600 NPP. A PCA based fault detection and diagnosis methodology is designed for selected AP600 NPP fault conditions of primary and secondary systems by Elshenawy *et al.* [17]. Current methods with commercially available software are not intelligent and automated in their functionality. Methods other than PCA such as deep learning are time consuming, complex and mathematically stiff.

In the present study, a dedicated PCTRAN developed specially for AP600 PWR NPP in Visual Basic is integrated with a newly designed fully Automated Intelligent Fault Detection, Identification and Diagnosis software (AI-FDID) developed in Visual Basic. The integrated AI-FDID-PCTRAN is a state-of-the-art dedicated software for AP600 PWR nuclear power plant designed and configured based on data-driven unsupervised machine learning technology. The software is highly robust, intelligent and smartly detects normal and abnormal operations of AP600 PWR nuclear power plant. A dedicated software database is designed and PCA is addressed computationally and simulation experiments prove the effectiveness of proposed scheme.

2. MATERIALS AND METHODS

2.1. AP600 PCTRAN

PCTRAN is a computer code specially designed for PWR type nuclear power plants. PCTRAN program is a generic software covers light water reactor technology [10]. It is a Westinghouse design with thermal output power about 1800 MW_{th} (i.e., 600 MWe). PWR simulators are available for different power ratings and variants. Parameters of AP600 PCTRAN are tuned for 600 MWe.

In a Pressurized Water Reactor's design, reactor coolant is light water which also acts like moderator. The reactor has vertical reactor core filled with fuel, coolant and control rods. Reactor coolant is a pressurized water and its quality and chemistry is maintained. Reactor coolant is passed through the tubes of steam generator while feed water is passed through the shell structure of steam generators. High quality steam is produced and passed through the steam turbines. Steam turbines generate mechanical energy that in turn rotates the electrical generator and thereby the electrical energy is produced. All the systems work together and form a PWR power generating station [12]. The AP600 PCTRAN Simulator interface is shown in Figure 1.

2.2. Unsupervised Machine Learning: PCA

Machine learning techniques are widely used in system identification and pattern identification. Principal Components Analysis (PCA) is one of the very popular unsupervised machine learning technique. The basic design of PCA is established on statistical analysis. This technique is well suited for detecting the precursors responsible for changes in data. The core idea behind the PCA is to estimate and formulate the variance in process data [4]. Now-a-days, PCA is used extensively used in all areas of sciences, engineering and computing for statistical analysis to produce useful information for process and its health conditions.

In model based approach, first principle problem modeling is adopted for which very accurate and insightful knowledge of the process

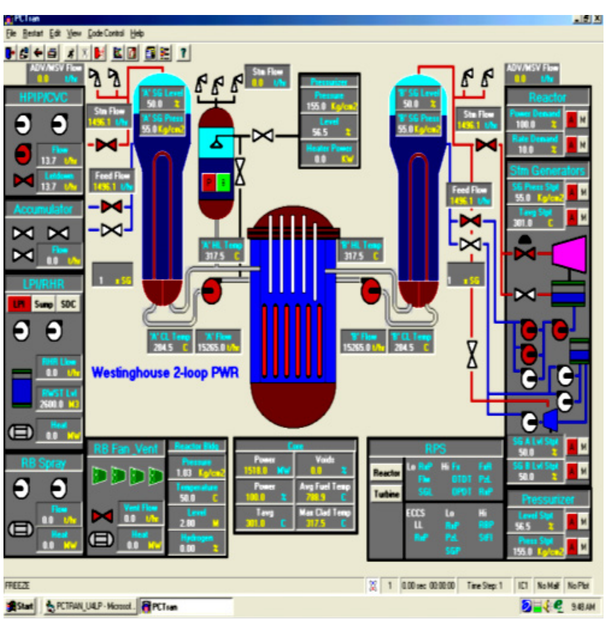

Fig. 1. AP600 PCTRAN simulator interface.

physics is required [5]. This modeling approach is governed on physical laws. Therefore, sometimes, it becomes very difficult to precisely model the process. In order to cater this problem, a data-driven approach is adopted. That's why data-driven PCA methodology is very popular for pattern identification. This method is further classified into statistical methods and non-statistical or signal based methods. PCA is a multivariate algorithms based technique and uses data driven approach which captures nonlinear patterns of datasets very well, so it is well suited for nuclear power plant applications in a much better and easier way then nonlinear techniques such as neural networks etc. That is why nonlinear techniques are not adopted due to complex structure and difficult configuration of software.

