공학박사학위논문

Realtime Process Monitoring and Fault Propagation Path Estimation using Principal Component Analysis and Granger Causality

주성분분석법과 그레인저 인과관계를 이용한 실시간 공정 모니터링 및 이상 전파 경로 계산

2017년 4월

서울대학교 대학원

화학생물공학부

하 대 근

Abstract

Real-time Process Monitoring and Fault Propagation Path Estimation using Principal Component Analysis and Granger Causality

Daegeun Ha

School of Chemical & Biological Engineering

The Graduate School of Seoul National University

Modern industrial process is a complex device industry consisting of a combination of numerous unit processes. Numerous process parameters such as flow rate, temperature, pressure, concentration and composition have strong linear or nonlinear correlation. Since improvement of computing power and process control systems in industrial processes, several board operator and field operator can manage huge amounts of data and whole process information from industrial plant. However, the number of processes and devices to be handled by a single operator will increase, and operators meets a limitation of cognitive ability due to flood of information,

causing problems such as process malfunction or instrumental failure. To solve this problem, we propose a PCA modeling procedures that aims to improve monitoring performance by variable selection, removing noise, operation mode classification and mode change detection. Fault diagnosis and causal analysis is also introduced. We calculated the causal relationship matrix between the process variables and find out the root cause of the unexpected process changes. The proposed approach was applied and validated to LNG plant located in Incheon and plasma condition monitoring in plasma etcher.

Chapter 2 discusses the application methodologies of signal processing to eliminate noises from OES signal and multivariate statistical techniques to improve monitoring sensitivity. Among the plasma sensors, optical emission spectroscopy (OES) has been widely utilized and its high dimensionality has required multivariate analysis (MVA) techniques such as principal component analysis (PCA). PCA, however, might devaluate physical meaning of target process during its statistical calculation. In addition, inherent noise from charge coupled devices (CCD) array in OES might deteriorate PCA model performance. Therefore, it is desirable to pre-select physically important variables and to filter out noisy signals before modeling OES based plasma data. For these purposes, this chapter introduces a peak wavelength selection algorithm for selecting physically meaningful wavelength in plasma and discrete wavelet transform (DWT) for filtering out noisy signals from a CCD array. The effectiveness of the PCA model introduced in this paper is verified by comparing fault detection capabilities of conventional PCA model under the various source power or pressure faulty situations in a capacitively coupled plasma

etcher. The PCA model introduced in this chapter successively detect even extremely small variation such as 0.67% of source power change even though the conventional PCA model fails to detect all of the faulty situations under the tests.

Chapter 3 discusses the application methodology of operation mode identification and multimode PCA to improve the performance of LNG mixed refrigeration (MR) process and prevent process shutdown. LNG MR process is usually used for liquefying natural gas. The compressors for refrigerant compression are operated with the high-speed rotating parts to create a high-pressure. However, any malfunction in the compressors can lead to significant process downtime, catastrophic damage to equipment and potential safety consequences. The existing methodology assumes that the process has a single mode of operation, which makes it difficult to distinguish between a malfunction of the process and a change in mode of operation. Therefore, k-nearest neighbor algorithm (k-NN) is employed to classify the operation modes, which is integrated into multi-mode principal component analysis (MPCA) for process monitoring and fault detection. When the fault detection performance is evaluated with real LNG MR operation data, the proposed methodology shows more accurate and early detection capability than conventional PCA.

Chapter 4 discusses PCA based fault amplification algorithm to detect both the root cause of fault and the fault propagation path in the system. The developed algorithm project the samples on the residual subspace (RS) to determine the disturbance propagation path. Usually, the RS of the fault data is superimposed with

the normal process variations which should be minimized to amplify the fault magnitude. The RS containing amplified fault is then converted into the co-variance matrix followed by singular value decomposition (SVD) analysis which in turn generates the fault direction matrix corresponding to the largest eigenvalue. The fault variables are then re-arranged according to their magnitude of contribution towards a fault which in turn represents the fault propagation path using an absolute descending order functions. Moreover, the multivariate granger causality (MVGC) algorithm is used to analyze the causal relationship among the variables obtained from the developed algorithm. Both the methodologies are tested on the LNG fractionation process train and distillation column operation where some fault case scenarios are assumed to estimate the fault directions. It is observed that the hierarchy of variables obtained from fault propagation path algorithm are in good agreement with the MVGC algorithm. Therefore, fault amplification methodology can be used in industrial systems for identifying the root cause of fault as well as the fault propagation path.

The application results show that the proposed multivariate statistical method can improve productivity and safety by providing useful information for process monitoring and fault diagnosis in various processes with distributed control system.

Keywords: Process monitoring, Fault diagnosis, Mode identification, Granger causality, Principal component analysis, *k*-NN classification

Student ID: 2013-30990

Contents

List of Figures

List of Tables

CHAPTER 1 Introduction

1.1 Research motivation

With an increase in the complexities in the industrial systems and high level process integrations allows the industry to obtain enormously huge amount of information on the process operation including thousands of sensors, analyzers and control loops have strong interactions among each other in terms of process variables. Therefore, in order to efficiently manage such complex processes, detect the instrumental malfunctions are required to improve product yield, reduce operational risk, ensure safe operation and achieve sustainable profit.

A fault is defined as a departure from an acceptable range of an observed variable or an unpredicted deviation of at least one characteristic property or variable of the system. (Chiang et al. 2001). The underlying causes of this abnormality, such as a failed compressor motor operation or controller failure, are called the basic events or root causes. The four procedures associated with process monitoring are: fault detection, fault identification, fault diagnosis, and process recovery. There appears to be no standard terminology for these procedures as the terminology varies across disciplines; the terminology given by Raich and Cinar (1996) is adopted here.

- Fault detection: determine whether a fault has occurred. Early detection may provide invaluable warning on emerging problems, with appropriate actions taken to avoid serious process upsets.

- Fault identification: identify the observation variables most relevant to

diagnosing the fault. The purpose of this procedure is to focus the plant operator's and engineer's attention on the subsystems most pertinent to the diagnosis of the fault, so that the effect of the fault can be eliminated in a more efficient manner.

- Fault diagnosis: determine which fault occurred, or determining the cause of the observed out-of-control status. The fault diagnosis procedure is essential to the counteraction or elimination of the fault.

- Process recovery: remove the effect of the fault.

When a fault is detected, the fault identification, fault diagnosis, and process recovery procedures are employed in the respective sequence; otherwise, only the fault detection procedure is repeated.

Process monitoring and fault diagnosis in industrial plant are significant task for improve efficiency and ensure the safety. When fault occur in industrial plant, the operator must quickly detect the fault, diagnose its root causes of abnormal events as soon as possible with ensure accuracy so corrective action can be taken in a timely manner. Due to the varying operational conditions, process and quality variables need to be monitored continuously to ensure a reliable and efficient operation and, thus, day average values are not sufficient to get early detections or warnings of deviating or abnormal conditions. Consequently, this calls for techniques to handle large data sets online. The methods for monitoring used today are normally based on time series charts, where the operator can view the different variables as historical trends.

However, it is very difficult for operators to take correct actions when facing faults because of complexity and integrity of modern industrial processes. Therefore, automatic fault detection and diagnosis systems for industrial plants have become extremely required to support operator or engineer's decision making for process control and monitoring.

To solve this problem, multivariate statistical techniques are powerful tools capable of compressing data and reducing its dimensionality so that essential information is retained and easier to analyze than the original huge data set; and they are also able to handle noise and correlation to extract true information effectively. PCA method, initially proposed by Pearson (1901) and later developed by Hotelling (1947), is a standard multivariate technique and has been included in many textbooks (Anderson, 1984; Jackson, 1991) as well as research papers (Wold, 1978; Wold, Esbensen, & Geladi, 1987). The main function of multivariate statistical techniques is to transform a number of related process variables to a smaller set of uncorrelated variables. PCA method is due to their ability to efficiency handle a large number of highly correlated variables, measurement errors, and missing data caused by sensor fault. More important is their ability to provide an operator with useful information on process improvement through projecting the high dimensional process data into the low dimensional space defined by a few latent variables.

Proposed process monitoring and fault diagnosis system is required to data collection, data preprocessing, analyze on-line process data, process trend monitoring, multivariate statistical process monitoring, fault diagnosis and information visualization.

1.2 Research objectives

The objective of this thesis is to suggest a real-time process monitoring and fault diagnosis of various industrial processes including LNG manufacturing process and semiconductor etching process based on the MSPC approach and causal relationship approach. The thesis deals with several real plant operation cases and achieve more accurate process monitoring and fault diagnosis techniques which is required to improve product efficiency, reduce operational risk, ensure safe operation. The detail objectives to achieve the principal goal are considered as follows: (i) Analyze the types of noise sources that occur due to process and device characteristics, and applies a filtering technique to remove noise; (ii) Adapt and combine the well-known multivariate statistical analysis model considering the target process dynamics and operation strategy; (iii) Develop extensions to existing modeling techniques to improve simplicity and accuracy of models for on-line process monitoring; (iv) Develop methodology of fault diagnosis and causal relationship model for plant-wide root cause diagnosis and fault propagation using the developed models.

1.3 Outline of the thesis

The thesis is organized as follows. Chapter 1 describes necessity of multivariate process monitoring and fault diagnosis with an introduction of this study and outline of this thesis. Chapter 2 describes noise reduction technique using discrete wavelet transform, automatically signal selection methodology and principal component analysis for OES data dimension reduction and plasma state monitoring. Chapter 3 describes a methodology of fault detection method using process mode classification and identification technique and multimode PCA to LNG Mixed Refrigeration Process. Chapter 4 describes a methodology of fault diagnosis and causal analysis using a statistical model for a LNG fractionation process. Further, in this chapter, we suggested MVGC methodology to find root cause of process malfunction.

CHAPTER 2 : Multivariate monitoring, variable selection and OES signal filter design of plasma process[1]

2.1 Introduction

With an advent of sub-20nm device era, Moore's law that predicts the number of transistor in an integrated circuit doubles approximately every sixteen months becomes difficult to follow. The reasons are physical difficulties to shrink gate width physically and electrically. Instead, the concept of more than Moore's law which proposes an equivalent scaling approach rather than shrinkage technology has become prevalent in semiconductor industry. (Kahng, 2010; Ramm et al., 2010; Waldrop, 2016) The equivalent scaling approach has three main directions such as introduction of innovative process schemes like 3D transistor, additional chip and system-level architectural design, and cost-effective manufacturing.

As a cost-effective manufacturing approach, plasma etching, which is one of the major processes in semiconductor manufacturing, has developed real-time monitoring and control techniques. (Gaman et al., 2015; Keville et al., 2013; Lynn et al., 2012) Specifically, inherent complexity of plasma has driven this activity to mostly focus on development of plasma sensors and their applications. (Baek et al., 2005; Booth et al., 2000; Klick et al., 1997; Lee et al., 2007; Oh et al., 2010; Sobolewski, 2006)

[1] The partial part of this chapter is taken from the author's published paper.

Among plasma sensors, optical emission spectroscopy (OES) which measures emission from plasma and analyzes chemical composition in plasma has been widely utilized in plasma etching for endpoint detection and process monitoring. (Litvak et al., 1996; White et al., 2000; Yue et al., 2000; Yue et al., 2001) Since OES detects light from ultra violet (UV) to visible ranges with less than 1nm spectral resolution, size of data generated during a typical plasma etching process is larger than 1 megabyte and the number of variable to consider is around 2048. Therefore, a multivariate analysis (MVA) technique such as principal component analysis (PCA) has been studied for dimension reduction and variable selection. (White et al., 1997; White et al., 2000; Yue et al., 2000; Yue et al., 2001)

PCA, however, might devaluate physical meaning of target process during its statistical calculation. In addition, inherent noise from charge coupled device (CCD) array in OES might deteriorate PCA model performance. For these reasons, it is desirable to pre-select physically important variables and to filter out noise signals in modeling OES based plasma data.

This paper introduces improvement of PCA performance by developing a pre-variable selection technique and by applying discrete wavelet transform (DWT) in modeling OES based plasma data. In Section 2-2, issues in PCA modeling of OES based plasma data are described especially in terms of huge data size and inherent noise. In Section 3, theoretical background of developed techniques in this paper is explained and how they are integrated into PCA is described. After a brief description of experimental conditions in Section 4, performances of conventional and this paper-introducing PCA model are compared under various test conditions in Section 5.

2.2 Issues in PCA Modeling of OES based Plasma Data

Figure 2-1 shows a schematic of a plasma reactor and OES installation in this paper. Emission in plasma is transported to an entrance slit in an OES system via an optical fiber which is attached on the viewport of plasma etcher. The incoming light at the entrance slit in OES system travels to the collimating mirror where the light is reflected to a diffraction grating. The diffraction grating then spreads the light to the focusing mirror where the spread light is transported to a CCD array. The CCD array which is assigned as each wavelength convert light to electrical signal.

In current plasma etch processes, a high resolution OES which can measure wider range of spectrum ranging from 150nm to 1,000nm with less than 0.8 spectral resolution is employed for process monitoring. Accordingly, huge data from the high resolution OES makes it difficult to manually identify useful features and key wavelengths. Therefore, these high volumes of OES data in plasma etching requires to develop robust and automated data reduction, feature extraction and analysis techniques.

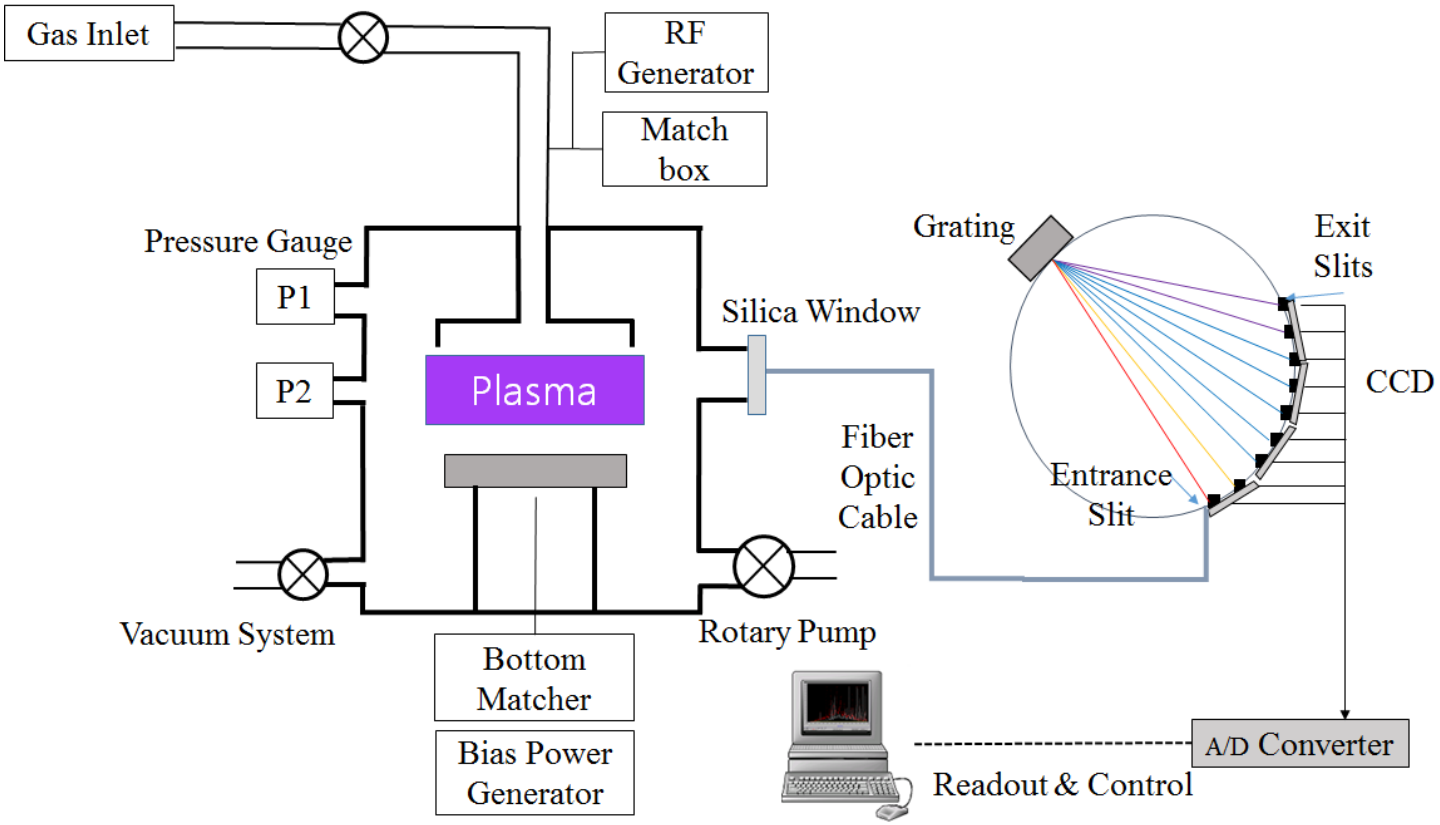

Figure 2-1. Schematic of a plasma reactor and an OES installation, in which inside of OES system is also described

Principal component analysis (PCA) has been applied for those purposes in plasma etching. (Ragan et al., 1997; White et al., 1997; Yue et al., 2000) PCA, however, might devaluate physically important parameters during its statistical calculation. Specifically, in consideration of the previous results that core plasma parameters can explain plasma etch processes well, it is desirable to include plasma variables in PCA. (Baek et al., 2014; Park et al., 2015) Thus variable selection techniques which can extract physically meaningful wavelength from OES data should be incorporated into PCA in modeling plasma etching processes.

In addition, inherent noise from a CCD array in OES system might deteriorate PCA model performance. There are several types of noise associated with a CCD array. (Ma et al., 2010; Goodlin, 2002) Firstly, photon shot noise arises from quantum detection of photons. Secondly, spontaneous photoelectron in the absence of photons is generated, which is called as dark current. Readout noise is also generated through electronic processing of signal. Lastly, there is pixel shift and drift, which is change in the location of the pixel as a function of time, possibly due to subtle mechanical vibrations.

Among these noise sources, the readout noise and the pixel shift and drift can be significantly reduced by optimum design of a CCD array. However, other existing noise such as shot noise and dark current can frequently deteriorate results of signal processing. Therefore, it is desirable to separate this noise effect in modeling.

2.3 Theoretical Background

2.3.1 Peak Wavelength Selection Algorithm

Sharp emission in plasma is the result of electron transition from excited to lower energy states within atom. Considering each atom has its own energy state, wavelength of sharp emission can estimate which atom exists in plasma. The intensity of the sharp emission can also estimate how much the atom exists in plasma.

A peak wavelength selection algorithm is devised to automatically search wavelength of peak emission in OES spectrum. Figure 2-2 shows a flow chart of the peak wavelength selection algorithm. Firstly, threshold value, α, is specified in consideration of signal to noise ratio of an employed OES system. Here, α is 300, which comes from the signal-to-noise specification of the employed OES. Then, emission intensities of three consecutive wavelengths are compared. If the emission intensity of wavelength in the middle is higher than both that of wavelength in the left side and that of wavelength in the right side to the extent of more than threshold α, the wavelength is selected temporarily.

