FSN-Based Fault Modelling for Fault Detection and Troublshooting in CANDU® Stations

E. Nasimi[1], H.A. Gabbar[2]
[12] Bruce Power LLP, elnara.nasimi@brucepower.com
[2] University of Ontario Institute of Technology

Abstract

An accurate fault modeling and troubleshooting methodology is required to aid in making risk-informed decisions related to design and operational activities of current and future generation of CANDU® designs. This paper presents fault modeling approach using Fault Semantic Network (FSN) methodology with risk estimation. Its application is demonstrated using a case study of Bruce B zone-control level oscillations.

1. Fault Semantic Network Methodology

1.1 Introduction to Fault Semantic Network (FSN) Methodology

In this study, a case of zone-control level oscillations at Bruce B will be used to demonstrate the proposed FSN approach. Development and application of a flexible fault knowledge structure in qualitative manner is validated using historical plant data.

1.2 Plant Object Oriented Model (POOM)

This study proposes to use Plant Object Oriented Modelling (POOM) methodology developed by Gabbar et al [1], [2]. A POOM-based plant model is represented as building blocks of static model elements; each is associated with operation and behaviours where all related process variables are classified based on their function as manipulated, control or measured variables. A high-level example of POOM model where various plant Process Variables (PVs) is shown for a typical CANDU® reactor below in Figure 1.

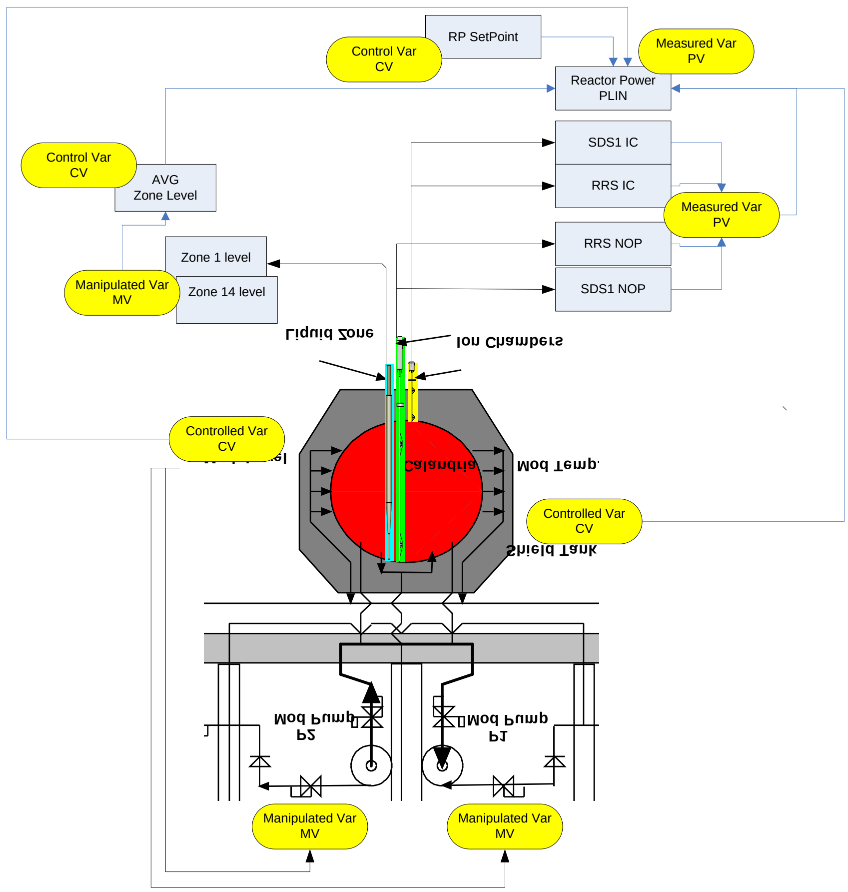


Figure 1. Process Variable (PV) interactions in Liquid Zone, Main Moderator Circuit and neutron detector systems are mapped and classified using POOM approach.

As shown above, there are many variables related to the plat operation, e.g. neutron detector current is a function of neutron flux, which in turn is a function of the current reactor power, moderator temperature, level, average and individual Liquid zone levels, etc. All process variables (PVs) identified in the POOM model are tabulated in one knowledge base. A portion of this structure is shown below along with the corresponding measurement instances.