The task of PCA is then, given a sample of *n* objects with *m* measured quantities for each, i.e. *m* variables, x_j ($j = 1, 2, ..., m$), find a set of *m* new, orthogonal (i.e. independent) variables, $t_1, \ldots\ldots, t_i, \ldots\ldots t_m$, each one a formulation of actual variables, x_j as:

$$t_i = a_{i1}x_1 + ... + a_j x_j + ... + a_{im} x_m \qquad (1)$$

Where, constants a_{ij} are determined such that the smallest number of new set of variables is observed with maximum possible variance present in the data. The t_i dimension is called principal component and each value of t_i is known as score. The feature matrix *P* is known as loadings, which is constructed by taking the eigenvectors in the columns. The variation in the measurement domain covered by the loading vectors coupled with smallest eigenvalues. Eigenvectors are computed either from covariance matrix or SVD [3]. For covariance, output data from the process is standardized. Covariance is $m \times m$ symmetric matrix which shows the linear dependence of one dimension on another dimension. Then eigenvalues and eigenvectors will be calculated numerically. Cumulative Percentage Variance (CPV) test is performed to find out the number '*a*' for reduced dimensions. After finding out the '*a*', loading matrix is selected by using first '*a*' dimension, as it captures the most of information represented by data. The new variables scores will be produced from loadings and scaled data, which represent the data in reduced order. By using eigenvectors and scores, original observed data can be reproduced with little bit error.

Whereas another technique which is associated with matrix representation called Singular Value Decomposition (SVD). SVD deals with orthogonal matrix U and diagonal matrix W of singular values and the transpose of an m x m orthogonal matrix V^T. Eigenvalues are the square of these singular values while eigenvectors can be calculated by taking the transpose of V^T.

Eigenvectors can be computed by using either covariance matrix or SVD which are then sorted by their magnitude. At this step, it is decided to neglect the values of least importance. By doing so, some information will be lost but there will be no loss in case of smaller eigenvalues. If some components are ignored, then the final data set contains fewer dimensions than the original [17]. To be precise, if originally have m dimensions in data, then m eigenvectors and eigenvalues are calculated, and then only the first eigenvector is chosen, later the final data set has 'a' dimension. This can be done by using CPV method. In this method, 'a' is computed using the smallest number of loading vectors required to cater CPV.

2.3. T^2 and Q-Statistics Tests

PCA technique is used to decompose the original data into two or more orthogonal subspaces which independently provide unique information about the health of the plant. The model subspace is derived by choosing the first k principal components which provides information about parameters that shows greater changes in variation. The residual subspace, orthogonal to the model subspace provides the inverse information which describes what happens behind the scenes. The Q statistic is used to quantify the inverse information from the residual subspace and helps to overcome the masking concerns of faulty parameters that does not exhibit large changes in variation. The detection of changes in the simulation scenarios accomplishes independently through the Hotelling T^2 and square prediction error (SPE) or Q-statistics respectively. A combined discriminant that uses both the T^2 and SPE is used, which is capable of improved diagnosis based on assign weight. Contribution plots are used to examine the contribution of each parameter to the T^2 and SPE statistics respectively. From the contribution plots, it gives an idea that which parameter is causing the changes or fault. The computations are based on the matrix algebra of the singular value decomposition of the data correlation matrix. This is found to be very useful because it simplifies and reduces the number of processing steps required to complete the algorithm.