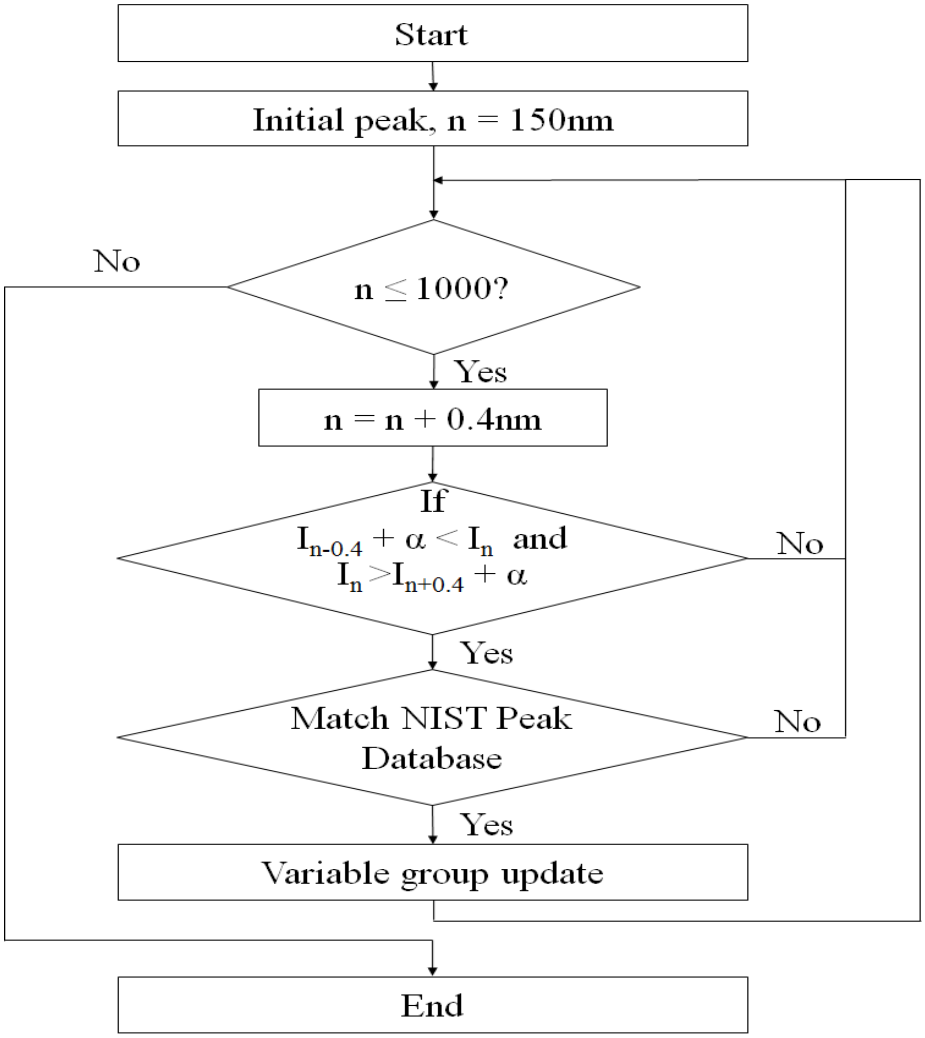

Figure 2-2. Flow chart of a peak wavelength selection algorithm

The temporarily selected wavelength is then compared with technical database from NIST (http://physics.nist.gov/PhysRefData/Handbook/element_name.htm) to check whether it is physically meaningful or not. If it is matched, then the wavelength is finally selected for PCA modeling. This process is done from the starting wavelength, 150nm, to the ending wavelength, 1000nm, sequentially.

2.3.2 **Discrete Wavelet Transform**

As described in Section 2-2, OES signal is composed of major information from plasma and several minor noise from a CCD array. Therefore, it is desirable to decompose OES data into major and minor signals before PCA modeling.

For this purpose, discrete wavelet transform (DWT) can be useful because it can decompose a signal in both time and frequency domain through variable window function. The variable window function uses a shorter time interval to analyze the high frequency components of a signal and a longer one to analyze the low frequency components of the signal. Many studies also utilized this characteristics of wavelet transformation for the purpose of improving performance of sensor fault detection and diagnosis. (Chen et. al., 2006; Ding et. al., 2003; Lada et. Al., 2002; Wang et. al., 2005; Xu et. al., 2008; Zuppa et. al., 2007)

According to Mallat's theory of multi-resolution analysis (Mallat, 1989), any signal $f(t) \in L^2(R)$ can be represented at multiple scales by decomposition on a family of scaling functions $\Phi_{m,n}(t)$ and wavelet functions $\psi_{m,n}(t)$ shown in Eq. (1).

$$\mathrm{f(t)} = \sum_{n\in Z} c_{m_0,n}\phi_{m_0,n}(t) + \sum_{m=m_0}^{\infty}\sum_{n\in Z} d_{m,n}\psi_{m,n}(t) \qquad (1)$$

where,

$$\phi_{m,n}(t) = 2^{-m}\phi\left(2^{-j}t - n\right),$$

$$\Psi_{m,n}(t) = 2^{-m/2}\psi\left(2^{-j}t - n\right),$$

$$c_{m_0,n} = \int_R f(t)\phi_{m_0,n}(t)dt,$$

$$d_{m,n} = \int_R f(t)\psi_{m,n}(t)dt.$$

When the signal is projected onto the scaling functions, one set of coefficients $c_{m_0,n}$ known as approximation coefficients is obtained. The approximation coefficients describe the trend or approximation of the signal. On the other hand, when the signal is projected onto the wavelet functions, another set of coefficients $d_{m,n}$ known as wavelet coefficients is obtained. The wavelet coefficients describe the details of the signals.

Figure 2-3 shows decomposition of an OES signal by the wavelet transformation. The original OES signal S is decomposed into approximation cA_1 and detail cD_1 coefficients at the first level. Application of the same transform on the approximation cA_1 causes it to be decomposed further into approximation cA_2 and details cD_2 coefficient at the second level. The decomposition process can continue to a level L as long as the length of approximation coefficients in cA_L is more than the length of coefficients in the wavelet filter.

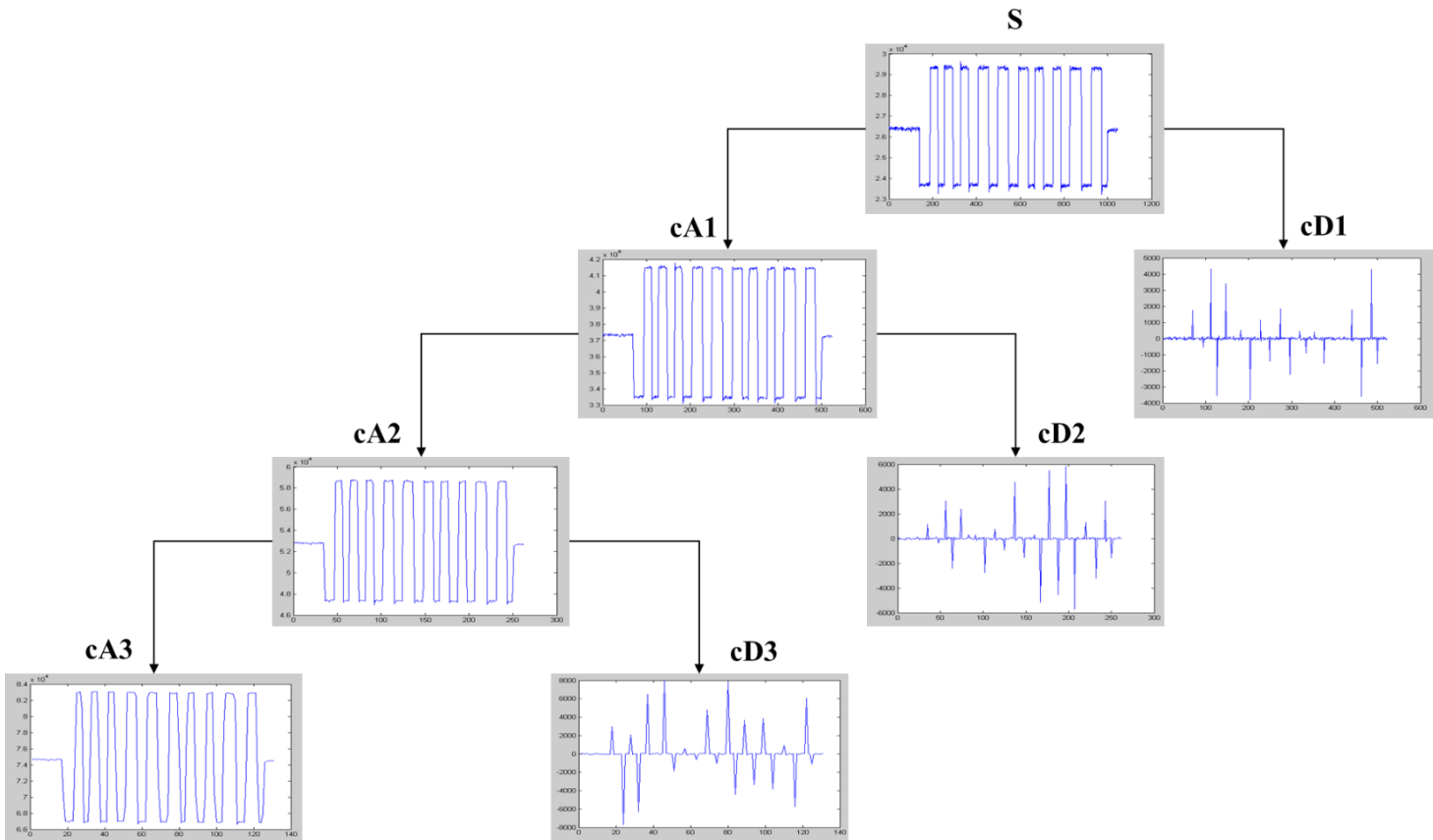

Figure 2-3. Results of DWT of OES signal in which 3 level decomposition

At this stage, there are details coefficients, cD_1, cD_2, and cD_3, and approximation coefficients, cA_3. The optimum level L is determined by checking a signal to noise ratio (SNR) and a root mean square error (RMSE) described in Eqs. (2) and (3).

$$\mathrm{SNR} = \frac{A}{\sigma} \tag{2}$$

where,

A: average of selected OES peak,

σ: standard deviation of selected OES peak.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \breve{y}_j\right)^2} \tag{3}$$

where,

n: number of samples,

y_j: modified OES signal after DWT filtering,

$\breve{y}_j$: original OES signal's average.

In OES signal case, the SNR are increased as the level increases from one to three and are saturated at the level 4 in Figure 2-4. Similarly, the RMSE are decreased as the level increases from one to three and are saturated at the level 4. Thus, level 3 is determined as an optimum.

As a result, OES signal can be decomposed into approximation coefficients and detail coefficients by discrete wave transformation. That is, noise from a CCD array are represented by detail coefficients and plasma conditions from process variable changes are represented with approximation coefficients.

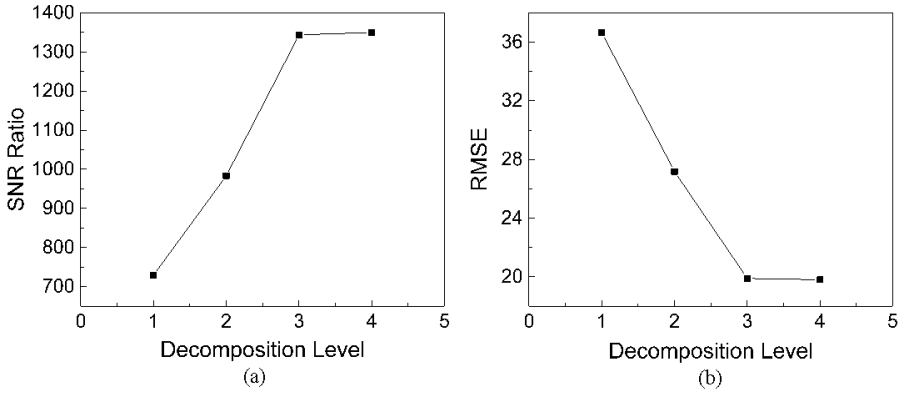

Figure 2-4. SNR (a) and RMSE (b) for Each Level after DWT decomposition

2.4 Experimental Set-up

Figure 2-5 shows a plasma etching reactor employed in this paper. It is a capacitively coupled plasma reactor for 300mm wafer, which is powered by three radio frequency (RF) generators such as 60MHz on top electrode, 13.56MHz and 2MHz on bottom electrode. The gap size between top and bottom electrodes is 25mm and the area ratio of top showerhead and bottom electrodes is 1.33. The employed optical emission spectrometer in this paper is capable of measuring light ranging from 150nm to 1000nm wavelengths with 0.4nm spectral resolution.

The reference plasma condition is 20mT of pressure, 300W of 60MHz RF power, 400sccm of Ar flow rate, and 16sccm of SF_6 flow rate. With the regard to the reference plasma condition, 60MHz RF power or pressure is changed intentionally to generate faulty situations.

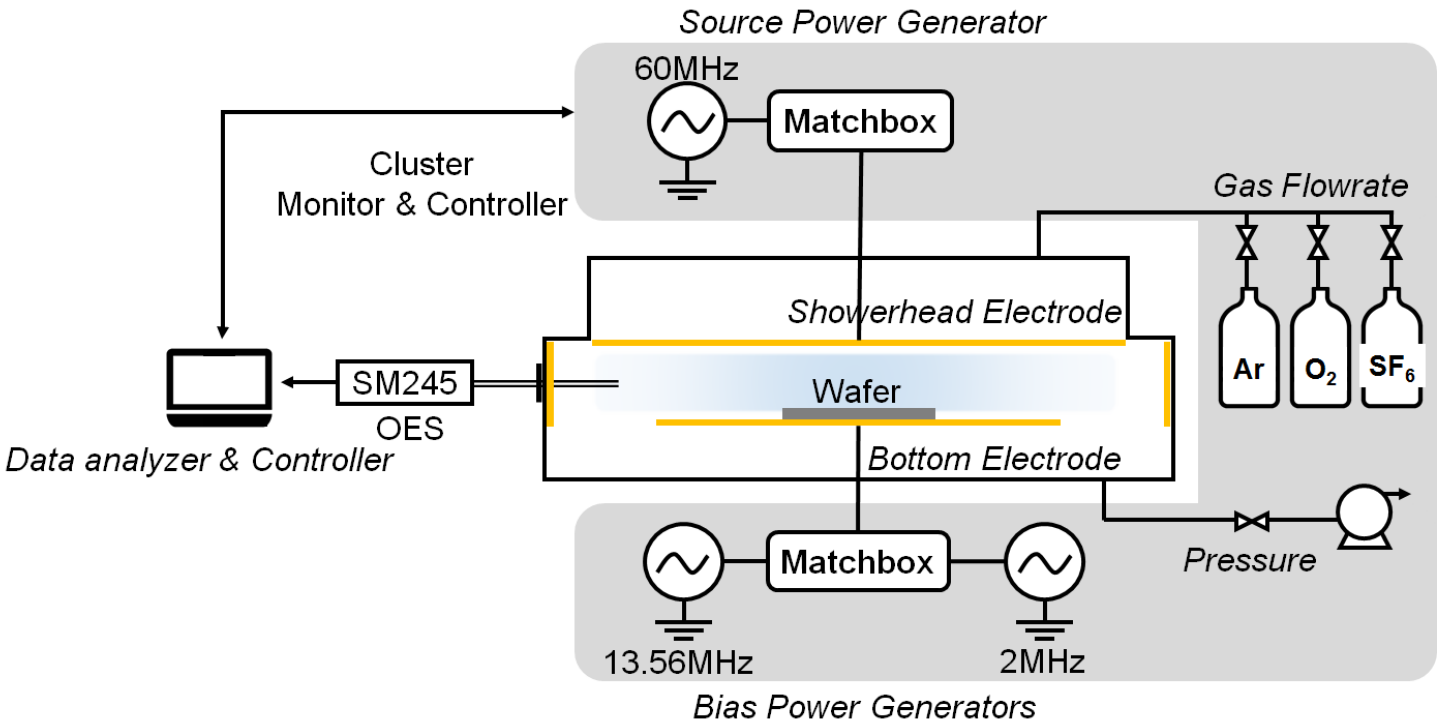

Figure 2-5. Schematic diagram of a plasma etching reactor

2.5 Results and Discussion

2.5.1 Pre-selected variables in OES data

Figure 2-6 shows reduction of OES spectrum data before and after applying peak wavelength selection algorithm. As described in Section 2.3.1, peak wavelengths with higher signal to noise ratio, α, are selected among neighbor wavelengths. Then, the wavelengths are matched with technical database of Ar and SF_6 from NIST. As a result, the number of wavelength variables is reduced from 2048 to 20.

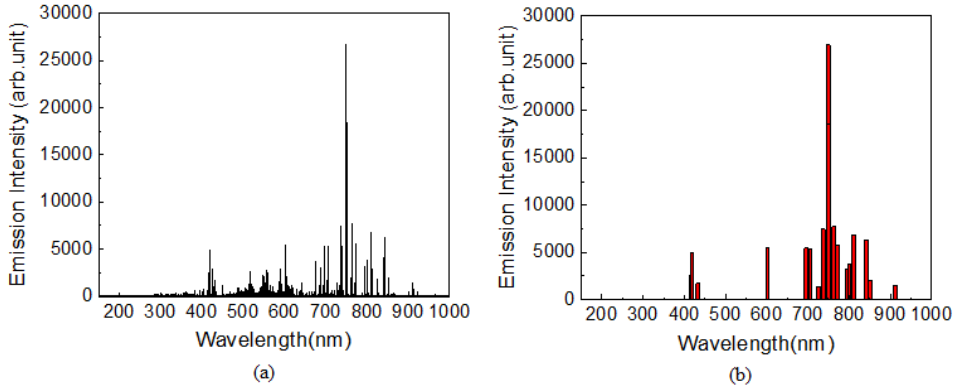

Figure 2-6. OES spectrum data before after applying peak wavelength selection algorithm

2.5.2 **Decomposition of OES signal by DWT**

In Section 2.5.1, it is described that one of issues in OES signal analysis, high dimensionality, can be mitigated by applying the peak wavelength selection algorithm. The selected signals are then decomposed by DWT in order to filter out noisy signals from a CCD array. Considering the fact that frequencies from OES measurement and those from noise in a CCD array are different, the threshold of detail coefficient is adjusted so as to separate noise in a CCD array.

Figure 2-7 shows the time resolved signal of Ar emission at 750nm wavelength before and after DWT. For the DWT filter design, soft threshold is applied and the setting values for the detailed coefficient are 105.48 for level 1, 99.636 for level 2, and 118.28 for level 3. It is clearly shown that the noise from the CCD array is excluded from the original signal in Figure 2-7-(b). Likewise, the other signals are filtered out by DWT before PCA modeling.