Moderator				Calandria		Power	Liquid Zone Level			SDS1 NOP		RRS NOP	
Temp	Level	PUMP 1/2 Temp		Level	Inlet Pressure	RRS PLIN	AZL	Zone 6	Zone 8	Zone 6	Zone 8	Zone 6	Zone 8
X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
61.4652	8770.83	44.2737	34.4238	8786.71	211.985	0.92821	31.5078	23.1104	30.7178	105.919	104.484	91.8321	91.7043
61.4595	8771.17	44.2737	34.4238	8783.75	211.035	0.92821	32.0286	23.466	30.8106	105.977	104.639	91.8493	91.8619
61.4538	8771.5	44.2736	34.4238	8781.57	208.922	0.92821	32.0894	23.3114	30.2501	106.024	104.639	91.8664	91.7581
61.4481	8771.83	44.2736	34.4238	8780.78	210.037	0.92821	32.0926	22.805	29.62	105.962	104.508	91.8836	92.0426
61.4423	8772.17	44.2735	34.4238	8784.1	206.912	0.92821	32.1953	23.006	30.1496	106.022	104.402	91.9008	92.0426
61.4366	8772.5	44.2735	34.4238	8784.43	210.373	0.92821	31.9414	22.1633	30.8299	107.016	104.296	91.9065	92.1888
61.4309	8771.98	44.2734	34.4238	8789.59	212.127	0.92821	31.9609	23.0601	31.4909	106.566	104.19	91.9093	91.9427
61.4238	8771.47	44.2734	34.4238	8782.84	207.073	0.92821	31.9822	22.1942	30.8724	106.421	104.089	91.9121	91.9619
61.4132	8770.95	44.2734	34.4238	8784.38	209.779	0.92821	32.0107	22.2252	31.0348	106.578	103.998	91.9149	92.208
61.4026	8770.43	44.2733	34.4238	8790.24	212.42	0.92821	32.0391	22.5112	31.2899	107.167	104.597	91.9441	92.5464
61.3921	8769.91	44.2733	34.4238	8786.15	207.288	0.92821	32.4922	22.6929	32.4573	107.939	104.802	91.9846	92.3003
61.3815	8769.4	44.2732	34.4238	8786.93	207.501	0.92821	32.2266	23.6979	32.7936	107.45	104.736	92.0251	92.0619
61.371	8768.88	44.2732	34.4238	8784.28	207.284	0.92821	32.2976	23.4042	31.978	106.707	104.267	92.0657	91.6504
61.3604	8768.36	44.2731	34.4238	8784.2	207.067	0.92821	32.1058	23.841	31.0851	106.555	104.295	92.1062	91.8696

Table 1. Portion of PV Data set where 14 variables are shown with their corresponding ID tags, e.g. moderator temperature – X1, moderator level – X2, etc.

1.3 Fault, Failure, Hazard and Accident Classification

Faults can be caused by a variety of reasons starting with design deficiencies, ingress of moisture and corrosion due to assembly structure failure or inadequate or incorrect maintenance strategy. A sample fault to accident propagation scenario is shown below for a generalized neutron detection system. Ingress of moisture into the detector assembly may occur due to manufacturing defect, incorrect handling, storage or maintenance. This may lead to detector corrosion and connector or cable faults resulting in misleading readings. For extreme cases, faulty performance may lead to reduction or loss of coverage which, in turn, may lead to overheating of fuel in that particular zone.

Selection of methods for fault detection depends on the nature of the root cause, e.g. calibration drifts are detected by monitoring equipment while degradation of the detector itself is identified during functional testing. Next, rules are created and associated with each transition of the causation model within FSN. For example, failures related to corrosion might be associated with rules such as follows:

```
IF   (Structure.Material = (X or Y))
       and (PV= Gas.Pressure)
       and (Dev = Very-Low)
THEN (FM = Failure.detector)
```

These rules are initially defined in generic form based on domain knowledge, i.e. regardless of plant specific knowledge, and then further explained for the specific case. Formal language is proposed to represent process domain knowledge and safety control rules, as explained in [3], [4] in order to facilitate synthesis and validation of fault models within FSN.