First, data is scaled then PCA is applied for fault detection. Proper scaling the data is very essential in PCA. Then T^2 and Q statistics are applied to draw control charts using calculated upper limits for fault detection. When an out-of-control signal is detected by T^2 or Q statistics control charts, then its identification is very essential. Fault diagnosis is a very critical function as generally while assessing variables, some variables of course move beyond very quickly during persistence of fault conditions. So, this information helps to further diagnose an actual cause of the fault. For this purpose, contribution plots can be used by using T^2 and Q statistics. For illustration purposes, a case study is selected to verify and validate the computations performed by software developed using MS Visual Basic 6.0. For this purpose, each of feature parameter, whole 150 observations data of Iris Flower are Scaled using mean and standard deviation and then 'Fault Detection' wizard in the software is applied to whole standardized data. The Fisher's Iris flowers dataset is designed consisting of data observations for three species. Four features are measured from each sample. Depending on mix of four features, a model is developed for analysis purposes. The plot of first three components of the dataset under consideration in 3D is shown in Figure 2.

Three principal components captures 99.5% of total deviation. Therefore, it mimics the maximum

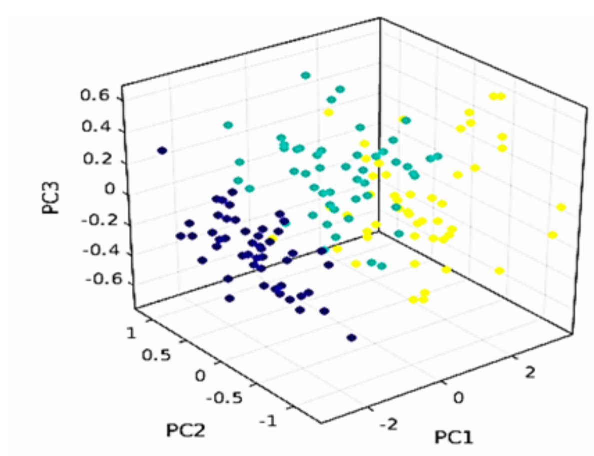

Fig. 2. Plots for first three principal components.

features through dimension reduction process. This proves the effectiveness and successful realization of the principal component analysis methodology adopted in extracting the features.

First two eigenvectors are selected in loading matrix for dimension reduction. The plot of scores computed by software for first two principal components is shown below in Figure 3.

The process in which fault is detected and then process variable is identified that contribute to fault which result in out-of-control status. In the present study, fault diagnosis is accomplished by categorizing different faults into different fault classes. Combined discriminant based on T^2 and Q statistics using Euclidean distance is used as a first choice for fault diagnosis, which is a pattern classification technique. Loading is stored into the database to measure the similarity between the two faults. The diagnosis results cover the similarity index and Euclidean distance calculated during match with each class stored in database. The statistical design data of PCA is summarized in Table 1.

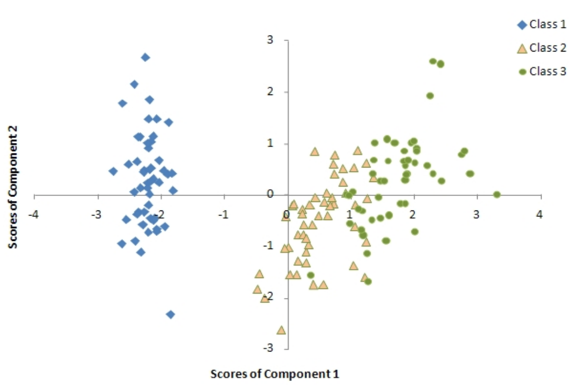

Fig. 3. Score plots for first two principal components.

Table 1. Statistical design data of PCA.

Statistical Design Parameters	Values
Total Observations	150
T_α	6.6447
T_β	4.9518
Similarity Index for Fault Class-1	0.5024
Similarity Index for Fault Class-2	0.9312
Similarity Index for Fault Class-3	1.0001
Euclidean Distance for Fault Class-1	0.1092
Euclidean Distance for Fault Class-2	0.0208
Euclidean Distance for Fault Class-3	0

2.4. Software Tool Design: AI-FDID-PCTRAN

In AI-FDID-PCTRAN design cycle, the first step is the design. In the present study, structural design approach is adopted. Its goal is to produce a model or representation of an entity that is decided to build. For this fault detection and diagnosis system classical structural design approach is used.

Modules are collection of functions, which perform same functionalities collectively and coherently. Module is described by its purpose and complete specification of functions included in it. In fault detection and diagnosis system two major modules are designed, which are fault detection training module and fault diagnosis testing module. These two modules share different functions. As both the modules constitute FDD system collectively, therefore, the common functions can be re-used from other module. Fault detection training module includes the functions which detect and identify fault in data.