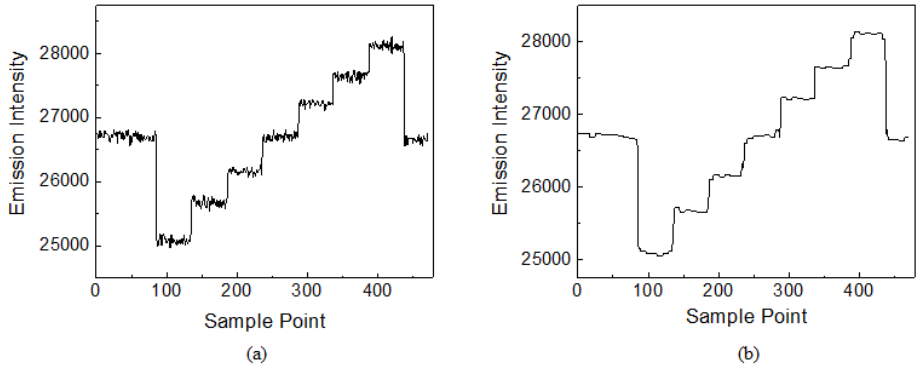

Figure 2-7. Time resolved signal of Ar emission at 750nm wavelength before and after DWT

2.5.3 Comparison of Fault Detection Performance in OES based PCA Models

To evaluate performance of OES based PCA models, various Ar + SF_6 plasmas are tested in the CCP etcher as shown in Figure 2-5. Based on the reference Ar + SF_6 plasma with which a reference PCA model is built, source power (60MHz RF power) or pressure is changed intentionally to see the detection capability of PCA models.

Tested PCA models are summarized in Table 2-1. The conventional PCA model is built with all wavelength variables in OES spectrum. On the contrary, the P-PCA model where the peak wavelength selection algorithm is applied and the PD-PCA model where the peak wavelength selection algorithm and DWT are applied.

Table 2-1. Summary of tested PCA models in this study.

PCA Model	No. of Variables	Noise Filtering by DWT	No. of PC to Capture 70% Variance	Upper Limit of Hotelling's T^2 with 99% Confidence Level
Conventional PCA	2048	No	26	90.8
P-PCA[1]	20	No	9	32.2
PD-PCA[2]	20	Yes	4	18.9

[1]P-PCA: peak wavelength selection algorithm is applied.

[2]PD-PCA: peak wavelength selection algorithm and DWT are applied.

Figure 2-8 shows Hotelling's T^2 for pressure fault cases where 5% or 10% of pressure with regard to the reference condition is increased. As is shown in Figure 2-8 (a), the conventional PCA model fails to detect any pressure faults in Hotelling's T^2. In contrast, the P-PCA and PD-PCA models detect both pressure faults like that shown in Figure 2-8(b), (c). As a result, missed detection rates for each case in Table 2-2 are 1, 0 and 0 for the conventional PCA, the P-PCA and PD-PCA, respectively.

Figure 2-9 shows SPE for pressure fault cases. Different from the Hotelling's T^2, all of the PCA models detect all faults in SPE. This might come from the fact that some of Ar emission peak is contained in residual space because captured variance of all PCA models is 70%. As a result, missed detection rates for each case in Table 2-3 are 1, 0 and 0 for the conventional PCA, the P-PCA and PD-PCA, respectively.

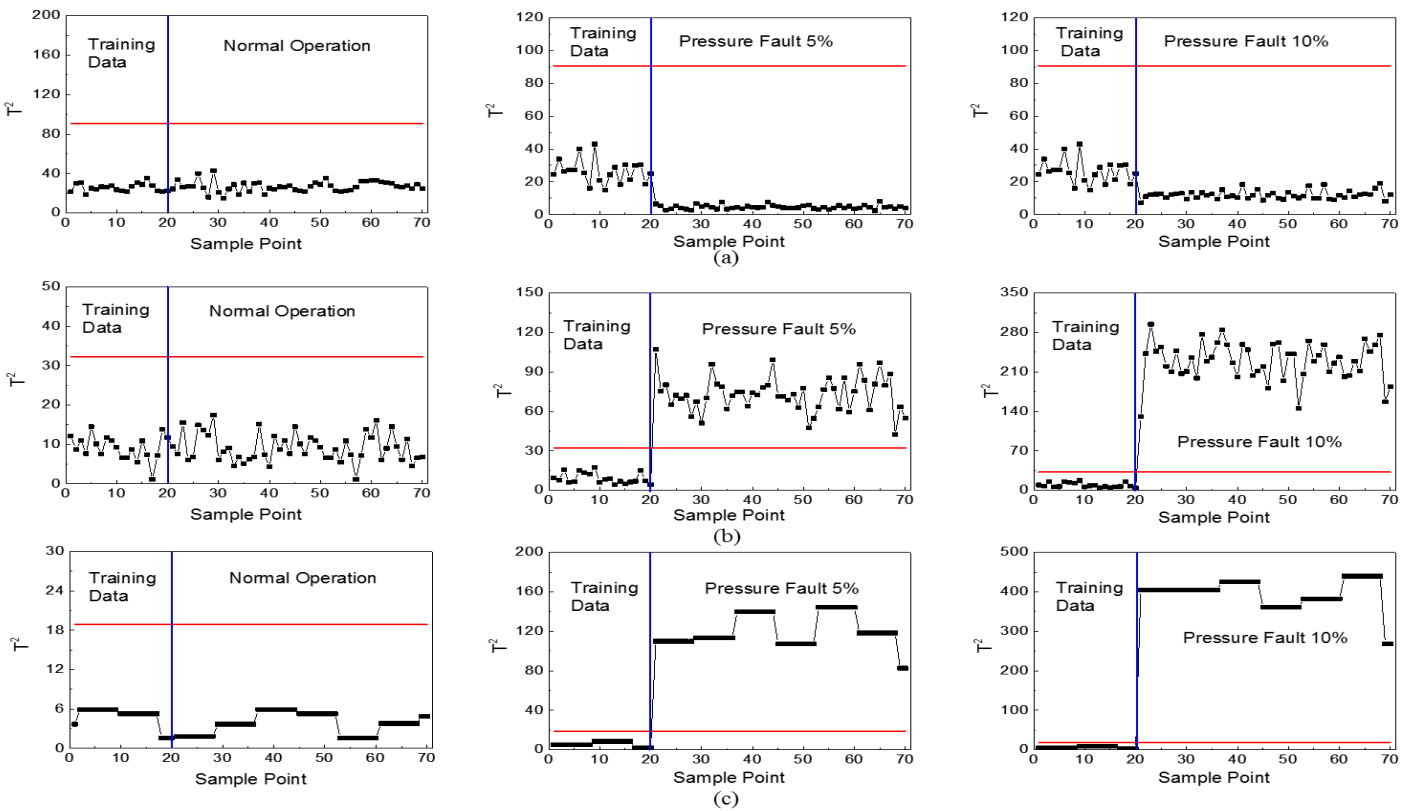

Figure 2-8. Hotelling's T^2 under pressure fault situations.

(a) Conventional PCA model, (b) P-PCA model, and (c) PD-PCA model

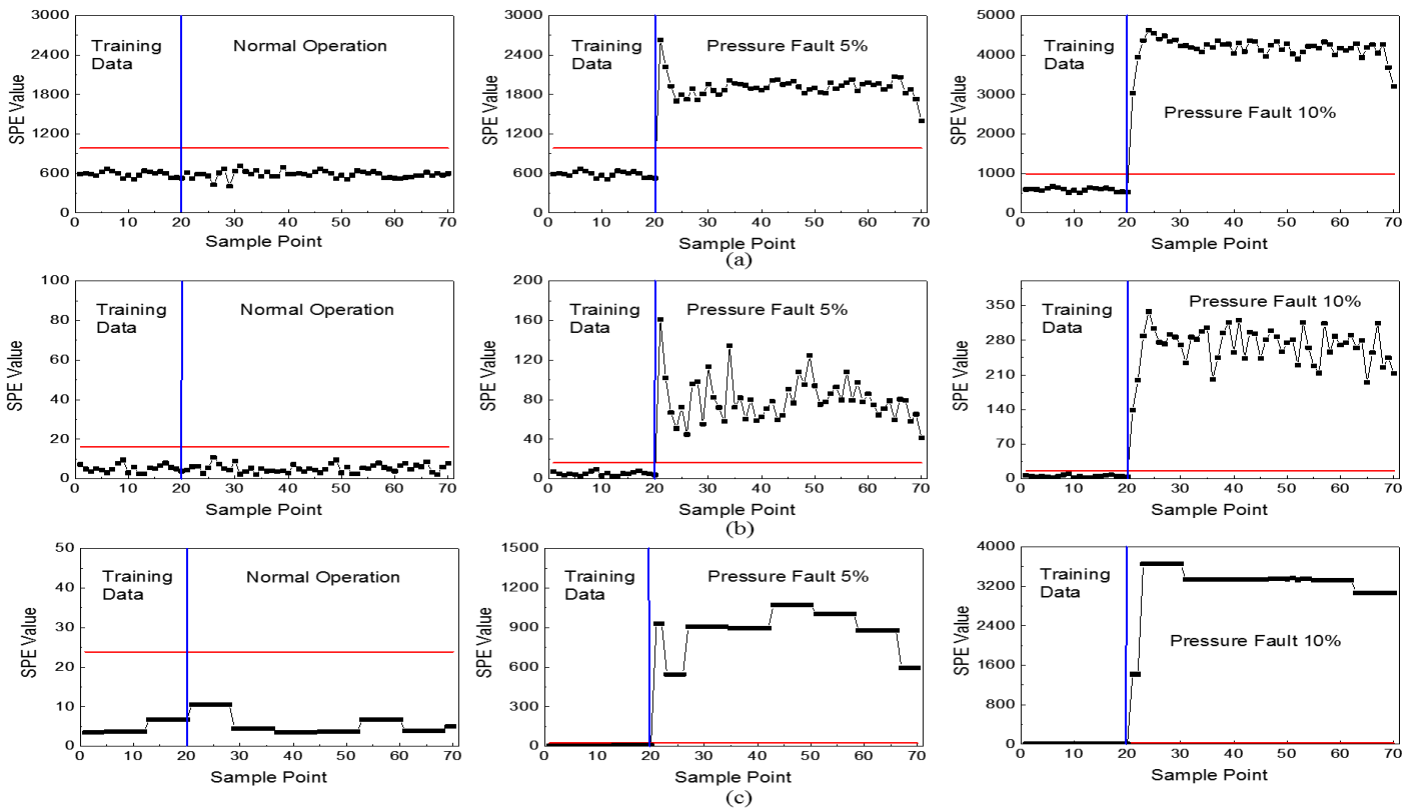

Figure 2-9. SPE under pressure fault situations. (a) Conventional PCA model, (b) P-PCA model, and (c) PD-PCA model

Table 2-2. Summary of missed detection rate in this study. (Hotelling's T^2)

	Conventional PCA	**P-PCA[1]**	**PD-PCA[2]**
Pressure Fault 5%	50/50 = 1.0	0/50 = 0.0	0/50 =0.0
Pressure Fault 10%	50/50 = 1.0	0/50 = 0.0	0/50 = 0.0
Source Power Fault 0.67%	50/50 = 1.0	31/50 = 0.62	0/50 = 0.0
Source Power Fault 1.33%	50/50 = 1.0	0/50 = 0.0	0/50 = 0.0
Source Power Fault 1.67%	50/50 = 1.0	0/50 = 0.0	0/50 = 0.0

[1]P-PCA: peak wavelength selection algorithm is applied.

[2]PD-PCA: peak wavelength selection algorithm and DWT are applied.

Table 2-3. Summary of missed detection rate in this study. (SPE)

	Conventional PCA	**P-PCA[1]**	**PD-PCA[2]**
Pressure Fault 5%	0/50 = 0.0	0/50 = 0.0	0/50 = 0.0
Pressure Fault 10%	0/50 = 0.0	0/50 = 0.0	0/50 = 0.0
Source Power Fault 0.67%	0/50 = 0.0	21/50 = 0.41	0/50 = 0.0
Source Power Fault 1.33%	0/50 = 0.0	1/50 = 0.02	0/50 = 0.0
Source Power Fault 1.67%	0/50 = 0.0	0/50 = 0.0	0/50 = 0.0

[1]P-PCA: peak wavelength selection algorithm is applied.

[2]PD-PCA: peak wavelength selection algorithm and DWT are applied.

In a similar way, Figure 2-10 shows Hotelling's T^2 for source power fault cases where 0.67%, 1.33% or 1.67% of source power with regard to the reference condition is increased. Similar to pressure fault cases, the conventional PCA model fails to detect any source power faults. The P-PCA model fails to detect 0.67% source power fault, as shown in Figure 2-10(b). On the contrary, the PD-PCA model surely detect all of the source power faults, as shown in Figure 2-8(c). As a result, the missed detection rate of the PD-PCA model in table 2-2 is 0.

SPE for the source power fault cases in Figure 2-11 shows a bit different trend from that for the pressure fault cases. Only the P-PCA model misses 21 faults in 0.67% of source power fault case and 1 fault in 1.33% of source power fault case. This might come from the fact that source power change (0.67%, 1.33%) is so slight that the noise from OES signal plays a role in the P-PCA case.

Consequently, PD-PCA model can detect faults from process condition variations more accurately and sensitively than P-PCA and conventional PCA model. The results so far propose that pre-selection of physically meaningful wavelength in OES spectrum and filtering of noise from a CCD array should be considered in modeling OES based plasma data by PCA.

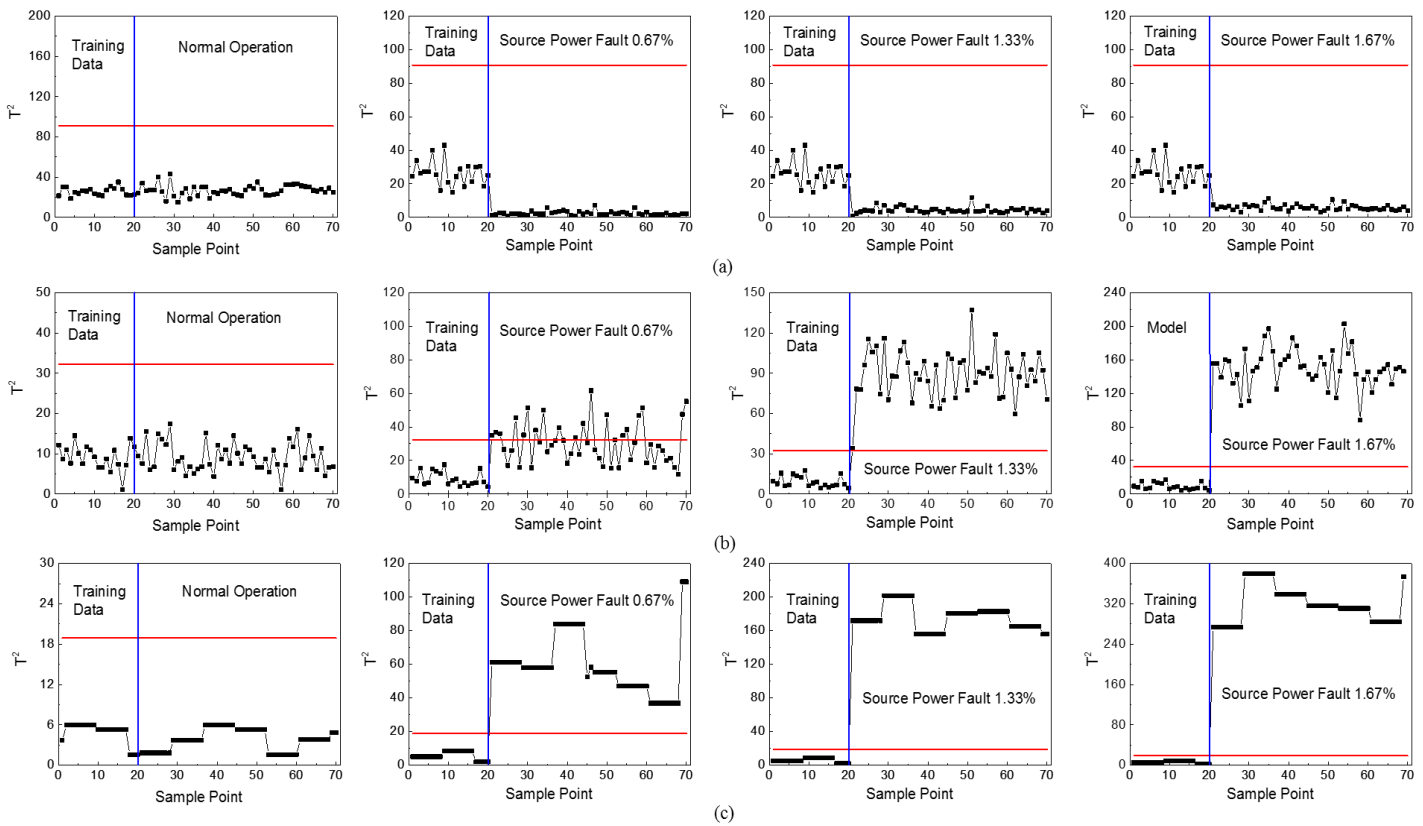

Figure 2-10. Hotelling's T^2 under source power fault situations. (a) Conventional PCA model, (b) P-PCA model, and (c) PD-PCA model

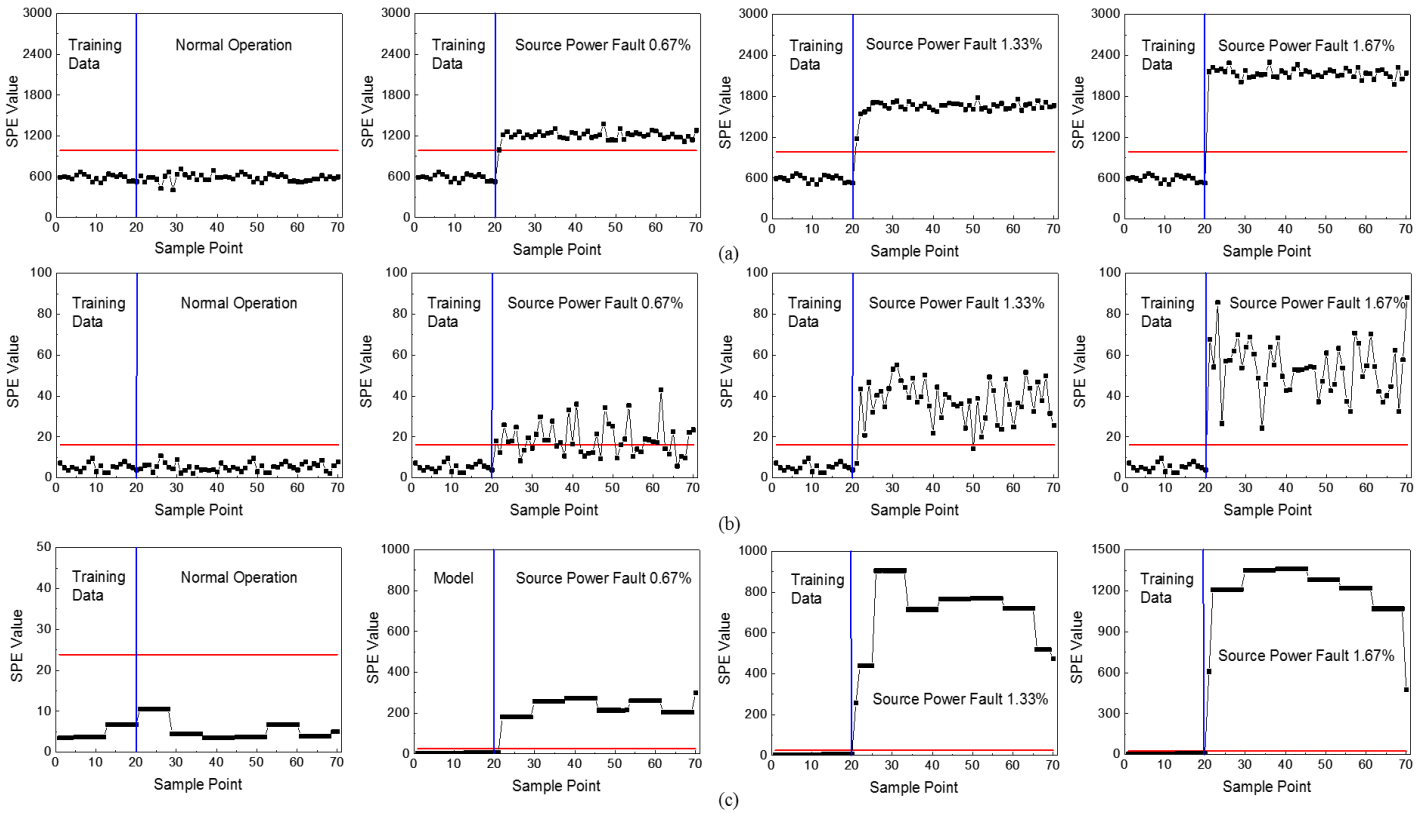

Figure 2-11. SPE under source power fault situations. (a) Conventional PCA model, (b) P-PCA model, and (c) PD-PCA model

2.6 Conclusion

Improvement of OES based PCA modeling is achieved by applying the peak wavelength selection algorithm and DWT in this paper. The peak wavelength selection algorithm effectively reduces high dimensional OES data into a group of physically meaningful wavelength variables. DWT successively decomposes OES signal into major signal and high frequency noise. Both techniques, in fact, enhance fault detection capabilities of the PCA model over the conventional PCA model under various source power and pressure faulty situations. The results introduced in this paper can be applied to other multivariate analysis (MVA) techniques, which is expected to contribute OES based plasma monitoring in plasma etching in such a way to enhance a cost-effective manufacturing approach for more than Moore's law era.