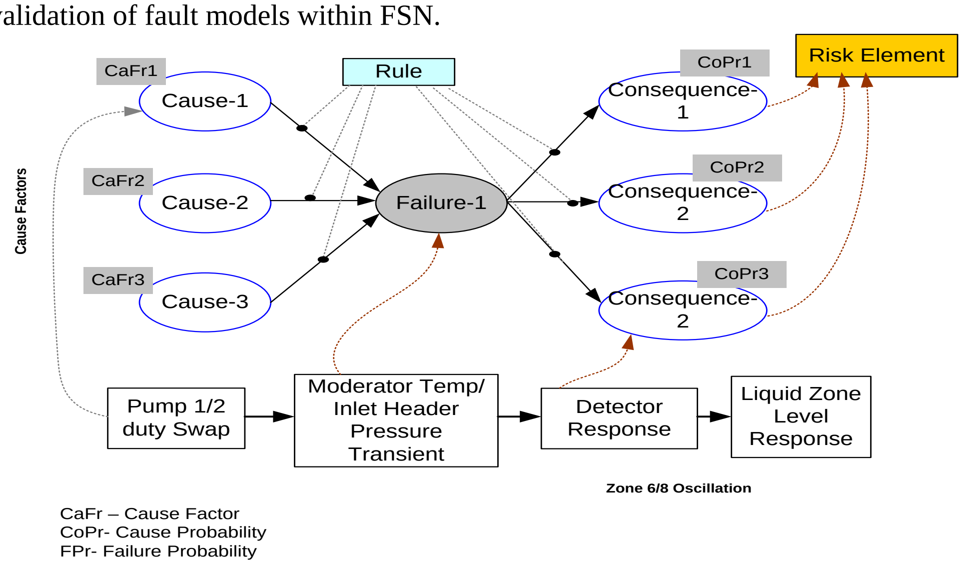


Figure 2. Fault Semantic Network shown for the selected case study [4].

Risk element is identified for each fault propagation scenario from the initiating events or root causes to the final consequences. For the example in Figure 3, there are three possible risk elements associated with consequence-1, namely:

```
[a] cause-1, failure-1, consequence-1;
[b] cause-2, failure-1, consequence-1;
[c] cause-3, failure-1, consequence-1.
```

where:

CaFr1 is causal factor used to define the frequency of cause-1.
FPr1 is the probability that failure-1 will occur due to any cause.
CoPr1 is the probability that consequence-1 will occur due to any cause & failure
CoIm1 is the total impact of consequence-1.

For independent events, total risk associated with consequence-1 is shown as:

$$\mathbf{R(Consequence\text{-}1) = [(CaFr1 + CaFr2 + CaFr3) * FPr1 * CoPr1] * CoIm1} \quad (1)$$

Total risk of consequence-2 and consequence-3 can be computed in a similar way. For case when events are co-dependent, Bayesian theorem can be applied to determine the total risk based on dependencies for cause-1, cause-2, and cause-3.

1.4 Implementation of FSN

Proposed FSN is comprised of two layers for static off-line and dynamic on-line modes. Static FSN includes faults, failures, hazards and accidents structured and linked in the form of causation models associated with process equipment. Dynamic FSN is constructed using dynamic simulated or real time data that can be obtained from operation, maintenance, safety, and control. Real time data are gathered for the selected case study and analysed in Matlab. CAPE-ModE [2] is developed within MS Visio to capture and structure process design models for Process Block Diagrams (PBD), Process Flow Diagram (PFD), and Piping and Instrumentation Diagram (P&ID), based on ISA-S95/88 [5]. ISA-S95 is an international standard for integration of enterprise, i.e. process, and control systems. It is widely used for standardization of a software layer used for information exchange, specification of user requirements, and functional requirements. Likewise, ISA-88 is a common standard for terminology and models used for batch control systems. Its main objective is to provide a hierarchical and modular categorization for the devices that carry out the process. ISA-88 defines physical model, which structures the plant hierarchically from the highest to the lowest level, e.g. Enterprise, Site, Area, Process cell, Unit, Equipment Module and Control module and is focused on the level of the Process Cell and the lower levels while ISA-95 is focused on the boundary between the Area and the Site. Thus, combination of these two standards forms the basis for managing the production process and ensures standardization within batch process automation.

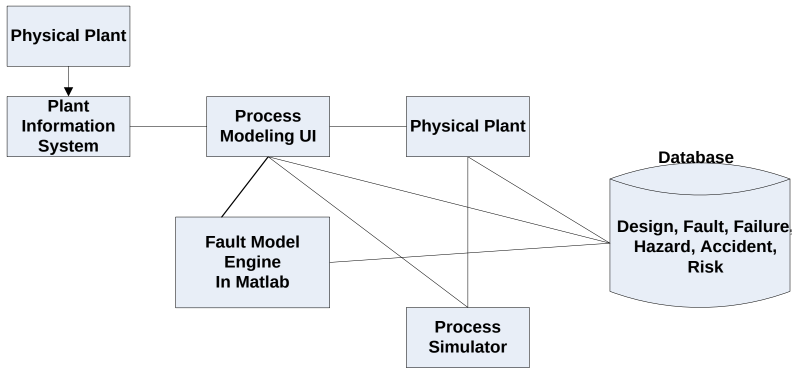


Figure 3. Proposed System Architecture for Failure, Fault, Hazard, and Accident Data Acquisition

Fault Diagnostic System (FDS) is developed to construct fault models using qualitative – quantitative, and deterministic and probabilistic techniques [6]. Fault modeling engine is developed within Matlab where fault propagation, equipment reliability, and material degradation are calculated and used to reconstruct and maintain fault/failure propagation models.