Firstly, the data is gathered from the process then this data is scaled using mean and standard deviation. Also, data is classified into different fault classes based on faults for fault detection and diagnosis. Therefore, for each class initially principal component analysis is performed and then T^2 and Q statistics are calculated. The last task in this module is to calculate combined discriminant of class for which training is in progress and then store loadings, eigenvalues, eigenvectors, combined discriminant and fault information into the database. So, this module allows the FDD system to be trained for normal as well as different fault conditions, and to be used for fault diagnosis in next stage.

Fault diagnosis testing module mirrors the fault detection architecture, it includes the functions which diagnose the type of fault. After analyzing the given data, it is compared with fault classes which are already stored in database. Whichever fault class has the closest match is basically the desired category of fault.

This module performs principal component analysis, compute T^2 and Q statistics on given set of data for which fault diagnosis process is initiated, therefore, it shares many functions of fault detection module. After performing analysis, it diagnoses fault by matching combined discriminant with already stores combined discriminant of fault classes. It also computes similarity measure by making use of loading of given test case and loadings are already stored in the database for different fault classes. Whichever fault class already stored has the closest match is basically the desired category of fault. Fault diagnosis module shares almost all the functions of fault detection module except functions used for fault identification purpose using T^2 and Q statistics. In this study relational database design is used.

3. RESULTS AND DISCUSSION

In the present study, the AI-FDID-PCTRAN software design interface, performance analysis of AI-FDID-PCTRAN software in normal and abnormal conditions are investigated. Obstacles exist for using the proposed model across different types of nuclear reactor designs are the variations of parameters, configuration of parameters and system loops, etc.

3.1. AI-FDID-PCTRAN Software GUI

The AI-FDID-PCTRAN software is a systematic and fully automated intelligent application program. Rapid Application Development (RAD) environment i.e. Visual Basic 6.0 is used in the development and evaluation of User Interfaces (UIs). Visual Basic is used as it is user friendly and it is adopted because Pakistan Nuclear Regulatory Authority (PNRA) already using Visual Basic for all type of simulators and software development. PNRA has licensed Visual Basic software that can updated and modules could be improved for specific requirements. Multiple Document Interface Design Style (MDIDS) are used to make it easier to work with system. The computational tools are made available in AI-FDID-PCTRAN software and embedded in GUI tabs. The software qualifies the compliance and regulatory adherence of PNRA, which is necessary for using the software in the nuclear industry. It meets the requirements of PNRA and tested against benchmark data.

3.2. Performance Analysis of AI-FDID-PCTRAN in Normal Operating Conditions

Now, the testing and validation of fault detection, identification and diagnosis system is carried out using a case study in which principal component analysis is performed using T^2 and Q statistics for PCTRAN Simulator for 600 MWe PWR type nuclear power plant Data. In order to detect, identify and diagnose faults, following classes are considered in this case study:

1. Normal operation
2. Loss of coolant accident at hot leg for 95% break
3. Main steam line break for 95%
4. Main feed water pump failure

Four above mentioned associated fault scenarios are chosen due to space limitation and scope of this research study.

The process variables, i.e., parameters under consideration are:

1. Primary pressure: P (Kg/cm^2)
2. Average temperature: T_{avg} (°C)
3. Hot leg temperature: THA (°C)
4. Cold leg temperature: TCA (°C)

For each class 3000 observations/samples are collected in 3000 seconds through PCTRAN simulator. Then for each class, software is trained through fault detection wizard to store loading and combined discriminant into database for similarity measure and Euclidean distance computation in diagnosis phase respectively.

Initially, when simulator starts there may be some fluctuations for a certain time. After some time, simulator starts to operate in normal condition under steady state. The values for variables under consideration in normal operation are mentioned in Table 2.

Eigenvalues and eigenvectors are computed for normal operation of AI-FDID-PCTRAN for parametric variation from steady state to fault conditions as shown in Tables 3 and 4, respectively. It means this data is the representation of variation from normal operating conditions to transient conditions capturing the deviation dynamics.