CHAPTER 3 : Multimode PCA and k-nearest neighbor algorithm for LNG mixed refrigeration process monitoring

3.1 Introduction

With an increase in the complexities in the industrial systems and high level process integrations, the need for more accurate process monitoring techniques to detect the instrumental malfunctions are required to improve product yield, reduce operational risk, ensure safe operation and achieve sustainable profit. In the industrial process, thousands of sensors, analyzers and control loops have strong interactions among each other in terms of process variables. In order to efficiently manage such complex processes, data-driven process monitoring methods have been utilized to handle large dimensional data from thousands of sensors. In this regard, MacGregor and Kourti (1995), Kano et al. (2001), Kourti (2002), Qin (2003), and Venkatasubramanian et al. (2003) have developed multivariate statistical tools for process control.

Process operating condition changes frequently due to set point changes, fluctuations in raw materials, composition of feed material, equipment aging and seasoning effects. In these situations, the application of traditional process monitoring methods based on the assumption that the process has only one stable operation region may cause false alarms, when the process is operated under another

steady-state nominal operating mode. This is because different modes of process usually have different statistical properties such as mean value, variance, and correlation between variables (Chiang et al. (2000), Lee et al. (2004), and Yin et al. (2014)).

To cope with these situations, multi-mode process monitoring methods were developed, which can be categorized into (a) global modeling, (b) multiple modeling, and (c) adaptive modeling approaches. Global modeling is an approach to developing the uniform model which can be applied to all operation modes. Hwang and Han (1999) proposed hierarchical clustering and super-principal component analysis (PCA) model. Lane et al. (2001) built a group model capable of generating the cluster of processes in terms of grades. Multiple modeling is a method to build local models which can match each operation mode. Zhao et al. (2004) and Yoo et al. (2007) proposed a multiple PCA model based process monitoring methodology. Adaptive modeling is a sort of a model update scheme for mode changes. Jin et al. (2006) and Choi et al. (2006) proposed a robust recursive PCA monitoring methodology for a process that includes frequent operation mode changes.

There are certain limitations of those approaches. For instance, global model which is based on multi-mode monitoring frequently shows low detection resolutions for a particular mode because it employs statistical mean of the whole data (Tan et al. (2010) and Ge et al. (2013)). In addition, a priori process knowledge is required in the preliminary step to manually segment the historical operating data according to different operating modes. Moreover, a similarity threshold has to be predefined by user to incorporate the similar data groups. Those conditions are not desirable for automatic process monitoring in industrial practice. Multiple modeling has a

detection capability with high resolution but the false alarm rate also increases in monitoring between-mode transition situations because it ignores cross-mode correlation (Zhang et al. (2013) and Haghani et al. (2014)). Adaptive modeling shows low reliability of monitoring results when process disturbance is included in model update (Qin et al. (1998), Li et al. (2000), and Wang et al. (2005)). Thus, it is still required to develop monitoring techniques of multi-mode operation cases.

This paper utilizes the k-nearest neighbor (k-NN) classification and multiple PCA models to enhance process monitoring performance under frequent operation mode change environments. The proposed methodology is evaluated with collected data from a real liquefied natural gas (LNG) mixed refrigeration (MR) process. The rest of paper is organized as follows. In Section 2, a LNG MR process is briefly described and the data collection procedure is explained. After describing the target process and data collection, theoretical background of PCA and *k*-NN is described in Section 3. In Section 4, mode identification techniques and fault detection models proposed in this study are explained as well as how they are integrated. In Section 3.5, performances of the proposed techniques are compared with those of general PCA by using real LNG MR process data.

3.2 Target process and data description

Figure 3-1 shows a flow of a general LNG process, which consists of pre-processing, refrigeration, liquefaction, and storage sections. In the pre-processing stage, impurities including CO_2 and H_2S in natural gas are removed in acid gas removal unit, whereas the moisture is removed in the dehydration unit. In the refrigeration and distillation stage, mixed refrigerant (MR) which comprises of

nitrogen, methane, ethane, propane, and sometimes butane is also used to cool down natural gas in a single main cryogenic heat exchanger (MCHE). In the liquefaction stage, liquefied natural gas reduces its volume by a factor of more than 600, making it more economical to transport over long distances by transport ships or pipelines.

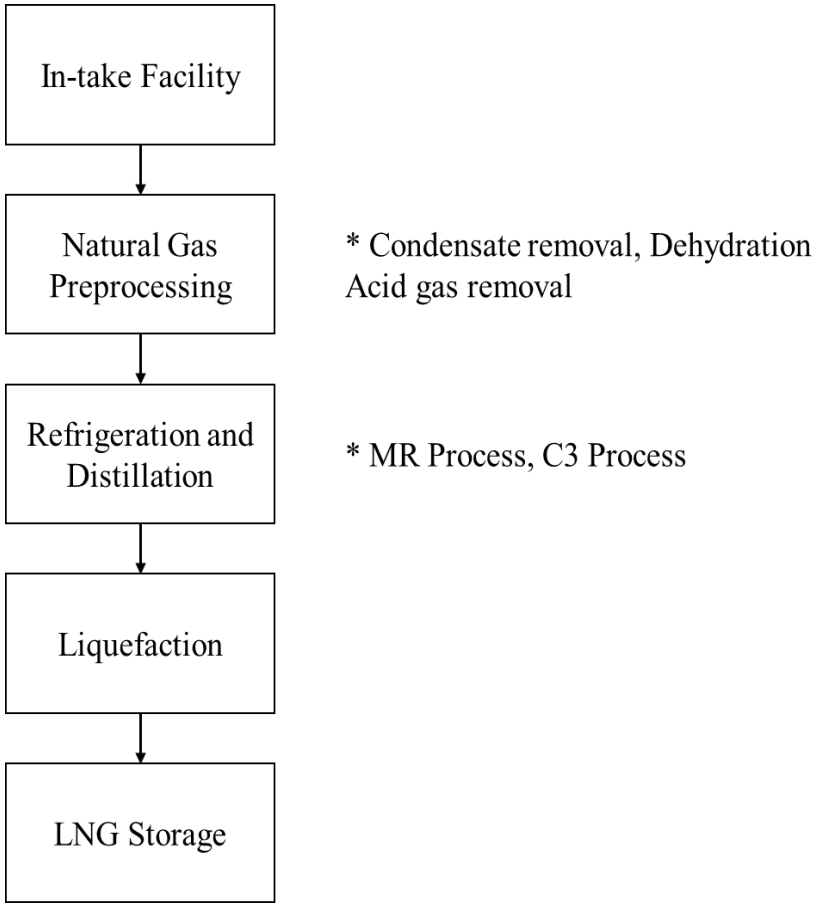

Figure 3-1. Process flow of a typical LNG process

This study focuses the MR compression process for process monitoring. The simplified MR compression process used in a LNG liquefaction plant is shown in Figure 3-2. The storage suction drums are placed before the 1st and 5th stage of the compressor. They separate the liquid refrigerant from gas and liquid mixture so as to prevent the liquid phase of refrigerant from entering the compressors. In the MR compression, multiple stages of gas compression are required to achieve the desired pressure (27 – 34 bar) to make compressed mixed refrigerant. The MR compressor operates at 3.8 bar and continually compresses until the pressure reaches to 20 bar at fourth stage. The discharge flow from the 4th stage compressor is then cooled to 32℃ by the 1st cooler and passes to the 5th stage compressor via the 5th stage suction drum. The discharge flow is then compressed to the designed pressure level in the 5th and 6th stage compressors. The discharge from this compressor is cooled again by the 2nd cooler and is partially condensed in propane and mixed refrigerant evaporators. The mixed refrigerant liquid is sub-cooled while the mixed refrigerant vapor is condensed in MCHE. The partially condensed mixed refrigerant liquid is routed to a MR separator through the warm Joule-Thompson valve while the condensed mixed refrigerant vapor is sent to MCHE.

Major process variables described in Figure 3-2 are summarized in Table 3-1. Total 17 variables are collected including compressor's pressure, temperature and flowrate variables in this study. Each pressure, temperature, and flow gauge is located at inlet or outlet of compressors. Among variables from each gauge, pressure

variables can represent compressors' efficiency and flowrate variables can monitor refrigerant flow. If the operating conditions of compressors fluctuate or some compressors have a malfunction, pressure and flowrate are directly affected.

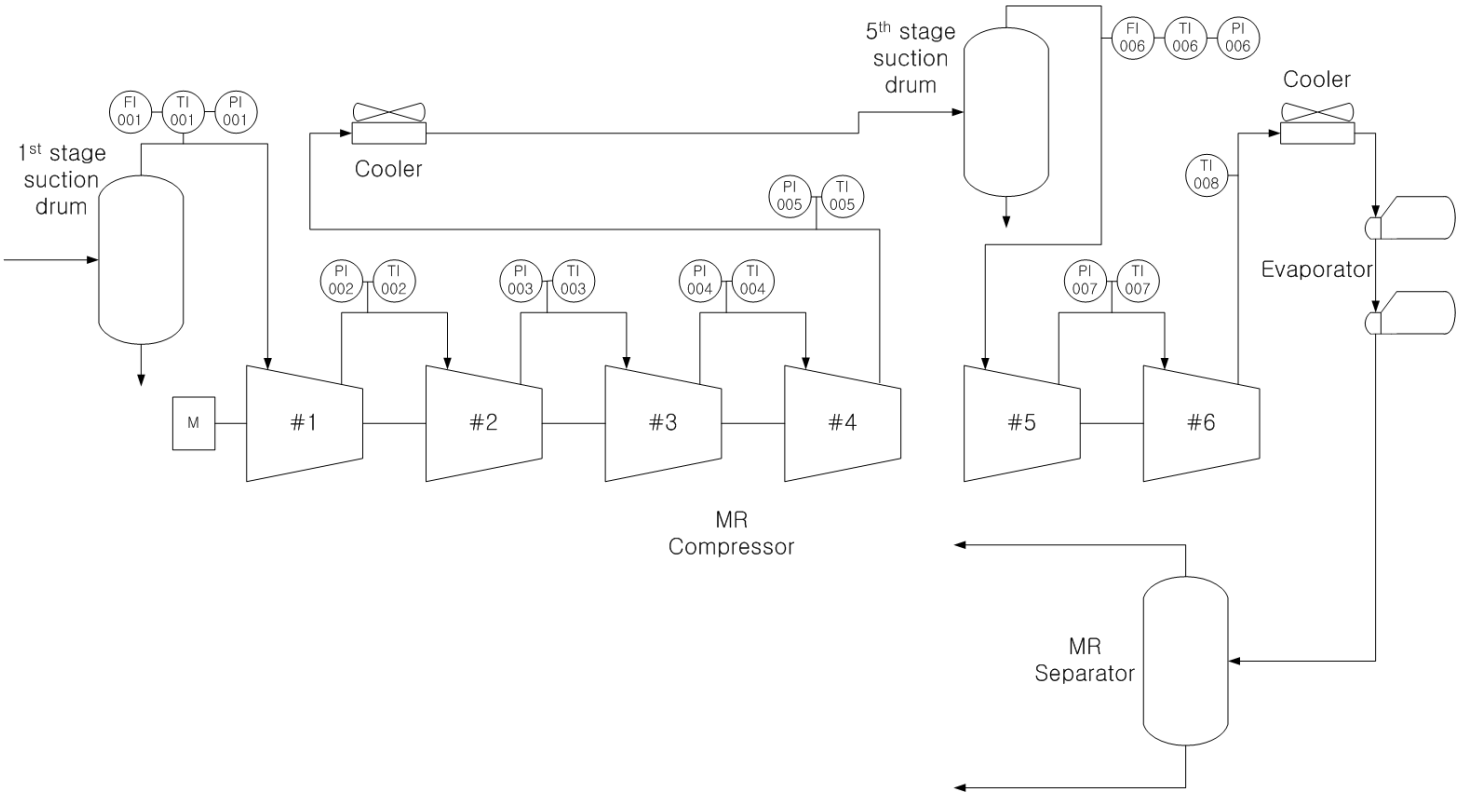

Figure 3-2. Schematic diagram of a LNG MR process

Table 3-1. Variable description in the LNG MR Process

Module	Inlet Variable	Outlet Variable	Description
Compressor #1	PI-001	PI-002	Pressure
	TI-001	TI-002	Temperature
	FI-001		Flowrate
Compressor #2	PI-002	PI-003	Pressure
	TI-002	TI-003	Temperature
Compressor #3	PI-003	PI-004	Pressure
	TI-003	TI-004	Temperature
Compressor #4	PI-004	PI-005	Pressure
	TI-004	TI-005	Temperature
Compressor #5	PI-006	PI-007	Pressure
	TI-006	TI-007	Temperature
	FI-006		Flowrate
Compressor #6	PI-007		Pressure
	TI-007	TI-008	Temperature

3.3 Theoretical Background

3.3.1 Principal component analysis based fault detection

3.3.1.1 Principal component analysis

Principal component analysis (PCA) is a linear data reduction method by capturing the optimal variability in multivariate dataset. PCA determines a set of loading vector matrix V and diagonal matrix Λ, ordered by the amount of variance explained in the loading vector directions and solving an eigenvalue decomposition of the sample covariance matrix S.

$$S = \frac{1}{n-1} X^T X = V \Lambda V^T \quad (1)$$

, where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_l\}$.

The training data matrix $X(n \times m)$, which has n observations of m measurement variables, can be decomposed to produce loading vectors corresponding to the largest singular values in order to optimally capture the variations. A PCA model is usually built from a few principal component. These principal components are the results of decomposing a data matrix X using PCA as follows:

$$X = TP^T + E = \sum_{i=1}^{k} t_i p_i^T + E \quad (2)$$

, where p_i is loading vector and t_i is score vector in the PCA model.

3.3.1.2 Fault detection index

On-line monitoring of measurement variables can be carried out with the help of the Hotelling's T^2 and Q statistics, this last, also known as the squared prediction error (SPE). These two statistics can be used to detect faults for multivariate process data.

The T^2 statistic for the lower-dimensional space can be calculated for each new observation x by:

$$T^2 = x^T P(\Sigma_a)^{-2} P^T x \quad (3)$$

, where $\Sigma_a$represents the non-negative real eigenvalues corresponding to the P principal components.

The upper confidence limit for T^2 is obtained using the *F*-distribution:

$$T^2_{a,n,\alpha} = \frac{a(n-1)}{n-a} F_{a,n-a,\alpha} \quad (4)$$

, where *n* is the number of samples in the data set, *a* represents the number of principal components and α shows the level of significance.

T^2 statistic can be interpreted as the measure of systematic variations of the process, and any violation of the threshold would indicate that the systematic variations are out of control. The portion of the measurement space corresponding to the lowest $m - a$ eigenvalues can be monitored using the squared prediction error (SPE) or Q statistic.

$$Q = x^T(I - PP^T)x \quad (5)$$

, where I is the identity matrix.

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

$$Q_a = \theta_1 \left(\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{\frac{1}{h_0}} \quad (6)$$

with:

$$\theta_i = \sum_{j=a+1}^{m} \lambda_j^i, \quad (7)$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \quad (8)$$

, where c_α is the value of the normal distribution with α representing the level of significance.

The Q statistic does not suffer from an over-sensitivity to inaccuracies in the smaller singular values and it is associated with noise measurements. A violation of the threshold would indicate that the random noise has significantly changed, or un-usual event has occurred, which had produced a change in the covariance structure of the model.

3.3.2 *k*-Nearest Neighbor classifier

k-nearest neighbor (*k*-NN) algorithm is a method to classify a new dataset by calculating its distance to the nearest neighbor by training samples in the feature space. For a given unlabeled sample *x*, the *k*-NN rule finds the *k*-nearest labeled samples in the training data set based on some distance metrics. The point is assigned to the class with most votes for class label amongst the *k*-neighbor points. The classifier is defined by its parameters. Setting parameter *k* depends on the data and affects the performance of the classifier. Parameter '*k*' must be large enough to reduce misclassification of an example point and must be small enough so that the sample point is close to the neighboring points, which results in better estimation of the point's class. For pattern classification, the *k*-NN algorithm only requires a set of labeled samples *k*, and a metric to measure distance. There are in general two steps in the *k*-NN algorithm:

1) Calculate the *k*-NN for each sample in the training dataset.

2) Calculate the *k*-NN squared distance for each sample, where *k*-NN squared distance of sample i $\left(D_i^2\right)$ is defined as the sum of squared distances of sample *i* to its *k*-NN.

$$D_i^2 = \sum_{j=1}^{k} d_{i,j}^2 \qquad (10)$$

, where $d_{i,j}^2$ denotes squared Euclidean distance from sample i to its j^{th} nearest neighbor.

3.4 Mode identification and fault detection

3.4.1 Operation mode identification and fault detection

In general, chemical processes are operated under several operation modes and frequently change due to the alteration of feed flow rate or composition, product specifications and utility supply conditions. Under these circumstances, multivariate statistical process monitoring methods should retain their monitoring capabilities against performance degradation from operating mode changes or disturbances. In order to do that, it is necessary to firstly determine which operation mode an incoming process belongs to. For this purpose, the *k*-NN classifier described in Section 3.2 is employed in this study.