1.5 Knowledge Base Modelling

Based on POOM model of the detector system and surrounding process equipment shown in Figure 1, plant structure, behaviour, operation, control, and functional views/relations are defined in hierarchical manner [6]. Failure data from the selected case study are collected in excel file which includes two major sheets: header and details. Failures can be either of mechanical, electrical, or material nature. For example, corrosion of a detector is a mechanical failure due to failure of structural integrity of a detector assembly. Noise induction due to adjacent electrical equipment (e.g. moderator pumps) constitutes an electrical failure. An example of a material failure is ingress of moisture into the detector assembly causing contamination of the detector lead wire. Human failure or fault could also be represented in a similar way. Incorrect execution of a calibration procedure is a human error (or failure), the underlying cause of which could be lack of training or high level of fatigue due to work load or stress.

First a selected set of equipment classes is identified along with their parent class, for example "*Detectors*" and "*Sensors*" Each class has its associated identifier (*ID*) and description. Equipment class is related to function, component, and process variables. Failure and faults are represented at a generic and equipment specific levels. Failure Mode Identification (*FM-ID*) is used to represent both faults and failures. Resulting hazard and accidents are modeled in header and details entities and are linked with "*Lessons Learned*" and barriers, or controls, available at each step in the causation mode. These barriers could be used for prevention, detection or mitigation purposes and could comprise engineering, administrative, or operational activities.

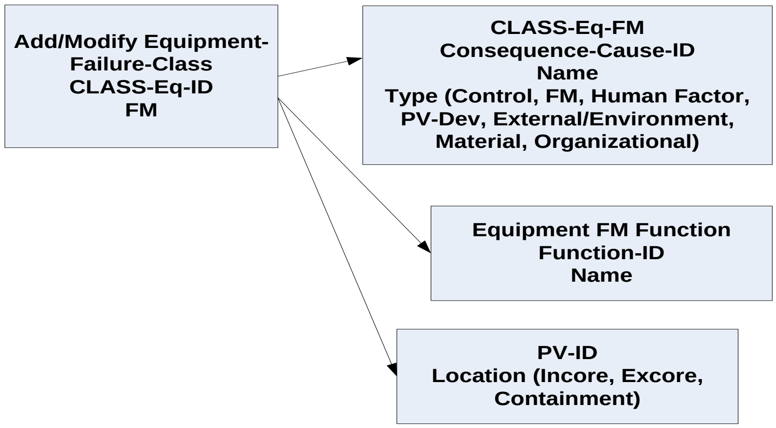


Figure 4. Data Acquisition Process Framework is shown.

2. Application of FSN for a Selected Case Study

2.1 Case Study – Bruce B Zone-Control Level Oscillations

A case of Bruce B zone-control level oscillations was selected for this study. Plant data for Unit-5 was reviewed dating back to 2008. A period of one hour on 20 February 2008, shown below, was selected for analysis.

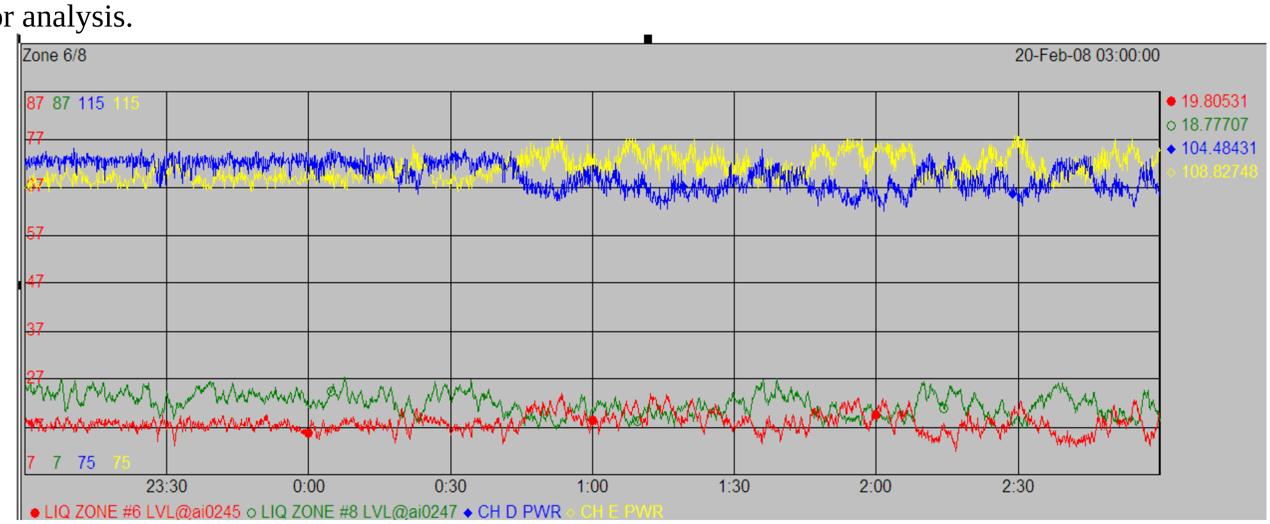


Figure 5. Bruce B Unit-5 data trend is shown below. Liquid Zone 6 and Zone 8 is shown in red and green, corresponding detector signals in those zones are shown in yellow and blue.