In Table 4, component 1 through component 4 are the first four principal components. Eigenvectors are direction cosines for principal components, while eigenvalues are the variance in the principal components. Sum of all eigenvalues

Table 2. Normal operation parametric values of variables.

Parameters	Values
P (Kg/cm^2)	154.96
Tavg (°C)	300.93
THA (°C)	320.43
TCA (°C)	281.42

Table 3. Eigenvalues computed by AI-FDID-PCTRAN for normal plant operation data.

Parameters	Eigenvalues	Difference
Component 1 (C-1)	3.1584	2.3882
Component 2 (C-2)	0.7702	0.7061
Component 3 (C-3)	0.0641	0.0568
Component 4 (C-4)	0.0073	-

Table 4. Eigenvectors computed by AI-FDID-PCTRAN for normal plant operation data.

Parameters	(C-1)	(C-2)	(C-3)	(C-4)
P (Kg/Cm^2)	0.3183	0.9389	-0.1292	-0.0228
Tavg (°C)	-0.5529	0.1947	-0.0895	0.8052
THA (°C)	-0.5366	0.2719	0.7159	-0.3546
TCA (°C)	-0.5523	0.0821	-0.6803	-0.4747

is equal to the sum of variances which are on the diagonal of the variance-covariance matrix.

Eigenvectors are the vectors indicating the direction of the axes along which the data varies the most. Each eigenvector has a corresponding eigenvalue, quantifying the amount of variance captured along its direction. PCA involves selecting eigenvectors with the largest eigenvalues.

The largest eigenvalues with positive sign and their corresponding eigenvectors is crucial in PCA because they capture the most significant patterns in the data, reducing dimensionality while retaining most of the variability in plant dynamics. The change in eigenvalue shows the variation in eigenvalue due to the difference between steady state and fault conditions of plant.

Fault detection in AI-FDID-PCTRAN software is carried out by using T^2 and Q statistics control charts. In both control charts, before AP600 NPP reaches its steady state, there are some faults at the start. But after reaching steady state, no faults are observed using these statistics. T^2 and Q statistics are applied to draw control charts using calculated upper limits for fault detection. The threshold values are chosen using known plant normal and abnormal fault conditions for selected scenarios using computational schemes mentioned by Elshenawy *et al.* [17].

The threshold values computed for T^2 and Q statistics are 6.0015 and 0.4456 respectively. The resultant plots for fault detection using these statistics are shown in Figures 4 and 5 respectively. The results are generated using Microsoft Visual Basic software and found graphically systematic and visually unique in nature as compared to results observed previously [1-6].

Through fault identification contribution of each variable in faults is calculated by using T^2 and Q statistics contribution plots. As T^2 statistic and Q statistic detect and identify different type of faults,

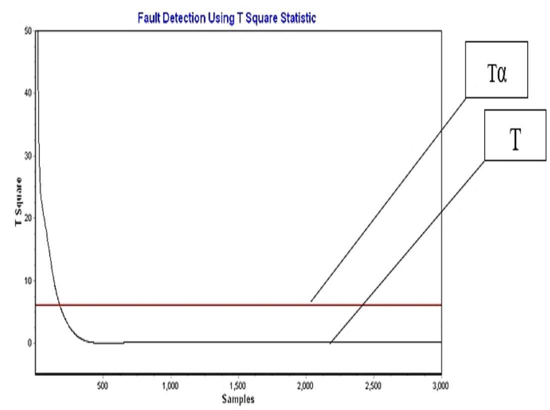

Fig. 4. Fault Detection through T^2 Statistics using Normal Plant Operation of AI-FDID-PCTRAN.

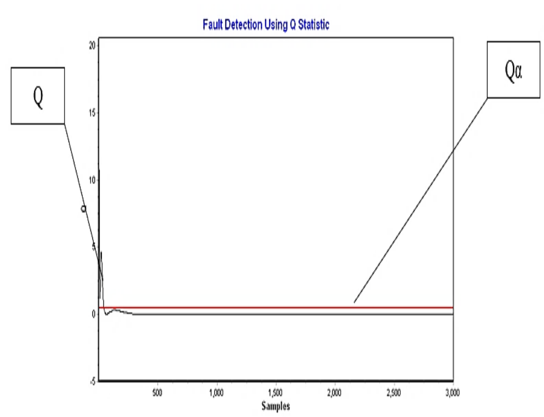

Fig. 5. Fault Detection through Q Statistics using Normal Plant Operation of AI-FDID-PCTRAN.