Figure 3-3 illustrates a proposed sequence of operation mode classification using the *k*-NN classifier. In a global PCA modeling stage, collected historical data is firstly pre-processed by filtering out trip or transition data which is defined as in-between data of operation mode changes. With this pre-processed data, PCA is then done to build a global PCA model and to reduce data space so that it can alleviate computational resource burden in calculating *k*-NN classifier. Of course, outliers which surpass pre-determined Hotelling's T^2 and SPE specification are excluded before *k*-NN classification.

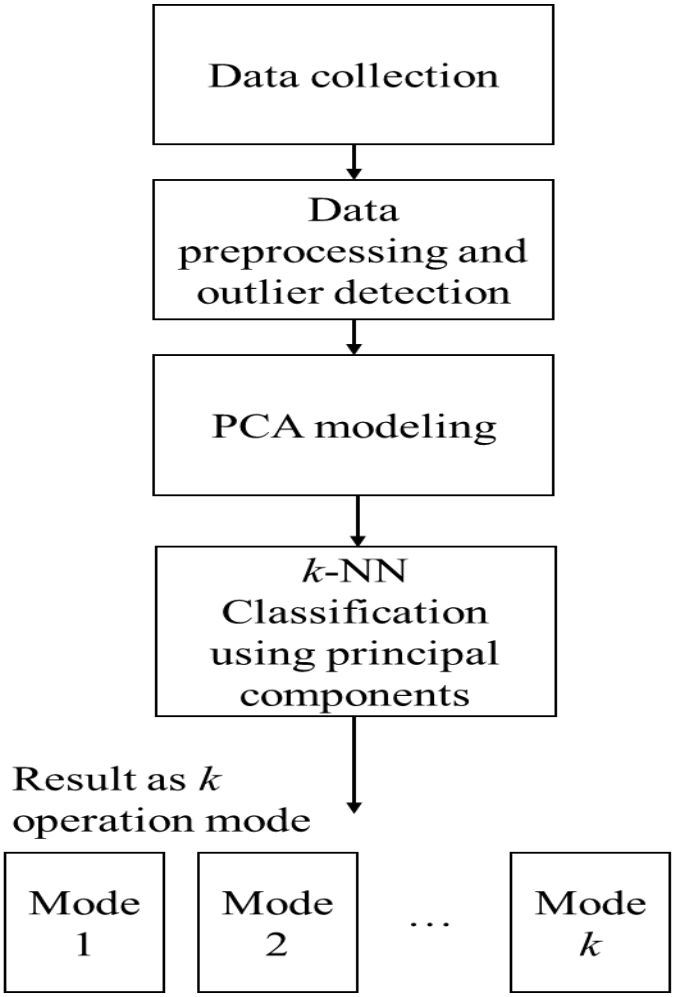

Figure 3-3. Flow chart of operation mode classification using *k*-NN classifier

Figure 3-4 shows score values in the 1st and the 2nd principal component spaces for normal operation data which is collected in a real LNG MR process. Data set is grouped into 6 modes by the *k*-NN classifier, which is shown in Figure 3-5. For each mode, sub-PCA model is built with new average and new standard deviation of each group and they are stored in a fault detection library together with the global PCA model. Figure 3-6 illustrates a flow chart of a fault detection procedure proposed in this paper. When it comes in, new data set is normalized with the average and the standard deviation of the global PCA model and it is projected into the global PCA model space. The score value is then matched with the nearest neighbor mode by the *k*-NN classifier algorithm. A sub-PCA model corresponding to the matched mode is applied to the incoming data to decide whether it is in normal operation or in a fault situation with its Hotelling's T^2 and SPE specifications.

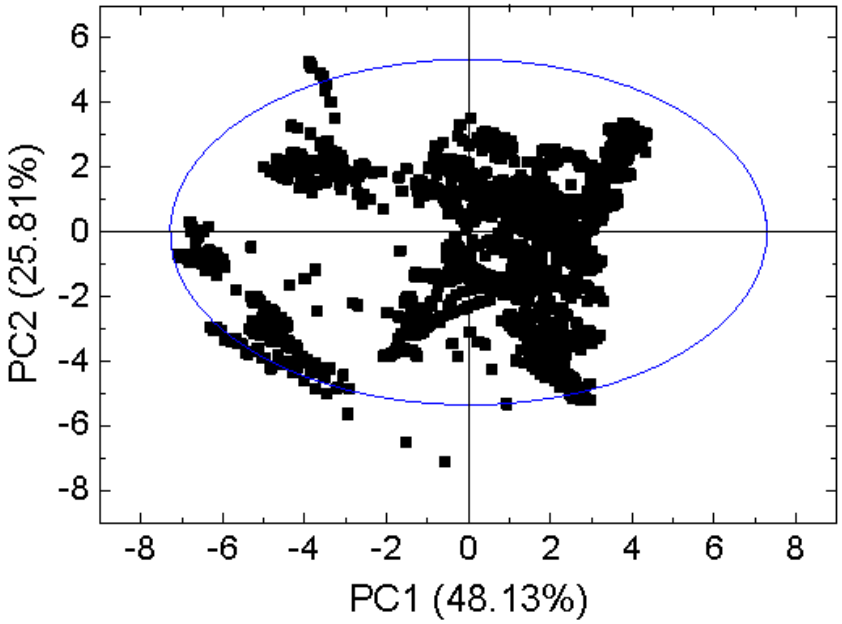

Figure 3-4. PCA score plot for normal operation data

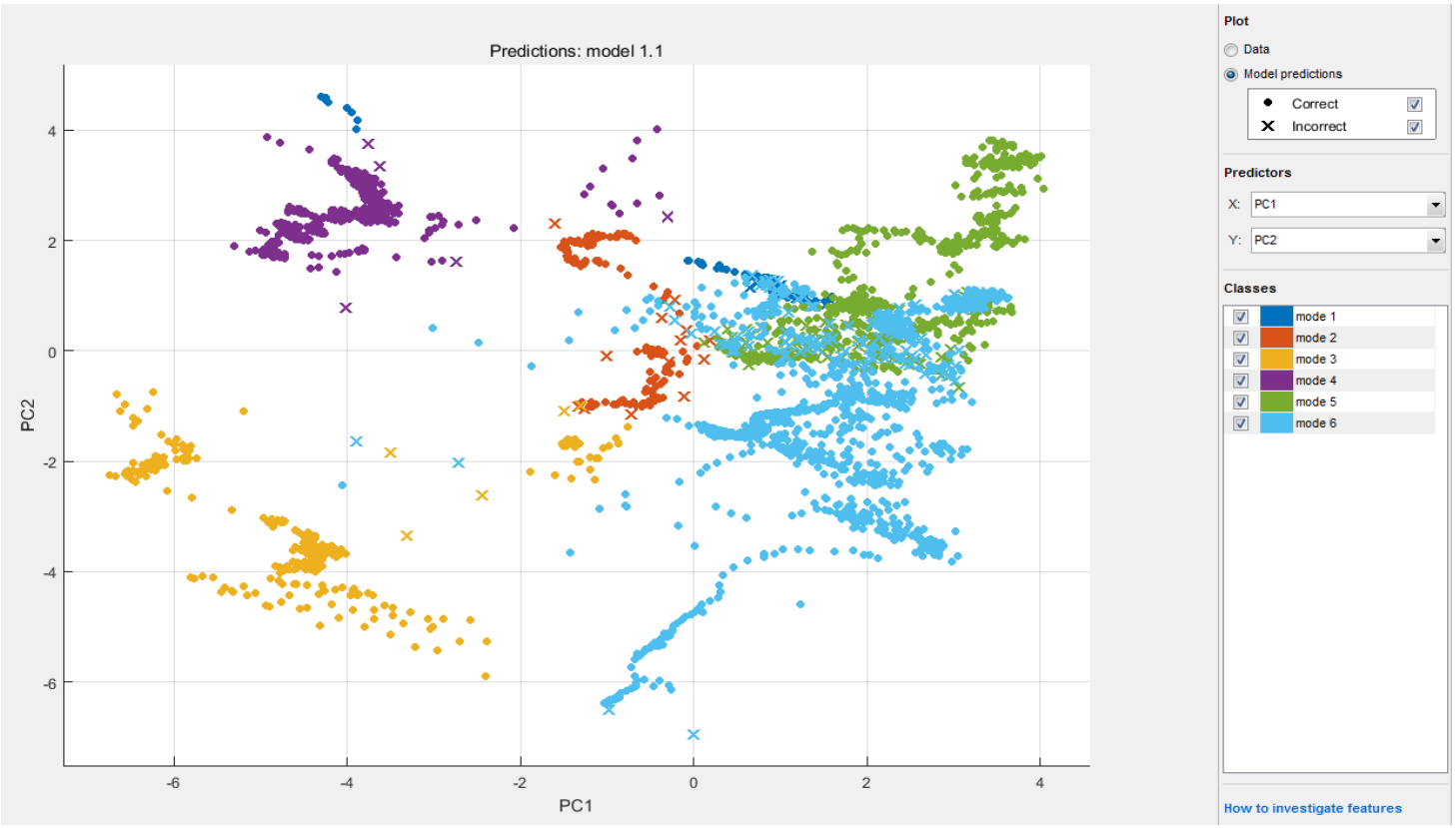

Figure 3-5. Results of operation mode classification using *k*-NN classifier algorithm

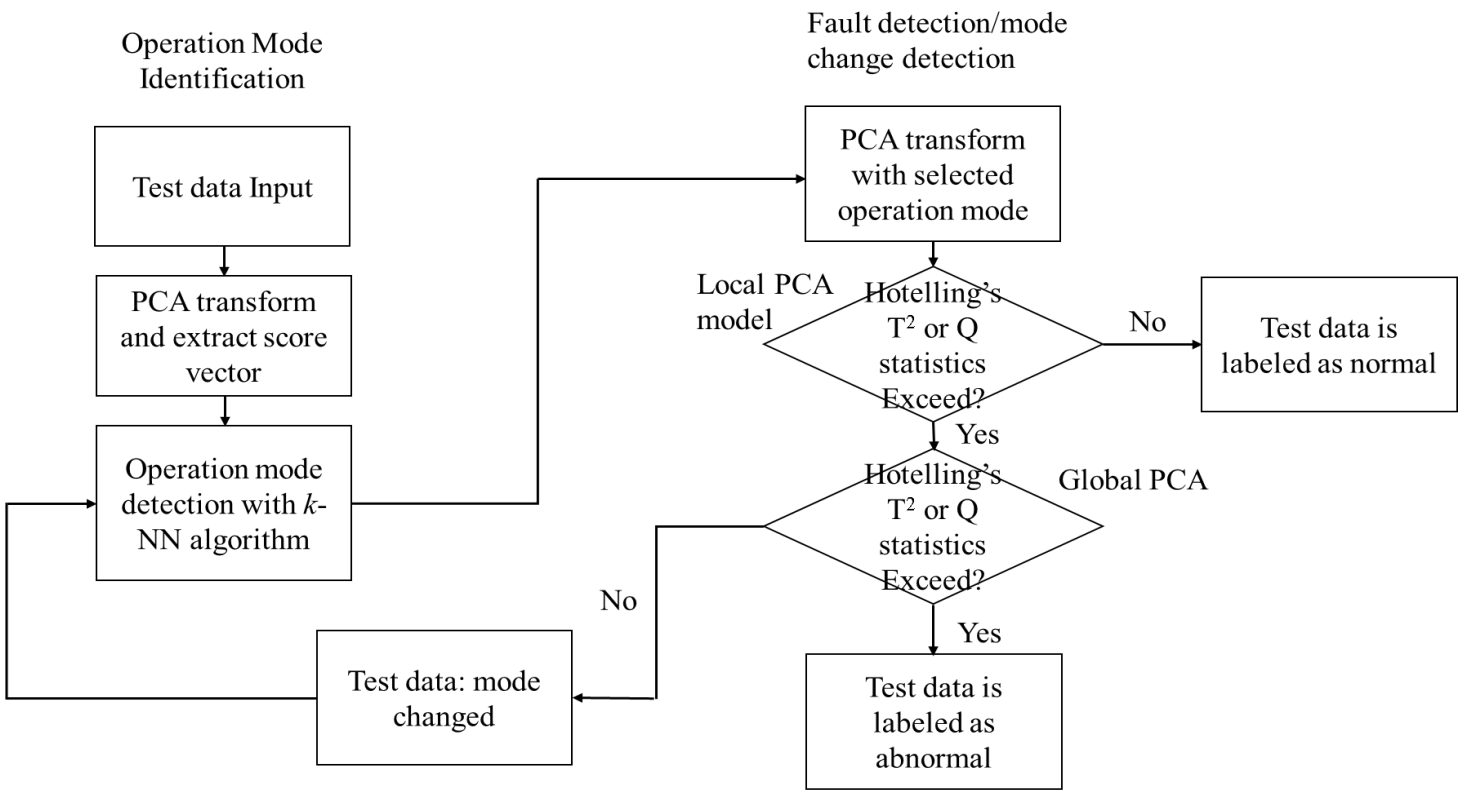

Figure 3-6. Flow chart of proposed mode identification and fault detection using *k*-NN and multimode PCA algorithm

3.5 Results and Conclusion

3.5.1 Consideration in LNG MR process monitoring

Compression processes manufacturing mixed refrigerant in LNG processes are generally operated under high pressure and high speed environment and thus accurate and speedy monitoring capabilities are inevitable for process safety. Major variables to be monitored in compressor operation are pressure, flowrate, temperature, pressure ratio and so on. Pressure or pressure ratio is a variable to monitor load to each compressor in series. Flowrate is a variable to monitor continuous flow in and out of compressor, which is a necessary parameter to prevent mechanical damage from sudden fluctuations. Temperature is a variable to check heat balance across whole compressors.

The main sources of possible faults in compressors are vibration from irregular and/or fluctuated gas flow. Increasing vibration might indicate blade, bearing, and/or shaft have problems, which can lead to catastrophic compressor failure and then potential unit shutdown. Fluctuations in gas flow induce instability of compressor's operation, which can lead to mechanical damage and/or compressor performance drift. These types of faults occur during frequent changes in the process state that occurs throughout the MR process. Figure 3-7 shows examples of pressure and flowrate trend in a MR process, which illustrates how frequently the processes are changed during 5 day trial operation; there are 6 different normal processes in blue shade, 3 trip periods in red shade and 3 transition periods before the normal processes. The set point for each normal process is summarized in Table 3-2. When there are various combinations of the set points of the variables like that shown in Figure 3-7

and Table 3-2, it is desirable to distinguish the difference of dynamic characteristics in advance. In addition, multivariate statistical methodologies, which can monitor whole process state changes by using measurable sensor variables are thought to be appropriate approaches to LNG MR process.

Under these circumstances, this paper proposes integration of *k*-NN algorithm into multi-mode PCA and evaluates its dynamic fault detection characteristics in LNG MR process. The results are shown step-by-step in following sections.

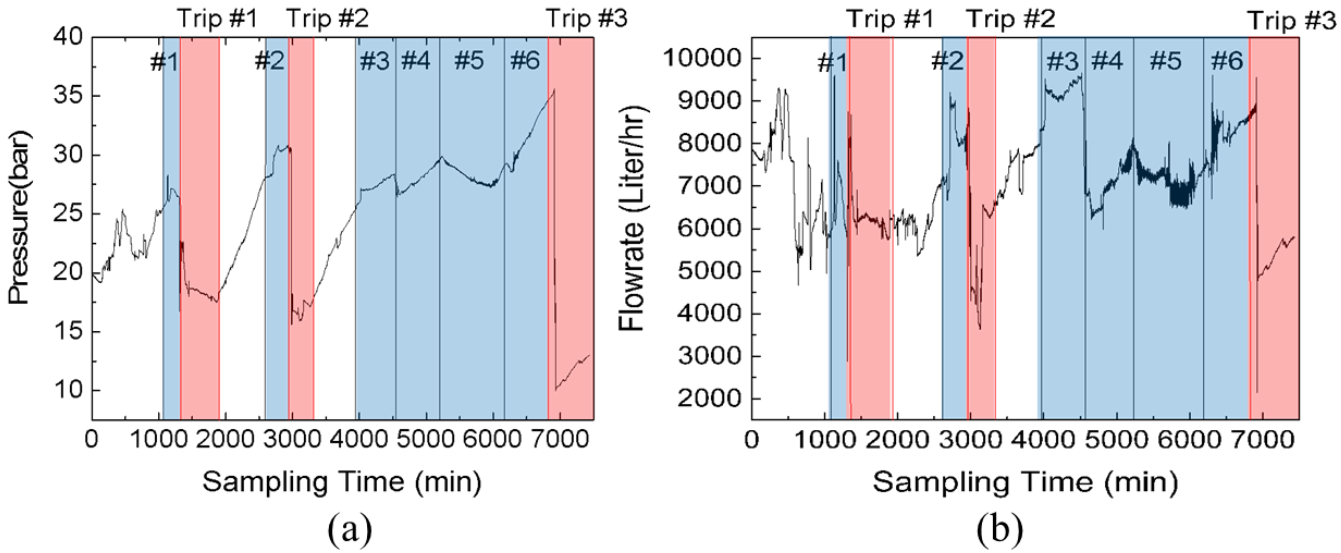

Figure 3-7. . Examples of pressure and flowrate trend in a LNG MR process: (a) PI-007 and (b) FI-006 in Table 3-1.

Table 3-2. Operation mode description (set point[1]).

Operation modes	Pressure (bar)	Flowrate (liter/hr)	Temperature (℃)
Mode #1	27.2, steady	7500, steady	73, steady
Mode #2	27 - 28.5	6800 - 7300	71.5, steady
Mode #3	28 - 30.5	8000, steady	72 - 76
Mode #4	27 - 28.5	9300, steady	69.5, steady
Mode #5	27 - 30	7500, steady	69.5, steady
Mode #6	30 - 35	7500, steady	69 - 73

[1] set point at the 6th compressor

3.5.2 **Global and local PCA modeling**

For the off-line analysis of the monitoring performance, data of 17 variables summarized in Table 1 are collected from a real LNG MR process plant during 5 day trial operation. This data contains several normal operation and trip history which is a period of sudden drops of sensor variables caused by malfunctioning or process changes. Data in Figure 3-7 is utilized for a global and each local PCA model. The global PCA model is built with all data from 6 normal operation periods in blue shade, but each local PCA model is built separately with each data in blue shade. Table 3-3 summarizes variance explained for the global and local PCA models. Every model employs only the first and second principal components and captures more than 70% ~ 80%

Table 3-3. Variance explained for the global and local PCA models.

	PC #1 (%)	PC #2 (%)	Cumulative Sum (%)
Global Model	48.12	25.81	73.94
Local (Mode 1) Model	71.18	18.49	89.67
Local (Mode 2) Model	67.83	15.25	83.08
Local (Mode 3) Model	54.86	30.38	85.24
Local (Mode 4) Model	53.08	21.26	74.34
Local (Mode 5) Model	41.30	35.28	76.58
Local (Mode 6) Model	45.79	35.14	80.93

3.5.3 **Detection of operation mode**

In this section, it will be verified how the proposed methodology can detect mode changes in the process operation. For the verification, two normal process periods are randomly selected and their data for 60 min is extracted to make two different operation modes. Those two operation modes are connected by transition period which is generated by data interpolation between them. Based on the procedure in Figure 3-3, the operation mode of the first and second cases is determined. For the real-time monitoring, moving window of 20 min with every 5 min upgrade scheme is applied for PCA modeling and *k*-NN classification. Figure 3-8 shows an example of operation mode detection result by using *k*-NN classification under moving window circumstance. It is detected that the first test case (Figure 3-8(a)) starts with operation mode 3 and transfers to operation mode 4. The matching ratio by *k*-NN classification is 100% of mode 3 or mode 4 for each mode period. The matching ratio for mode 3 is decreasing but that for mode 4 is increasing during the transition period. The similar result is obtained for another example (Figure 3-8(b)).