Detector responses for Zone 6 and Zone 8 are plotted in yellow and blue, with the onset of oscillations noticeable at approximately 00:30 and becoming particularly pronounced after 00:45 hours. Zone response is plotted in green and red and is noticeably following the same trend.

The root cause of this behaviour is unknown at this time. There are several potential causes that have been identified via analytical methods and engineering analysis.

2.2 Multivariate Analysis Techniques – Principle Component Analysis (PCA)

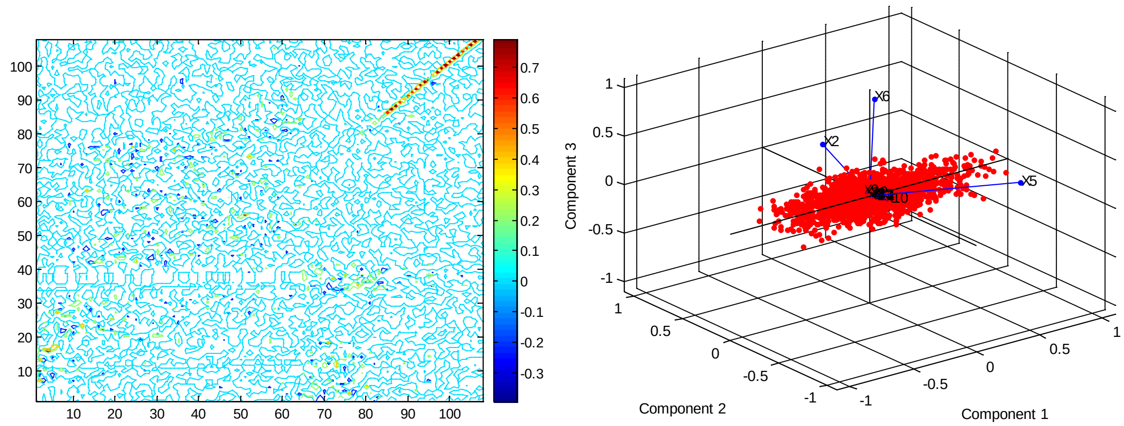
Due to the large number of parameters affecting the selected system performance, unravelling the relational pattern and connections among the data by analytical methods is challenging [7]. It is possible that more than one individual variable affects the selected detector behaviour at the same time, thus they should not be analysed independently. Traditional statistical data analysis methods focus on just one or two variables and are not suitable in this case. Multivariate Analysis (MVA) techniques appear to be a more preferred option. Two statistical methods, namely Primary Component Analysis (PCA) and weighted Principal Component Analysis are proposed for data extraction and pattern recognition in this case study.

Primary Component Analysis (PCA) is one of the oldest and most versatile methods of MVA. Historically, the bulk of applications of multivariate techniques have been in the behavioural and biological sciences [8] [9], [10], and [11], but have also been used in other applications, e.g. pharmaceutical industry for tablet development and manufacturing [12]. It's been commonly used in other industries, e,g, chemical, process, wastewater treatment, for dimensionality reduction for detecting and diagnosing faults [13].

A key application of PCA is to reduce the dimensionality of the problem in order to identify a potential root cause [8]. PCA algorithm decomposes a data table with correlated measurements into a new set of uncorrelated (i.e., orthogonal) variables, which are called primary components. Each unit is assigned a set of scores which show the magnitude of its effect on the rest of the components.

2.3 Application of PCA algorithm for Data Extraction and Pattern Recognition

For the first iteration of PCA algorithm 104 channelized measurements were obtained at 2 second sampling frequency resulting in *104x2200* dataset, which was converted into *104x104* principal component space as shown below in Figure 6 (a). This identified a challenge with traditional PCA analysis as it proved difficult to determine which specific component has the most weight.