therefore, in this case study, it is observed that T^2 statistics are more sensitive to pressure surges than temperature surges whereas Q-statistic is sensitive to temperature changes. So, using T^2 statistic, it is observed that initial faults before achieving steady state condition are mainly due to pressure changes while fault detected by Q-statistics are mainly due to hot leg temperature changes. The contribution plots for fault identification using T^2 and Q statistics are shown in Figures 6 and 7 respectively. The normal operation plant data is successfully diagnosed from all other fault classes with similarity measure as 1.001 and Euclidean distance as 0.

3.3. Performance Analysis of AI-FDID-PCTRAN in Abnormal Operating Conditions

Realistic applications of AI-FDID-PCTRAN are Different abnormal conditions of plant and here in this research work as case study a scenario such as MSLB has been chosen for its specific usage.

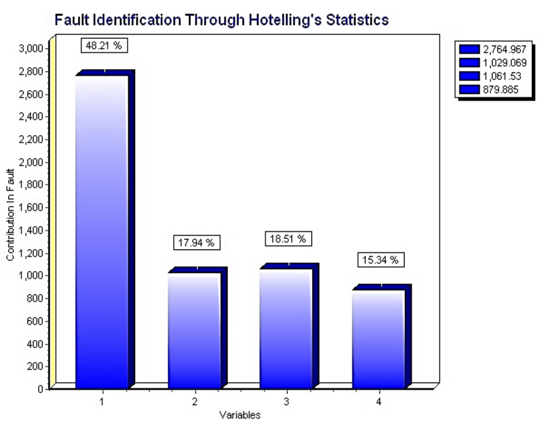

Fig. 6. Fault Identification through T^2 Statistics using Normal Plant Operation of AI-FDID-PCTRAN.

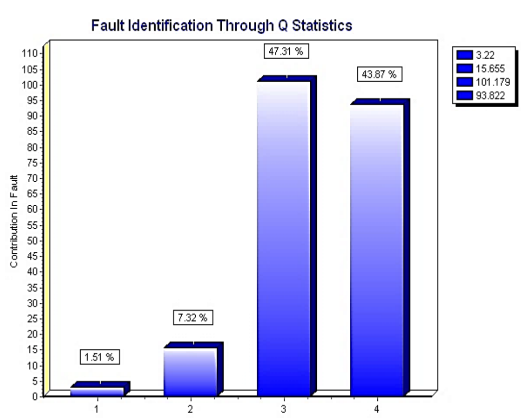

Fig. 7. Fault Identification through Q Statistics using Normal Plant Operation of AI-FDID-PCTRAN.

In abnormal operation, MSLB case is considered as typical scenario and presented in this study. When main steam line break occurs, initially T_{avg} is increased. This initial increment is due to the fact that steam flow rate is increased, and then reactor coolant system (RCS) temperature deceases as temperature difference increases in steam generator. Therefore, positive reactivity arises in the system, therefore temperature is increased due to increased fission rate. Initially at 100% reactor power T_{avg} is 302 °C. As turbine trips, neutron flux becomes zero due to the absorption of thermal neutrons in control rods. In other words, heat generation is stopped. The increment in the RCS temperature reduces the density of coolant. Therefore, water level of coolant in the pressurizer increases. This level increment produces small peak of pressure rise in coolant. As a result, the reactor is tripped, and RCS temperature continues to decrease further. This decrement in temperature reduces the system pressure.

Eigenvalues and eigenvectors are computed for MSLB introduced in AP600 PCTRAN using AI-FDID-PCTRAN for parametric variation from steady to fault conditions are mentioned in Tables 5 and 6 respectively.

Eigenvalues and eigenvectors are computed for abnormal operation of AI-FDID-PCTRAN for parametric variation from steady state to fault conditions as shown in Tables 5 and 6 respectively. It means this data is the representation of variation from abnormal operating conditions to transient conditions capturing the deviation dynamics.