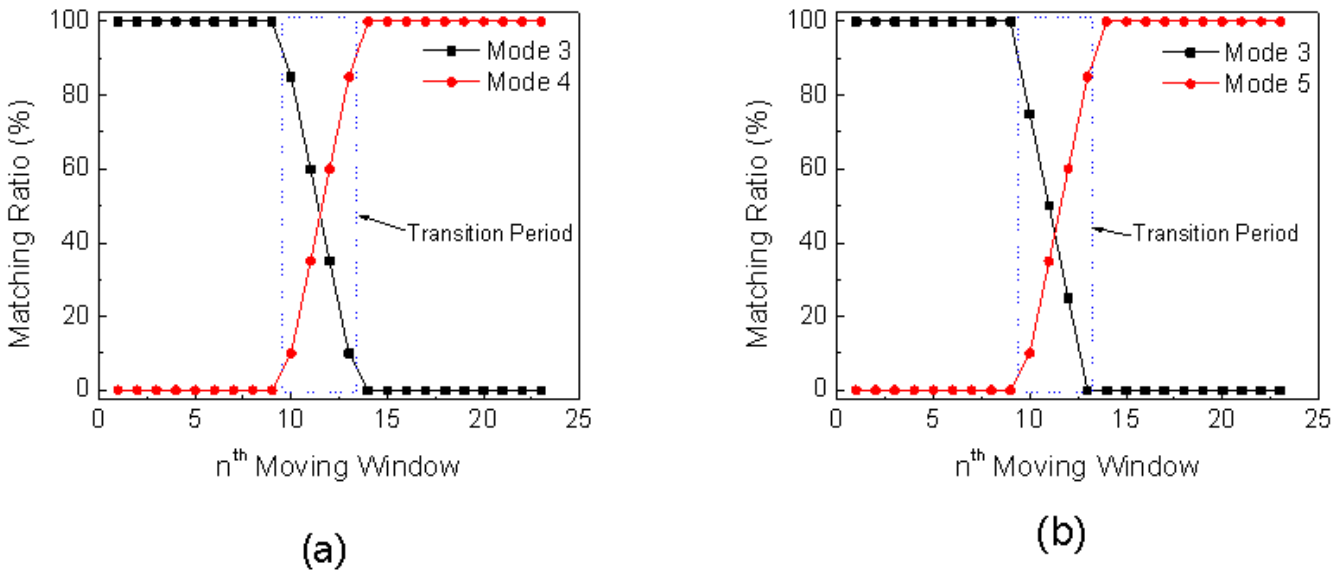

Figure 3-8. Results of mode detection according to n^{th} moving window: (a) mode 3 to 4 change and (b) mode 3 to 5 change.

Operation mode change might increase false alarm rate unless it is recognized in advanced. Figures 3-9 and 3-10 show Hotelling's T^2 and Q statistics of the global, local and *k*-NN + multi-mode PCA models for the first and second test cases. Here, the global PCA model is described in Section 5.2. The local PCA model is the model only with data in mode 3. The procedure in Figure 3-3 is followed with 6 local PCA models in Table 3-3 to build the *k*-NN + multi-mode PCA model.

In the global PCA model case, there is no false alarming like that shown in Figures 3-9(a) and 3-10(a). Of course, it does not distinguish the operation modes since the data including all operation mode is employed in modeling.

In the local mode PCA model case (Figures 3-9(b) and 3-10(b)), there is false alarming from transition to operation mode 4 period, which is because the model is built with operation mode 3 data set and it does not consider operation mode changes. The local PCA model is subject to false alarming under other operation modes.

On the contrary, the proposed *k*-NN + multi-mode PCA model detects mode changes and monitors the processes with the changed model. (Figure 3-9(c) and 3-10(c)). Thus, it will be possible to successfully distinguish real faults in normal operation even under operation mode change environment.

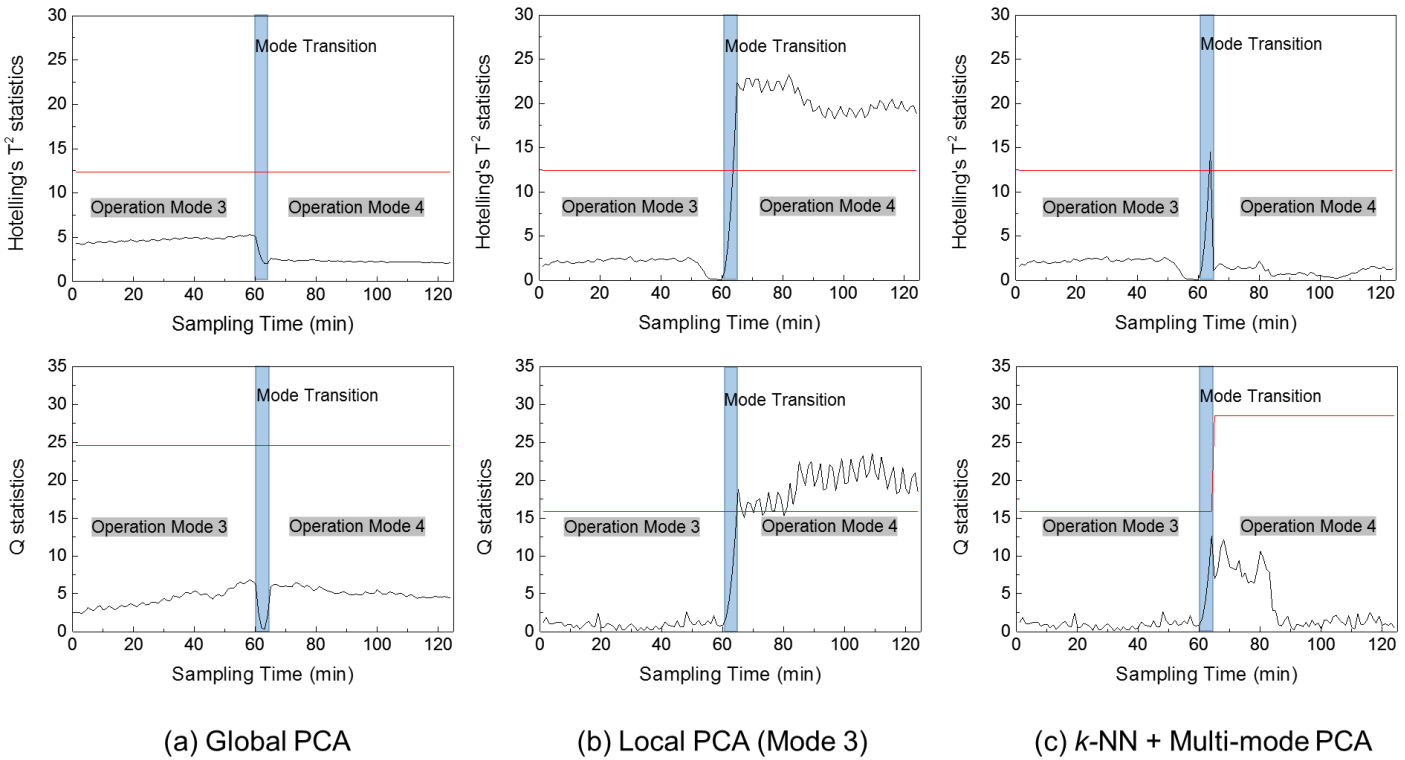

Figure 3-9. Monitored Hotelling's T^2 and Q statistics under mode change from 3 to 4: (a) global PCA case, (b) local PCA (mode 3) case, and (c) *k*-NN + multi-mode PCA case

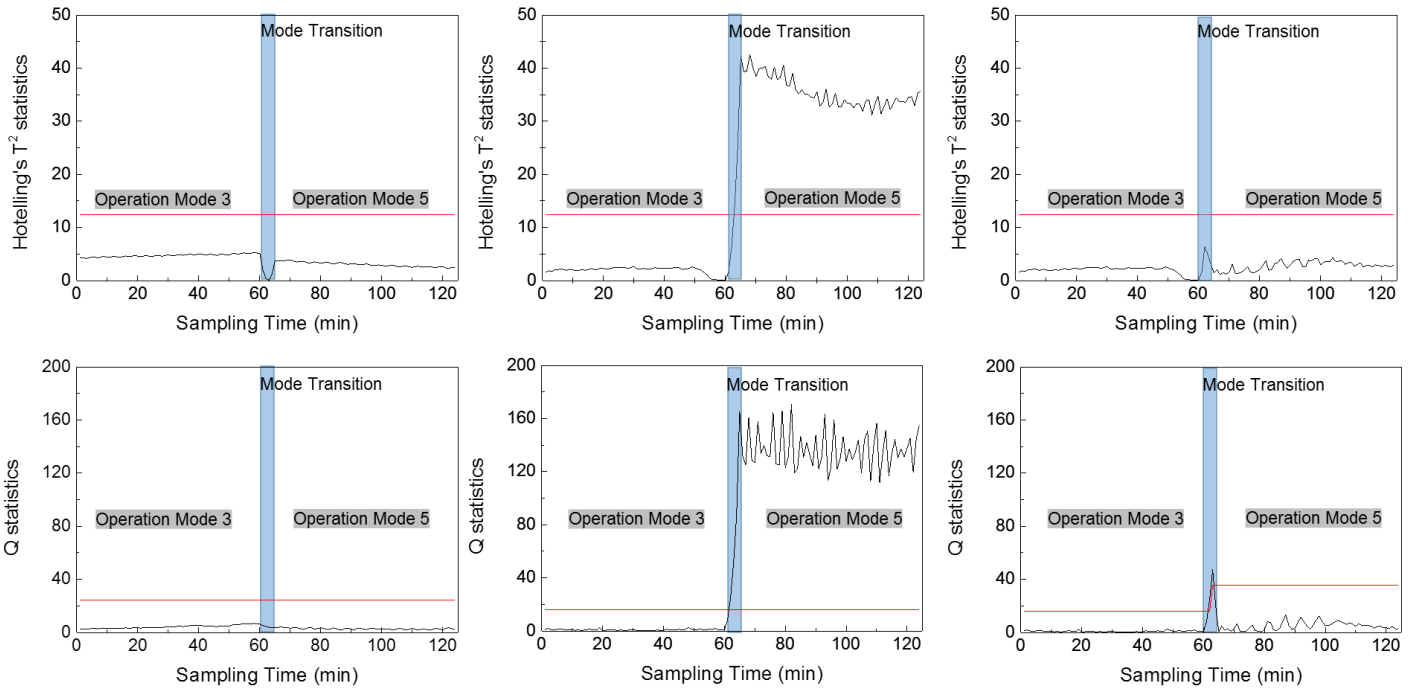

Figure 3-10. Monitored Hotelling's T^2 and Q statistics under mode change from 3 to 5: (a) global PCA case, (b) local PCA (mode 3) case, and (c) k-NN + multi-mode PCA case

3.5.4 **Comparison of fault detection performance**

Two critical fault cases from a real LNG MR process plant are selected to demonstrate how compressor shutdown can be detected accurately and in advance. Both cases are compressor vibration caused by unstable flow, which eventually leads to process shutdown. In order to compare fault detection performance under operation mode change, the proposed *k*-NN integrated multi-mode PCA and conventional PCA methodologies are applied to the cases.

The historical data which contains from the stable operation region of the MR compression process to the emergency shutdown are retrieved. The Hotelling's T^2 statistics and Q statistics for each PCA model are shown in Figures 3-11 and 3-12.

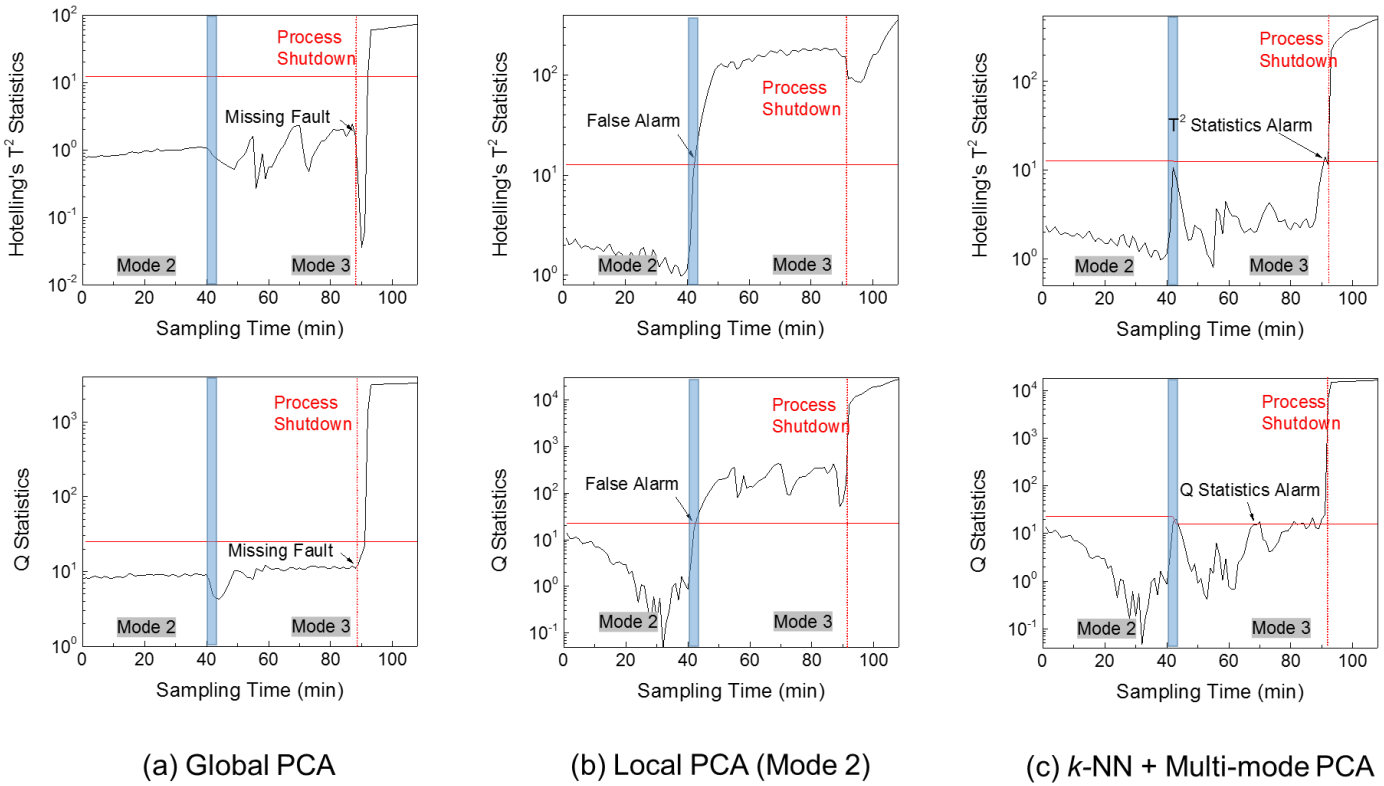

Figure 3-11. . Monitoring performance for the fault case in mode 3: (a) global PCA, (b) local PCA (for mode 2), and (c) *k*-NN + Multi-mode PCA

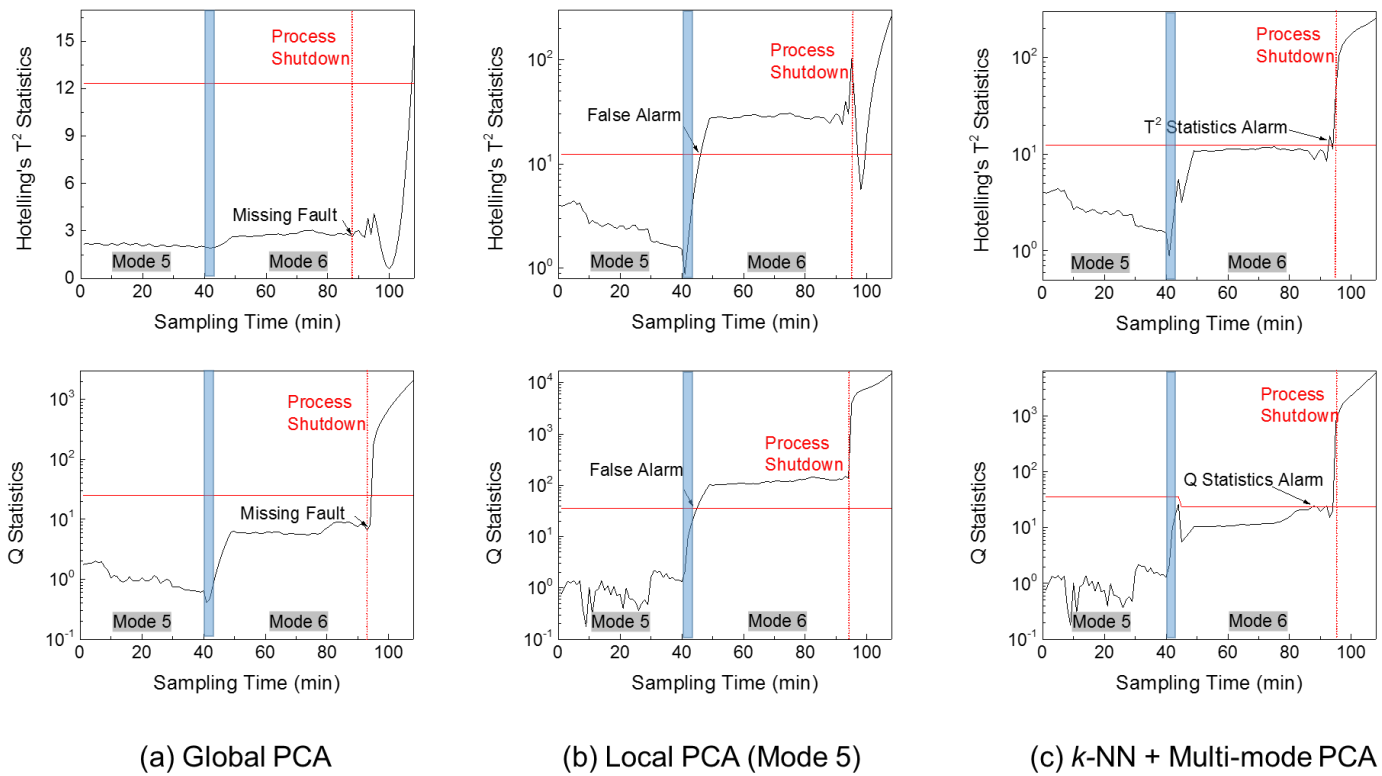

Figure 3-12. Monitoring performance for the fault case in mode 6: (a) global PCA, (b) local PCA (for mode 2), and (c) *k*-NN + Multi-mode PCA

When the operation condition changes from mode 2 to 3 and eventually process shutdown occurs at around 90 minutes (Figure 3-11), each PCA model shows different fault detection performance. The global PCA (Figure 3-11(a)) detects fault after the process shutdown. This is because detection sensitivity of the global PCA covering all of the operation modes is relatively low. The local PCA (Figure 3-11(b)) detects changed mode 3 as a fault. This comes from the fact when operation mode is mismatched with that in local PCA model, the local PCA can be subject to false alarming. The k-NN + PCA (Figure 3-11(c)), however, shows accurate fault detection at around 90 min in terms of Hotelling's T^2 statistics and at around 70 min in terms of Q statistics. The similar results are obtained when the operation condition changes from mode 5 to 6 and eventually process shutdown occurs (Figure 3-12).