(a) (b)

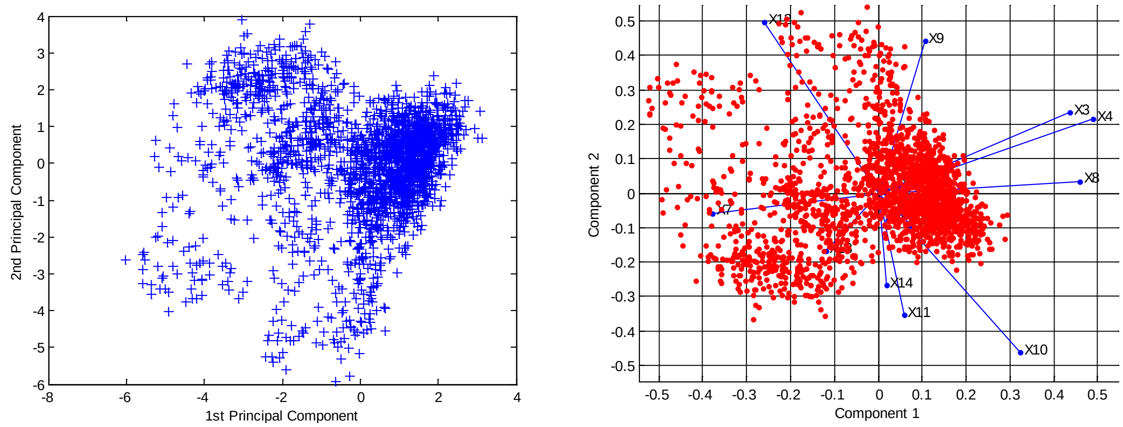
Figure 6. 104x104 principal component space after first PCA rendition of the sample dataset.

Since all principal components (PCs) are supposed to be orthogonal to each other, there should be no need for redundant information. Therefore, in order to address the challenge of representation all channelized measurements were removed from the historical sample of data. The time interval was shortened to 1 hour and the sample frequency was increased to 10 second. A second iteration of PCA algorithm was conducted on the modified dataset with results shown in Figure 6 (b).

A principal component calculated by PCA algorithm is a single axis in space. The second and the third principal components are other axis in space, perpendicular to the first. All the principal components, as a whole, form an orthogonal basis for the space of the data. Each of the instances of PV measurements is represented by a red point, and their locations indicate the score of each measurement for the two factors. The points are scaled to fit within the unit square, so only their relative locations can be determined from the plot. Process Variables are represented in this plot by blue vectors, and the direction and length indicate how each variable depends on the underlying factors. PCA algorithm identified process variables X2 and X6 corresponding to moderator temperature and inlet pressure, as having the most effect.

2.4 Weighted Principal Component Analysis – Improving Robustness

While PCA algorithm shows promising results, it is known to be less than optimal for large sets of data where data clusters and outliers are not identified in advance or their existence is unknown. A recently proposed generalization of PCA based on Weighted PCA (wPCA) increases robustness by assigning different weights to data objects based on their estimated relevancy [14]. The same general approach is used as in the PCA method; however the data is standardized by assigning weights to the mean and the outer products which form the covariance matrix. This puts a weight on every point in the training data and makes identification of extreme points easy in a large sample set. wPCA was performed on the selected data set by using the inverse variances of the ratings as weights, and scores for each principal component were calculated by transforming the original data into the space of the principal components.



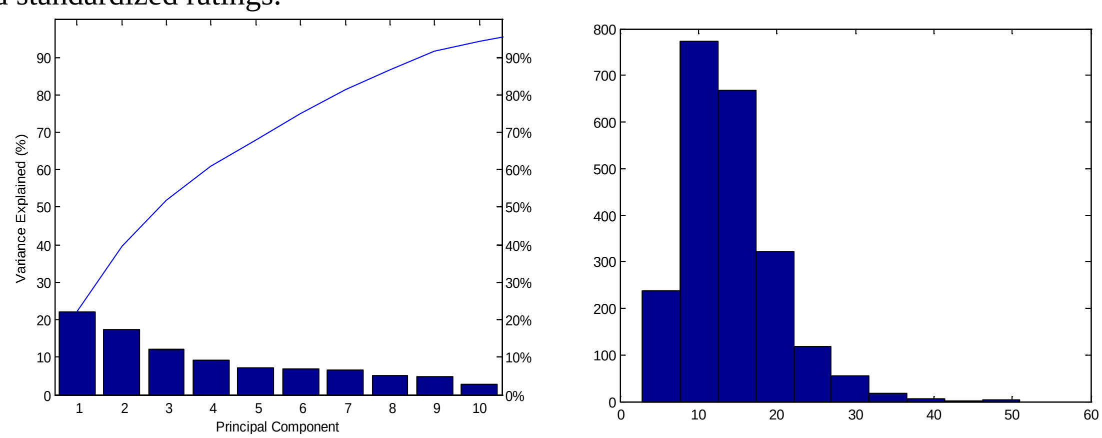
(a) (b)

Figure 7. Principal component space (a) and Weighted PCA results (b).

wPCA algorithm identified X3, X4, X8 and X9 vectors corresponding to moderator Pump Bearing Temperature for Pump1 and Pump 2 respectively. X8 corresponds to the average Liquid Zone level (AZL) measurement and X9 to Zone 6 level.