In Table 5, the difference is calculated numerically by considering the eigenvalues at steady state and abnormal plant conditions at faults. The reference data or benchmark is the plant design data and design conditions describing plant normal operating conditions and fault condition as mentioned in plant design document available with PNRA.

The change in eigenvalues drops from first principal component through fourth principal component as the impact of plant dynamics reduces from first through fourth principal component. Table 6 has the similar interpretation for abnormal operation of plant as that of Table 4 discussed earlier.

Table 5. Eigenvalues computed by AI-FDID-PCTRAN for abnormal plant operation data (MSLB).

Parameters	Eigenvalues	Difference
Component 1 (C-1)	2.993	1.9884
Component 2 (C-2)	1.0046	1.0022
Component 3 (C-3)	0.0024	0.0024
Component 4 (C-4)	0	-

Table 6. Eigenvectors computed by AI-FDID-PCTRAN for abnormal plant operation (MSLB) data.

Parameters	(C-1)	(C-2)	(C-3)	(C-4)
P (Kg/Cm^2)	-0.0534	0.9934	-0.1011	-0.0011
Tavg (°C)	0.5775	0.0404	0.1005	-0.8091
THA (°C)	0.575	0.097	0.644	0.4952
TCA (°C)	0.577	-0.0452	-0.7516	0.3163

Fault detection in AI-FDID-PCTRAN software is carried out using T^2 and Q statistics control charts and only PCA method has been adopted for this study due to computational ease, unsupervised learning and flexibility of usage in visual software. T^2 statistics is more sensitive to pressure surges so when main steam line break occurs. The pressure fluctuation is observed not severe enough to be uncontrollable, so after that pressure begins to increase towards 155kg/cm^2. After some time, pressure again starts to decrease which is indicated by the rising trend at the end of scenario. The threshold value computed for T^2 statistic is 6.0015 while Q-statistic is more sensitive to detect temperature changes. As main steam line break occurs, initially temperature rises and then starts decreasing. The threshold value for Q-statistic computed by AI-FDID-PCTRAN software is 0.0158. Through fault identification contribution of each variable in faults is calculated by using T^2 and Q statistics. As T^2 statistic and Q statistic detect and identify different type of faults, therefore, in this case study, it is observed that T^2 statistics are more sensitive to pressure surges than temperature surges whereas Q-statistic is sensitive to temperature changes. So, using T^2 statistic, it is observed that faults are mainly due to pressure changes while fault detected by Q-statistics are mainly due to cold leg temperature changes. The main steam line break data is successfully diagnosed from all other fault classes with similarity measure as 0.9999 and Euclidean distance as 0. Alternative methods such as deep learning, SVM, or hybrid models with commercially available software are not intelligent and automated in their functionality. Hence, methods other than PCA are time consuming, complex, mathematically stiff as reported previously [7-15] and are beyond the scope of this research work. The case studies or scenarios presented in this study are unique and precise with great degree of reliability as compared to results reported by Ren *et al*. [12].

4. CONCLUSIONS

AP600 PWR nuclear power plant is attempted to extract the parametric innovative operational data using dedicated PCTRAN. PCA technique is used as unsupervised machine learning technique to decompose the original plant data into two or more orthogonal subspaces which independently provided unique information about the health of the plant. Fault detection is carried out using Q statistics while fault identification is accomplished through T^2 and Q statistics and fault diagnosis is established from designated fault classes. The performance of the proposed design scheme is tested for normal and abnormal plant operating conditions. Simulation experiments proves that the designed and developed AI-FDID-PCTRAN software is robust and fully automated as valid for normal and abnormal conditions with demonstrated variations in scenarios having entirely different plant conditions. The AI-FDID-PCTRAN software is intelligent and capable enough to predict the health monitoring and condition monitoring of any other system of AP600 available in PCTRAN that can be extended for different variants of PWR in future.

5. ACKNOWLEDGEMENTS

The support of the Sindh Institute of Management and Technology, Information Support Division of KNPGS and Mehran University of Engineering and Technology, is gratefully acknowledged.

6. CONFLICT OF INTEREST

The authors declare no conflict of interest.

7. REFERENCES

1. J. Ma and J. Jiang. Application of fault detection and diagnosis methods in nuclear power plants: A Review. *Progress in Nuclear Energy* 53: 255-266

(2011).