As a result, it is proved that the proposed *k*-NN integrated multi-mode PCA methodology is more sensitive to the change of covariance structure caused by operation condition than conventional PCA. Therefore, when this proposed methodology is applied, an abnormal operating condition can be quickly detected and the generated information will help operators better recognize potential fault situations or symptoms. Correct actions, then, can be taken by the operators and process engineers to prevent a process shutdown, which can avoid catastrophic instrumental damage.

3.6 Conclusion

This paper proposes a systematic method for classifying process operation modes using *k*-NN algorithm and detecting process malfunction using multiple PCA models. Applying the operation mode classification based on the *k*-NN algorithm, it is possible to effectively classify the process operation modes according to changes in the process operating conditions. When the fault detection performance is evaluated with real LNG MR process data, the k-NN integrated multi-mode PCA shows more accurate and early detection capability even under operation mode change situations than conventional PCA. This indicates that false alarm rate can be reduced and enhance early fault detection capabilities even under operation mode changes.

Since LNG MR process contains various compressors in series and operates under high pressure and speed environments, early and accurate fault detection capability is a must for process safety and cost-effectiveness. In addition, process operation conditions change frequently. Under these circumstances, it is expected that the proposed *k*-NN integrated multi-mode PCA methodology will play an important role in monitoring the LNG MR process.

CHAPTER 4 : Estimation of disturbance propagation path using PCA and multivariate Granger Causality

4.1 Introduction

With the developments in the automation and control systems in the process and product industries, it is possible to collect enormous amount of data. However, analysis and interpretation of data is a key issue. Using various statistical tools, it is possible to quickly analyze the data to enhance both the process performance and reduce industrial waste, thereby, improving process economics. Online and offline process monitoring techniques are widely used in the industries to statistically analyze the process behavior. To ensure the smooth and safe operation of chemical plants, large number of sensors are usually used to record and analyze the data. However, with an increase in number of sensors, the chances for sensor faults in addition to the process faults have also been increased. Moreover, the occurrence of any fault in a system affects all the associated variables and disturbs their normal correlations which makes it difficult to detect the actual root cause of fault. Therefore, instant fault detection through analyzing the root cause and estimating fault propagation path in the system has always remained a key issue in process monitoring.

Data-driven models (Ge et al., 2013) have been widely used in semiconductor

manufacturing (Yue et al., 2000), chemical (Kosanovich et al., 1996) and steel industries (Miletic et al., 2004) exhibiting multi-level control hierarchy during the last few decades. Several univariate and multivariate statistical tools have been developed that can be efficiently used for both the fault detection and diagnosis. However, multivariate statistical methodologies are more preferred over the univariate techniques for analyzing the complex industrial system where process variables exhibits a strong correlations. Principal component analysis (PCA) is one of the effective multivariate statistical technique that finds its applications for process monitoring and control (Kresta et al., 1991; De Veaux et al., 1994), fault detection and diagnosis (Koutri and MacGregor, 1994; Raich and Cinar, 1996), and sensor validation(Tong and Crowe, 1995; Dunia et al., 1996) in various process industries. The PCA converts the higher dimensional correlated data into lower dimensional un-correlated data while retaining most of the original information. It can linearly reduce the dimensionality while ignoring the nonlinearities in the process data. PCA models the process behavior in terms of process variables during normal operation and compares the variation in those variables during fault situation. Hotelling's T^2 statistics and Q-statistic (squared prediction error (SPE)) are the two fault detection indices commonly used for analyzing the variation in process variables. The T^2 statistics represents the systematic part of process variation in principal component subspace (PCS), whereas, SPE shows the residual part of the process variation in the residual subspace (RS). The T^2 statistics and SPE can be calculated for each sample and compared with the confidence limits to monitor the

process faults. The occurrence of any fault in the process either affects the T^2 statistic or SPE of the samples or even changes both the statistics in some cases. Usually both the indices exceed their critical values during fault situation, whereas, the process is considered to be normal if both the indices remain under the control limits. T^2 statistics is usually used for overall process monitoring as it measures the variation among samples and indicate their distance from the center of the model. As multiple sensors are simultaneously affected by a certain fault, the contribution plot of samples shows the involvement of many variables towards a fault that makes it difficult to identify the actual fault variable. On the contrary, SPE measures the difference/residuals between the sample and its projection on the model. It measures the sum of variations in the RS which is not explained by the principle components (PC). SPE is highly sensitive to even minor process variations due to its smaller value compared to the T^2 statistic (Mujica et al., 2010). Qin also performed the detailed comparison for both the indices and showed that SPE can detect the fault more readily compare to the T^2 statistics. (Qin et al., 2003)

The PCA models combined with the fault detection indices finds it applications in various industries (Qin et al., 2003; Chiang et al., 2001; Bezergianni et al., 2008; Villegas et al., 2010; Yoon et al., 2004; Qin et al., 2012) to monitor the process variations and diagnose process faults. For an instance, Pehna applied the PCA model to the nuclear power plant and used the fault detection indices to monitor the temperature variation in the nuclear reactors. (Penha and Hines, 2001) Similarly, Ferrer used the T^2 statistics to analyze the process shift in the automobile

manufacturing industries. Landells et al. used the PCA methodology along with the statistical indices for early fault detection in refineries and other chemical production industries. (Ferrer et al., 2007; Landells and RaWi, 2012)

Several studies have relied on PCA models and used fault detection indices for early fault detection in the industrial systems. However, little attention has been paid towards identifying the fault propagation path in addition to the fault detection. As the fault in one variable affects the other associated variables, it is important to estimate the fault propagation path direction moving across the variables. Recently, Hong et al. developed the fault propagation path estimation algorithms based on the progressive PCA models. (Hong et al., 2010; Hong et al., 2011; Hong et al., 2014) The algorithm uses the SPE contribution plots to detect the variable having high contribution and identifies it as the fault variable. The detected variable is then eliminated and new PCA model is developed to identify the next variable. The detection of fault variable from each model in turns represent the order of variables affected by a certain fault. This methodology can be applied to the small systems with limited number of variables, however, it can become more complex and time consuming with an increase in number of process variables or unit processes.

Granger causality (GC) algorithms based on time series hypothesis is another tool that can evaluate the cause and effect relationship among variables. (Granger et al., 1969) Developing the logical interpretation of causal analysis along with the process knowledge can help in detecting the root cause of the faults. GC algorithms not only find its application in process and energy industries (Yuan et al., 2014; Yuan et al.,

2016; Landman et al., 2014) but also in economics (Granger et al., 1988) due to its ease of implementation and reliable interpretation of results. It uses a statistical hypothesis to predict whether one time series can affect the other time series or not. Additionally, the PCA models can be combined with the GC algorithms for efficient process monitoring. For instance, Li et al. developed the framework of locating the root cause of fault using both PCA and GC algorithms. (Li et al., 2016) Similarly, Landman et al. integrates the process causality and topology information to develop a causal model that determine the disturbance propagation path. (Landman et al., 2014) Moreover, Yuan et al. used the multilevel GC framework with clustering techniques for determining the causal relationship among plant-wide oscillatory variables. (Yuan et al., 2016) Ladman et al. also compared various statistical techniques for different industrial process and suggested that only one specific method is not powerful enough to predict the actual causal and effect relationship. (Landman et al., 2013) Therefore, several techniques should be used in addition with the process knowledge to obtain successful results.

The PCA models usually require a training data set for developing the model and to determine confidence limits for the fault detection indices. Then the real plant data or fault data can be used for detecting and diagnosing the fault. On the other hand, GC uses only the time series information of variables irrespective of the normal or fault data to evaluate the causal relationship among variables. The aim of this study is to develop a robust algorithm that can predict the fault propagation path in a system while verifying the causal relationship among those fault variables. Therefore, the

fault amplification methodology is first applied to the conventional PCA model to amplify the fault magnitude to identify the disturbance propagation path in terms of process variables. Then, the standalone multivariate GC (MVGC) methodology is used to identify the cause and effect relationship among the fault variables. The results from both the models are then compared to ensure the reliability of the fault propagation path obtained and verifying the causal relationship among variables at the same time. The developed models are applied to the LNG fractionation process and distillation columns operation where some of the common fault case scenarios are assumed to estimate the fault directions.

The paper is divided into four major sections. First section briefly explains the PCA model development process and the methodology for estimating the fault propagation path. The following section explains the time series multi-variate granger causality algorithm to determine the causal analysis among variables. The third section explains the LNG fractionation process taken as an exemplary case and discusses the results for the simulated fault case scenarios to determine the fault propagation path. Finally, the last section embodies the conclusion of the paper.

4.2 Theoretical Background

4.2.1 Fault propagation path detection

Occurrence of any fault in a system changes the value of variables in a way that the resultant contains sum of normal portion and fault portion of data. Therefore, it is important to reduce the impact of normal portion of data to amplify the fault effect. Qin and Valle developed an expressions to estimate the fault directions in the residual subspace of a fault data as given in Eq. (1). (Qin et al., 2003; Valle et al., 2001) For an instance, sample vector x under fault situation can be projected on the RS where $\boldsymbol{x}^*$ presents the fault free portion of the subspace and $\boldsymbol{\Xi_i f}$ represents the actual fault.

$$\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{\Xi_i f} \tag{1}$$

, where $\boldsymbol{\Xi_i}$ represents the orthonormality and ||f|| denotes the fault magnitude that subjects to change over time as the fault develops. The contribution of $\boldsymbol{x}^*$ projected on the RS is usually very small compared to the fault magnitude so it can be eliminated as given in Eq. (2).

$$\| \boldsymbol{x}^* \|^2 = SPE(\boldsymbol{x}) < \delta_\alpha^2 \tag{2}$$

On the contrary, moving data average techniques can be used to reduce the impact of normal variations in situations where the fault magnitude is not too large. (Granger, 1988) In an actual fault situation, $\boldsymbol{x}^*$ is usually unavailable and is overlapped with the fault data. Therefore, the removal of $\boldsymbol{x}^*$ from RS can be achieved by rescaling the residual matrix of the fault data with mean zero and unit variance of residuals of the normal matrix. The algorithm developed for PCA model to estimate the fault propagation path is given in Figure 4-1. The PCA model is developed by using the

normal operation data which determines the principal component loadings, principal component scores and the limits of fault detection indices. The developed model can be further utilized for the process monitoring using real time operational data or it can be compared with any hypothetical fault case scenario to estimate the fault direction. As fault detection is more accurate in the RS so the normal process variations are eliminated from the RS of the fault data to amplify the fault magnitude. The removal of $\boldsymbol{x}^*$ amplifies the fault portion of the data as given in Eq. (3).

$$\boldsymbol{x} = \boldsymbol{\Xi}_i \boldsymbol{f} \tag{3}$$

For an illustrative example, Xi in Eq. (4) represents the fault data collected for any fault case scenario containing n and m number of samples and variables, respectively.

$$\boldsymbol{X}_i = \begin{bmatrix} \boldsymbol{x_{11}} & \cdots & \boldsymbol{x_{1m}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{x_{n1}} & \cdots & \boldsymbol{x_{nm}} \end{bmatrix} \equiv [\, \boldsymbol{x_1} \;\; \boldsymbol{x_2} \;\; \ldots \;\; \boldsymbol{x_n}]^T \tag{4}$$

, where xi represents the row vector showing i_{th} sample corresponding to m variables. The projection of these samples on the RS containing the amplified fault data can be achieved from the above Eq. (3) as shown in Eq. (5).

$$\widetilde{\boldsymbol{X}}_i^T = \widetilde{\boldsymbol{\Xi}}_{\mathbf{i}} [\, \boldsymbol{f_1} \;\; \boldsymbol{f_2} \;\; \ldots \;\; \boldsymbol{f_n}] \tag{5}$$

, where $\widetilde{\boldsymbol{\Xi}}_{\mathbf{i}}$ and $\widetilde{\boldsymbol{X}}_i^T$ shares the same range space. The covariance matrix of fault data can be used to analyze the covariance among different variables (σ_{ij}) for

a particular number of samples using Eq. (3).

$$\text{Covariance } [\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}^{T}] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1m} \\ \vdots & \ddots & \vdots \\ \sigma_{m1} & \cdots & \sigma_{mm} \end{bmatrix} = [\sigma_{ij}]\,; \quad i,j = 1,2 \ldots m \quad (6)$$

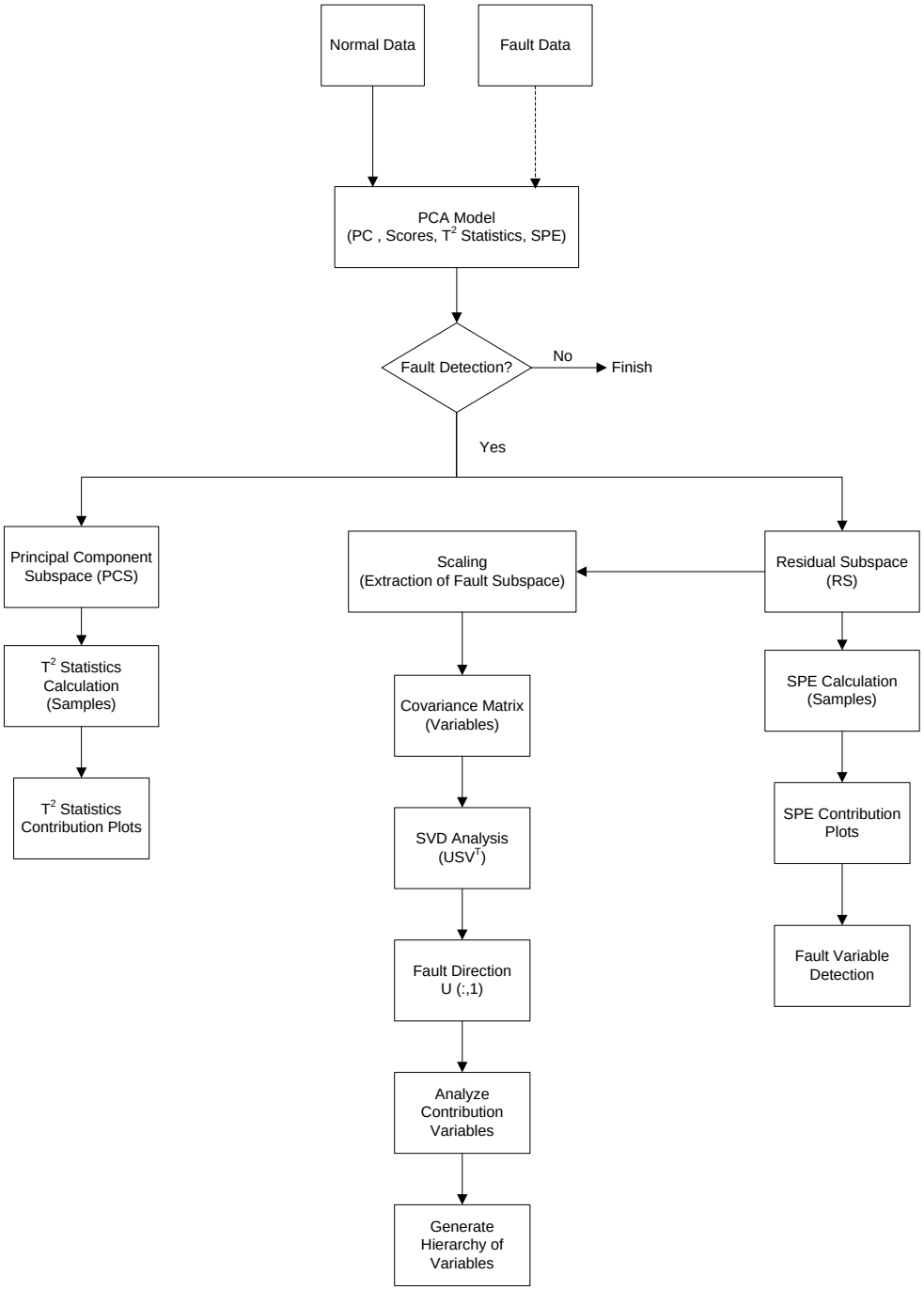

Figure 4-1. Algorithm for fault propagation path estimation

Performing SVD on the covariance matrix $\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{T}}$ and retaining singular values can help in transforming correlated variables into un-correlated variables. SVD decomposes the covariance matrix ($\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}$) into product of three matrices as given in equation 22.

$$\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}^{\boldsymbol{T}} = \boldsymbol{U_i D_i V_i^T} \tag{7}$$

, where $\boldsymbol{U_i}$ is the orthogonal matrix and represents the fault direction, $\boldsymbol{D_i}$ represents the diagonal matrix containing nonzero singular values arranged in descending order and $\boldsymbol{V_i^T}$ is the transpose of orthogonal matrix such that $\boldsymbol{U_i U_i^T} = \boldsymbol{U_i^T U_i} = \boldsymbol{I}$ and $\boldsymbol{V_i V_i^T} = \boldsymbol{V_i^T V_i} = \boldsymbol{I}$. The fault direction matrix can be chosen as $\widetilde{\boldsymbol{\Xi}}_{\mathbf{i}}$ where the first column of the orthogonal matrix $\boldsymbol{U_i}$ corresponding to the highest eigenvalue in the diagonal matrix $\boldsymbol{D_i}$ represents the maximum variation in the variables as represented in equation 5.

$$\widetilde{\boldsymbol{\Xi}}_{\mathbf{i}} = \boldsymbol{U_i}\,\mathbf{(:,1)} \tag{8}$$

The historical data of various fault case scenarios can be used in a similar way to extract the fault directions that could be used in future to identify the particular fault.