3. Discussion of Results and Conclusions

PCA method was used for 104x104 parameter matrix where each parameter had 2200 instances in order identify primary factors co-variant with zone-control level oscillations in the selected case study. A simple line segment plot, also called Pareto Scree plot, was created for the eigenvalues of a correlation matrix where the results are sorted in descending order of magnitude. This shows the fraction of total variance in the data as represented by each primary component. It can be noted that there is a large amount of variance is between the first and second components. It can also be noted that the first two principal components explain roughly two-thirds of the total variability in the obtained standardized ratings.



(a) (b)

Figure 8. Pareto Scree plot of variance distribution shows components that explain 95% of the total variance (a) and Hotelling's T^2 statistic.

This is consistent with X2 and X6 factors, corresponding to moderator temperature and calandria inlet pressure, identified by PCA algorithm. To determine the percent variability explained by each principal component, Hotelling's T-square was calculated. Hotelling's T-square provides is a statistical measure of the multivariate distance of each observation from the center of the data set, which allows identification of the most extreme points in the data [15].

Although, the initial rendition of PCA algorithm showed promising results, it also highlighted challenges of representation traditionally associated with statistical methods. Often, as in this case, graphical representation of results may be difficult to understand and/or require significant effort to translate into a format suitable for human recipients. In order to address this challenge, PCA method was adjusted to use weighted coefficients (wPCA) to improve robustness of the algorithm and the analysis was repeated for the same set of data. This study determined four (4) variables with the most-covariance, namely X3, X4, X8 and X9 corresponding to Pump Bearing Temperature for Pump 1 and 2, Average Zone Level and Zone 6 level.

Once these results were obtained, fault diagnosis and reconstruction of fault propagation scenario was conducted to validate the approach. First, the factors identified by weighted PCA algorithm were mapped with the plant Process Object Oriented model (POOM) as shown below.

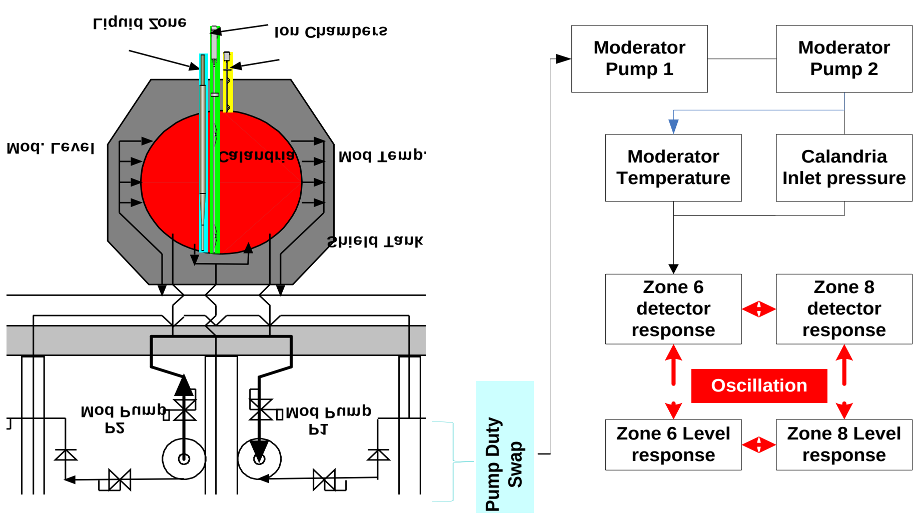


Figure 9. Fault propagation scenario based on wPCA results.

This shows that wPCA method identified fault initiating factors X3-X4, namely change in Moderator Pump bearing temperature, that lead to moderator temperature and/or calandria inlet pressure transient. This propagates into detector response and corresponding zone level oscillation. Engineering methods used for troubleshooting in the past showed that the oscillations occur simultaneously in Zone pairs 6/8 and 7/9 located on the North West (NE) side of the reactor. This is coincidental with the location of moderator Pump 2, shown in the POOM model of the plant. To prove this finding and validate wPCA results, zone-control level oscillations were plotted against the moderator pump bearing temperature data as shown below.