2. G. Hu, T. Zhou, and Q. Liu. Data-Driven Machine Learning for Fault Detection and Diagnosis in Nuclear Power Plants: A Review. *Frontiers in Energy Research* 9: 663296 (2021).
3. F.D. Maio, P. Baraldi, E. Zio, and R. Seraoui. Fault Detection in Nuclear Power Plants Components by a Combination of Statistical Methods. *IEEE Transactions on Reliability* 62(4): 833-845 (2013).
4. V.S. Yellapu, V. Vajpayee, Akhilan, and P. Tiwari. Online Fault Detection and Isolation in Advanced Heavy Water Reactor using Multiscale Principal Component Analysis. *IEEE Transactions on Nuclear Science* 66(7): 1790-1803 (2019).
5. J. Ma and J. Jiang. Detection and Identification of Faults in NPP Instruments using Kernel Principal Component Analysis. *Journal of Engineering for Gas Turbines and Power* 134: 103990 (2012).
6. Y. Yu, M. Peng, H. Wang, Z. Maa, and W. Li. Improved PCA Model for Multiple Fault Detection, Isolation and Reconstruction of Sensors in Nuclear Power Plant. *Annals of Nuclear Energy* 148: 107662 (2020).
7. C. Lai, I. Ahmed, E. Zio, Wei. Li, Y. Zhang, W. Yao, and J. Chen. A Multistage Physics-Informed Neural Network for Fault Detection in Regulating Valves of Nuclear Power Plants. *Energies* 17: 2647 (2024).
8. C. Gallo and V. Capozzi. Feature Selection with Non-Linear PCA: A Neural Network Approach. *Journal of Applied Mathematics and Physics* 7: 2537-2554 (2019).
9. P. Kazemi, A. Masoumain, and P. Martin. Fault Detection and Isolation for Time Varying processes using Neural based Principal Component Analysis. *Processes* 12(6)*:* 1218 (2024).
10. E.A.M.B. Ihrayz. PCTRAN Westinghouse AP1000 Power Control of Pressurized Water Reactor using Simulink of MATLAB. *Open Journal of Energy Efficiency* 12: 25-35 (2023).
11. B. Liu, J. Lei, J. Xie, and J. Zhou. Development and Validation of a Nuclear Power Plant Fault Diagnosis System based on Deep Learning. *Energies* 15*:* 8629 (2022).
12. C. Ren, H. Li, J. Lei, J. Liu, W. Li, K. Gao, G. Huang, X. Yang, and T. Yu. A CNN-LSTM based Model to FAULT Diagnosis for CPR1000. *Nuclear Technology* 209(9): 1365-1372 (2023).
13. C. Zhang, P. Chen, F. Jiang, J. Xie, and T. Yu. Fault Diagnosis of Nuclear Power Plant based on Sparrow Search Algorithm Optimized CNN-LSTM Neural Network. *Energies* 16*:* 2934 (2023).
14. M. Xin, W. Jiao, and L. Da-zi. Fault Diagnosis of Nuclear Power Plant based on Simplified Signed Directed Graph with Principal Component Analysis and Support Vector Machine. *Chinese Automation Congress, 22nd-24th November 2019, Hangzhou, China* (2009).
15. Y. Liu, F. Xie, C. Xie, M. Peng, G. Wua, and H. Xia. Prediction of Time Series of NPP Operating Parameters using Dynamic Model based on BP Neural Network. *Annals of Nuclear Energy* 85: 566-575 (2015).
16. L. Yong-kuo, A. Abiodun, W. Zhi-bin, W. Mao-pu, P. Min-jun, and Y. Wei-feng. Rajesh. A Cascade Intelligent Fault Diagnosis Technique for Nuclear Power Plants. *Journal of Nuclear Science and Technology* 55(3): 254-266 (2018).
17. L.M. Elshenawy, M.A. Halawa, T.A. Mahmoud, H.A. Awad, and M.I. Abdo. Unsupervised Machine Learning Techniques for Fault Detection and Diagnosis in Nuclear Power Plants. *Progress in Nuclear Energy* 142: 103990 (2021).