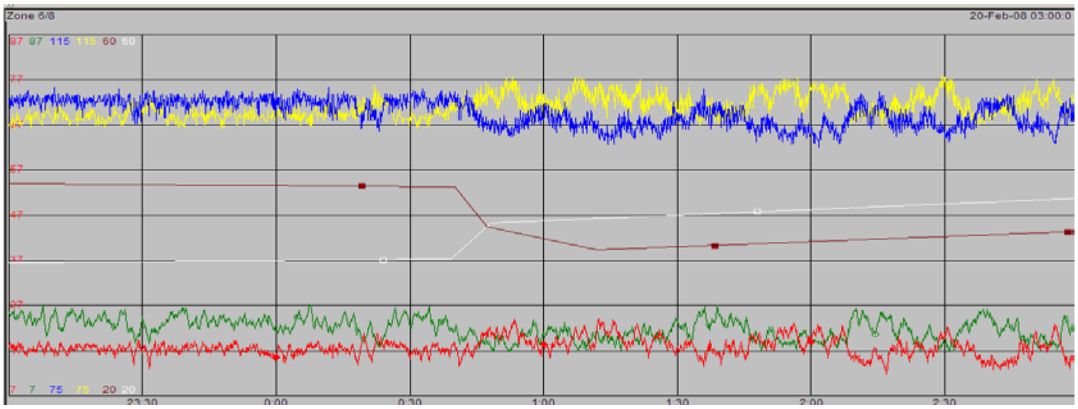


Figure 10. Bruce B NOP and Zone 6/8 oscillation plotted against moderator pump bearing temperature. Pump 1 shown in brown, Pump 2 shown in white.

Activation of Pump 2 is indicated by the increase in the pump bearing temperature; while Pump 1 temperature drop indicates that the pump was swapped out of duty. The onset of oscillations is coincidental with pump duty swap timestamp. The wPCA algorithm suggests that main moderator duty pump swap is intimately related to the cause of zone-control level oscillations in this case study.

4. References

[1] H. A. Gabbar, P.W.H., Chung, K. Suzuki, and Y. Shimada, "Utilization of unified modeling language (UML) to represent the artifacts of the plant design model", Proceedings of International Symposium on Design, Operation and Control of Next Generation Chemical Plants (PSE Asia), PS54, 2000, pp. 387-392.

[2] H. A. Gabbar , K. Suzuki, and Y. Shimada "Study on plant object-oriented model formalization - case study: HDS plant design", *Journal of Design Studies*, Vol. 24, No. 1, 2003, pp. 101-108.

[3] H.A. Gabbar, "Qualitative fault propagation analysis", *Journal of Loss Prevention in the Process Industries*, Vol. 20, No. 3, 2007, pp. 260–270

[4] H.E. Sayed, H. A. Gabbar, and S. Miyazaki, "Faults forecasting system", *World Academy of Science, Engineering and Technology*, Vol. 29, 2009, pp.1210-1216

[5] V.J. Pohjola, R. Koivisto, and M.K. Alha, "Conceptual association of safety with process design and operation", in D.W.T. Rippin, J.C. Hale, and J.F. Davis (Eds), "Foundations of computer-aided process operations", FOCAPO '93 Conference, Mount Crested Butte, Colorado, 18-23 July, 1993

[6] H. A. Gabbar, "Intelligent topology analyzer for improved plant operation", *Industrial Management + Data Systems*, Vol. 107. 2, 2007, pp.229

[7] A.C. Rencher, "Methods of Multivariate Analysis", Second Edition, John Wiley & Sons, Inc, 2002, pp. 727

[8] M. Scholz, M. Fraunholz, J. Selbig, "Nonlinear principal component analysis: neural network models and applications", Lecture Notes in *Springerlink Computational Science and Engineering* Volume 58, 2008, pp 44-67.

[9] U. Demšar, P. Harris, C. Brunsdon , A. S. Fotheringham and S. McLoone, "Principal component analysis on spatial data: an overview", *Routledge Annals of the Association of American Geographers*,

July 2012
[10] P. Sanguansat, Ed. "Principal component analysis - multidisciplinary applications", *InTech*, February 29, 2012, pp.212
[11] S. Yin, "On PCA-based fault diagnosis techniques", in Proceedings of IEEE Conference on Control and Fault-Tolerant Systems (SysTol), 6-8 Oct. 2010, pp. 179 - 184
[12] S. Zhou, J. Zhang, and S. Wang, "Fault diagnosis in industrial processes using principal component analysis and hidden Markov model", in Proceedings of American Control Conference Boston, Massachusetts June 30 - July 2, 2004, Vol.6, pp.5680 - 5685
[13] L.H. Chiang, E. L. Russell, R.D. Braatz, "Fault diagnosis in chemical processes using fisher discriminant analysis, discriminant partial least squares, and principal component analysis", *Chemometrics and Intelligent Laboratory Systems* Vol.50, No. 2, 13 March 2000, pp. 243–252.
[14] H.P. Kriegel, P. Kröger, E. Schubert, A. Zimek, "A general framework for increasing the robustness of PCA-based correlation clustering algorithms", Proceedings of 20th International Conference on Scientific and Statistical Database Management (SSDBM), Hong Kong, China, 2008, pp. 418 – 435.
[15] H. Hotelling, "The generalization of student's ratio", in *The Annals of Mathematical Statistics*, 1931 Vol.2, No.3, pp. 360-378.