4.2.2 **Causal analysis based on Granger Causality (GC)**

In this section, the concept of time domain granger causality (GC) is discussed to determine the cause and effect relationship among various time series and variables. Barnett developed the multivariate granger causality (MVGC) toolbox in Matlab® that has been successfully used in various process industries to perform the causal analysis. (Barnett et al., 2014) The function of the GC is to determine whether one time series can causally affect the other time series or not. (Granger et al., 1969) The outcome of the causal analysis is the causality matrix, where each element (i, j) in the data matrix represents its causal relationship e.g. effect of variable i on variable j. For instance, if one time series (X_2) affects the other time series (X_1), than the knowledge of X_2 should help in determining the future values of X_1. For an illustrative example, consider two time series $X_1(t)$ and $X_2(t)$ and their bivariate autoregressive (AR) models as shown in equation 9 and equation 10.

$$X_1(t) = \sum_{t=1}^{k} a_{11,l} X_1(t-l) + \sum_{t=1}^{k} a_{12,l} X_2(t-l) + e_1(t) \quad (9)$$

$$X_2(t) = \sum_{t=1}^{k} a_{21,l} X_1(t-l) + \sum_{t=1}^{k} a_{22,l} X_2(t-l) + e_2(t) \quad (10)$$

, where $a_{ij,l}$, k and $e_i(t)$ represents the AR co-efficients, model order and the prediction errors of the model, respectively. Equation 9 and equation 10 are regarded as an un-restricted or full models. On the other hand, restricted model[35] can be generated by omitting all the co-efficient of a_{12} or a_{21} from the above

equations shown in equation 11 and equation 12 as follows.

$$X_1(t) = \sum_{t=1}^{k} b_{1,l} X_1(t-l) + e_{1(2)}(t) \quad (11)$$

$$X_2(t) = \sum_{t=1}^{k} b_{2,l} X_2(t-l) + e_{2(1)}(t) \quad (12)$$

After developing both the un-restricted and restricted model, the GC from X_2 to X_1 can be quantified by the ratio of variances as given in equation 13. Where, $e_{i(j)}$ and e_i represents the residuals/error terms of the restricted and un-restricted models, respectively.

$$F_{X_2 \rightarrow X_1} = \ln \frac{\text{var}(e_{i(j)})}{\text{var}(e_i)} \quad (13)$$

Bivariate GC concept can be extended to the multivariate case by defining the AR co-efficients as given in equation 14, where q^{-1} represents the backward shift operator.

$$A_{ij}(q^{-1}) = \sum_{t=1}^{k} a_{ij,l} q^{-1} \quad (14)$$

Equations 9 and 10 can be transformed for the multivariate case containing m number of variables as shown in equation 15.

$$\begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_m(t) \end{bmatrix} = \begin{bmatrix} A_{11}(q^{-1}) & A_{12}(q^{-1}) & \dots & A_{m1}(q^{-1}) \\ A_{21}(q^{-1}) & A_{22}(q^{-1}) & \dots & A_{m2}(q^{-1}) \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}(q^{-1}) & A_{m2}(q^{-1}) & \dots & A_{mm}(q^{-1}) \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_m(t) \end{bmatrix} + \begin{bmatrix} e_1(t) \\ e_2(t) \\ \vdots \\ e_m(t) \end{bmatrix} \quad (15)$$

It can be said that X_j causes X_i if excluding the X_j reduces the ability to predict future values of X_i and vice versa, even though all the other variables remain included in the regression model. If each X_j is removed once for predicting all other variables X_i such that $\forall i \neq j$, there will be (m-1)×m prediction error sequences and correspondingly (m-1)×m covariance matrix which can be denoted as $\sum_{i(j)} \equiv cov\left(e_{i(j)}\right)$. The cause and effect matrix of variances can be generated as shown in Table 4-1 , where the ij^{th} element is the co-variance matrix predicts the i^{th} row variable by excluding the j^{th} column variable for i≠j. Moreover, the diagonal elements shows the residual covariance of the k full models represented as $\sum_i = cov(e_i)$

Table 4-1. Structure for cause and effect matrix

	$\mathbf{X_1}$	$\mathbf{X_2}$	...	$\mathbf{X_m}$
$\mathbf{X_1}$	$\sum_1$	$\sum_{1(2)}$	...	$\sum_{1(m)}$
$\mathbf{X_2}$	$\sum_{2(1)}$	$\sum_2$	...	$\sum_{2(m)}$
⋮	⋮	⋮	⋱	⋮
$\mathbf{X_m}$	$\sum_{m(1)}$	$\sum_{m(2)}$	...	$\sum_m$

The diagonal elements of the matrix are usually smaller than the other elements of the same row since un-restricted (full) model use more inputs for the regression to fit the data. If the variable X_j has more causal effect on the X_i then the ij^{th} element in the matrix must be greater than the ii^{th} element. Therefore, the GC from X_j to X_i can be determined by the GC index as shown in equation 16.

$$F_{j\rightarrow i} = \ln \frac{\sum_{i(j)}}{\sum_i} \quad (16)$$

Moreover, the statistical significance can be established by using the F-statistical test prior making any inference from GC results as given in equation 17.

$$F = \frac{RRS_r - RRS_{ur}}{RRS_{ur}} \times \frac{T-2p-1}{p} \quad (17)$$

, where RRS_r and RRS_{ur} are the residual sum of squares for the restricted and un-restricted models, respectively. On the other hand, T and p represents the total number of observations and the model order, respectively.

4.3 Application to the Liquefied Natural Gas (LNG) Process

4.3.1 Process Description

LNG (Liquefied Natural Gas) fractionation process has been selected in this study to evaluate the stability and accuracy of the developed methodologies. LNG plant consists of three major sections i.e. pretreatment section, liquefaction section and post treatment section. Pre-treatment section mainly includes the acid gas removal (AGR) unit and dehydration unit. Liquefaction unit contains the cryogenic heat exchanger network to liquefy the gas streams and the post-treatment unit mainly include the fractionation unit. The natural gas (NG) from the extraction wells contains higher concentration of methane and natural gas liquids (NGL's) such as ethane, propane, butane and fraction of higher hydrocarbons in addition to the acid gases (H_2S, CO_2). Fractionation units are usually used to obtain high purity methane and to separate the NGLs. In this study, the dynamic model of LNG fractionation process is developed in Aspen HYSYS® for multivariate statistical process monitoring and fault diagnosis. The data is generated through computer simulations and various fault case scenarios are developed to statistically examine the fault propagation path algorithms. The process flow diagram of the fractionation process is represented in Figure 2. Four cryogenic distillation columns namely de-methanizer, de-ethanizer, de-propanizer and de-butanizer are installed in series to achieve high percentage of methane, ethane, propane and butane, respectively. Table 4-2 shows the composition of product streams from each column of fractionation unit during normal operation.

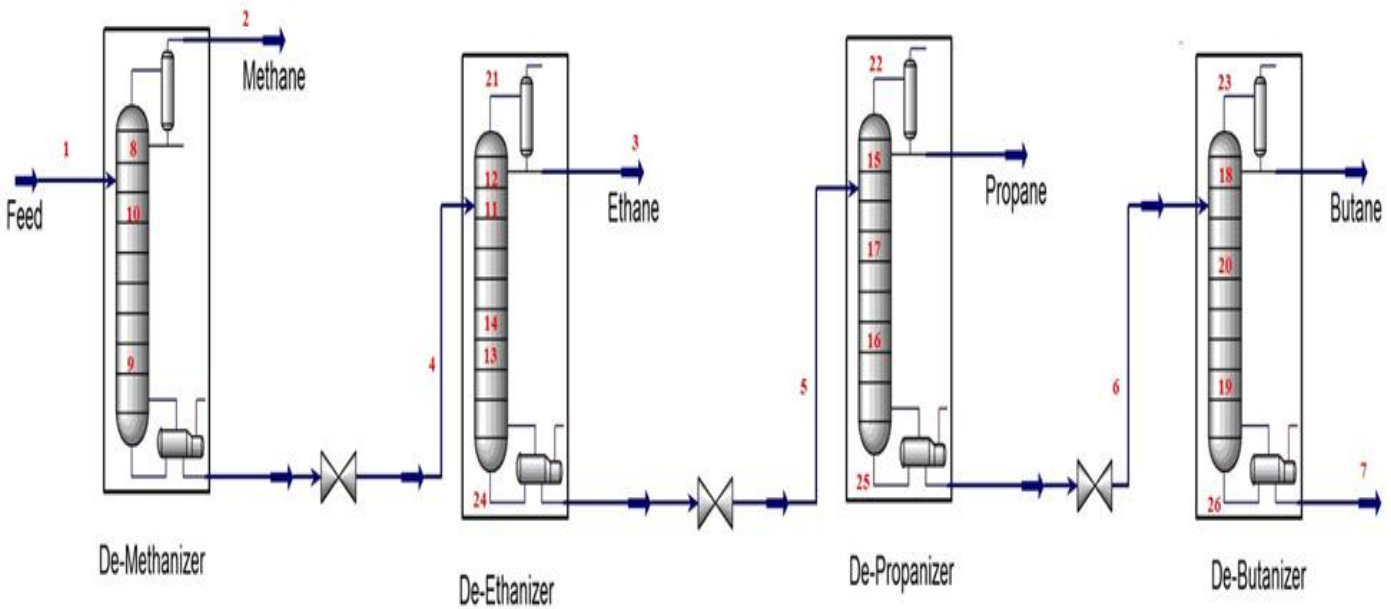

Figure 4-2. Schematic Diagram of LNG Fractionation Proces

Table 4-2. Stream compositions for fractionation units.

Component	Quality		Stream
Methane	C1	91.60%	Top of De-Methanizer
	C2	5.31%	
	C3	2.06%	
	C4	0.80%	
	Others	0.23%	
Ethane	C2	99.55%	Top of De-Ethanizer
	Others	0.45%	
Propane	C3	99.61%	Top of De-Propanizer
	Others	39.00%	
Butane	C3	1.23%	Top of De-Butanizer
	C4	98.00%	
	Others	0.77%	

4.3.2 Development of fault case scenarios

Number of fault cases scenarios can be developed by manipulating various variables during dynamic simulation. The change in any process variables (stream flow rates, temperatures, pressures or the reboiler's heat duty) readily affects the other associated variables. Figure 4-2 represents the NGL fractionation process considered in this study. The numbers in red shows the location of specific variables.

4.3.2.1 Malfunction in the de-ethanzier column

In this case, de-ethanizer column malfunctioning has been considered to analyze its effect on overall NGL fractionation process. Thirty two different variables are selected from various sections of the fractionation process where the prefix and suffix denotes the fractionation column and the specific process variables, respectively, as shown in table 4-3. For an instance, the prefix dM, dE, dP and dB represents the de-methanizer, de-ethanizer, de-propanizer and de-butanizer column respectively. Whereas, the suffix 1T, 2P and 5F represent the temperature of first stream, pressure of second stream and flow rate of fifth stream, respectively.

Table 4-3. Stream variables used for PCA model and fault case scenario

No	Stream Number	Variable Name
1	dM-1T	De-methanizer Stream 1 Temperature
2	dM-2P	De-methanizer Stream 2 Pressure
3	dM-2T	De-methanizer Stream 2 Temperature
4	dE-3T	De-ethanizer Stream 3 Temperature
5	dE-4F	De-ethanizer Stream 4 Flow Rate
6	dE-4T	De-ethanizer Stream 4 Temperature
7	dP-5F	De-propanizer Stream 5 Flow Rate
8	dP-5T	De-propanizer Stream 5 Temperature
9	dB-6F	De-butanizer Stream 6 Flow Rate
10	dB-6T	De-butanizer Stream 6 Temperature
11	dB-7T	De-butanizer Stream 7 Temperature
12	dM-8P	De-methanizer Stream 8 Pressure
13	dM-9P	De-methanizer Stream 9 Pressure
14	dM-10T	De-methanizer Stream 10 Temperature
15	dE-11P	De-ethanizer Stream 11 Pressure
16	dE-12P	De-ethanizer Stream 12 Pressure
17	dE-13P	De-ethanizer Stream 13 Pressure
18	dE-11T	De-ethanizer Stream 11 Temperature
19	dE-14T	De-ethanizer Stream 14 Temperature
20	dP-15P	De-propanizer Stream 15 Pressure
21	dP-16P	De-propanizer Stream 16 Pressure
22	dP-17T	De-propanizer Stream 17 Temperature
23	dB-18P	De-butanizer Stream 18 Pressure
24	dB-19P	De-butanizer Stream 19 Pressure
25	dB-20T	De-butanizer Stream 20 Temperature

No	Stream Number	Variable Name
26	dE-21P	De-ethanizer Stream 21 Pressure
27	dE-21T	De-ethanizer Stream 21 Temperature
28	dP-22P	De-propanizer Stream 22 Pressure
29	dB-23P	De-butanizer Stream 23 Pressure
30	dE-24T	De-ethanizer Stream 24 Temperature
31	dP-25T	De-propanizer Stream 25 Temperature
32	dB-26T	De-butanizer Stream 26 Temperature

Some of the common factors that can affect column's operation includes the un-controlled reboiler heat duty, failure of the steam valves and sensor faults[36]. In this case, fault is introduced in the distillation column by increasing the reboiler's heat duty by a step change to analyze its effect on the other variables. The data from the dynamic simulations is recorded to statistically analyze the process variations and detecting the fault propagation path. Figure 4-3 represents the dynamic behavior of some process variables for both the normal and abnormal operation. It can be seen from the results of dynamic model that the process variables were kept within their limits with small oscillation around the set point, whereas, the fault graph shows a steep rise in process variables for the fault case scenario.

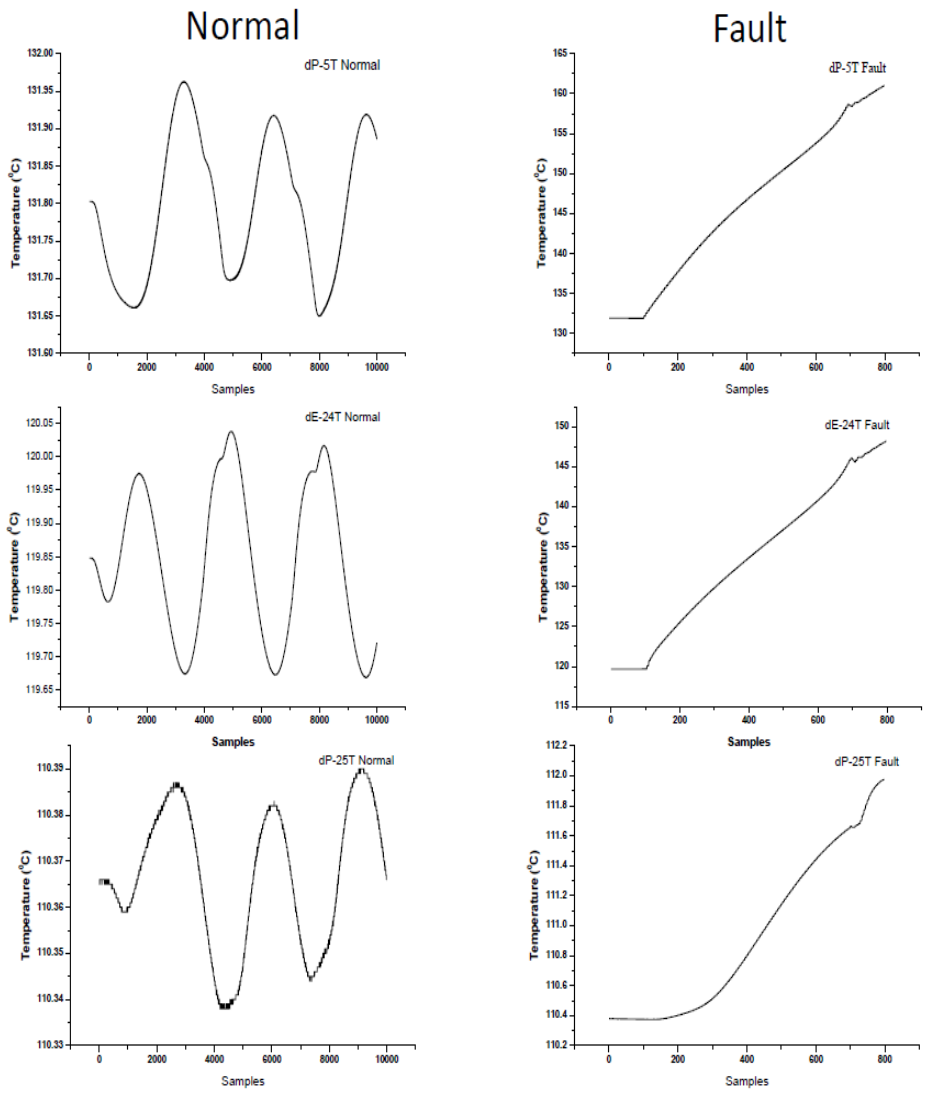

Figure 4-3. Dynamic behavior of process variables for normal and fault operation.

The PCA algorithm is developed using a training data and tested for fault propagation path estimation, whereas, MVGC analysis is performed to evaluate the reliability of fault propagation path through analyzing the cause and effect relationship among variables. Following the algorithm given in figure 1, the scores are generated for both the normal and abnormal operation which are represented in Figure 4-4. The results showed that all the scores are plotted within the process limits on PC1 and PC2, however, the scores tend to plot outside the control limits in case of fault situation. For detailed analysis, the fault detection indices are used to determine the process limits for a normal operation and to compare them with the sample's value for detecting the fault variable. Using equation 9 and equation 10, the confidence limits for the T^2 statistics and SPE for the normal process is calculated as 14.715 and 17.61, respectively.

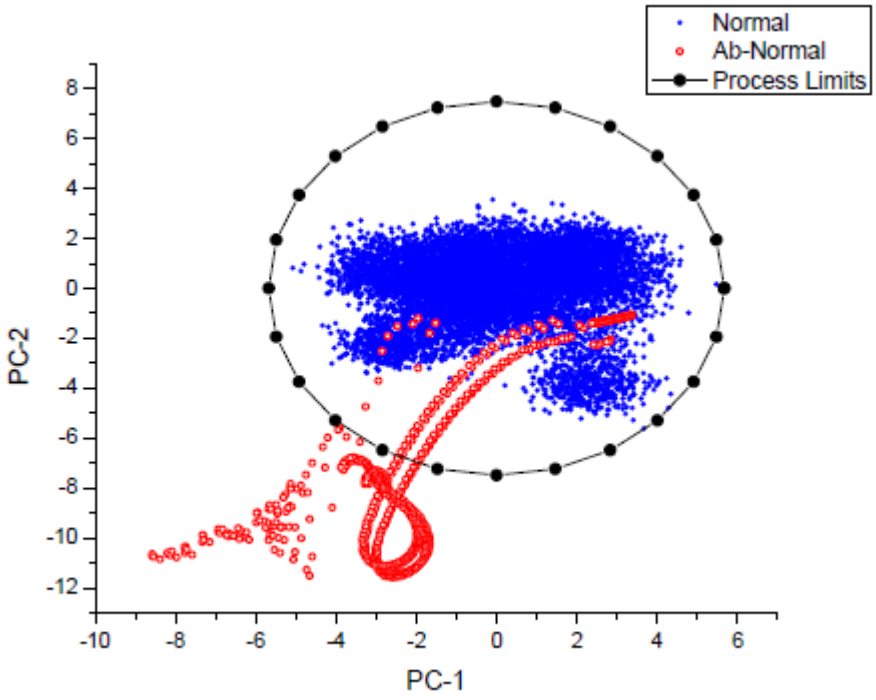

Figure 4-4. Score plot for normal and fault operation on principal components.

The Figure 4-5 shows the T^2 statistical limit for normal operation (red line), whereas, the samples T^2 statistics is represented as blue line. The results showed that the 225^{th} sample exceeds the upper confidence limit of T^2 statistics. However, the contribution plot of 225^{th} sample given in Figure 4-6 shows the involvement of many variables towards a fault that makes it difficult to accurately identify the fault variable. On the other hand, the SPE index can be alternatively used for detecting the fault variable. The Figure 4-7 shows the upper confidence limit (red line) of SPE for the normal process as well as the SPE for each sample. The results showed that the sample 103 exceeds the upper confidence limit of SPE, where, the contribution plot in Figure 4-8 shows that the variable dP-5T has the highest contribution towards a fault. After developing the fault detection indices, fault propagation path estimation methodology is applied to determine the hierarchy of variables affected by a particular fault. As fault amplification is an important section of the fault propagation path estimation algorithm, fault is amplified in the RS as given in Figure 4-9.

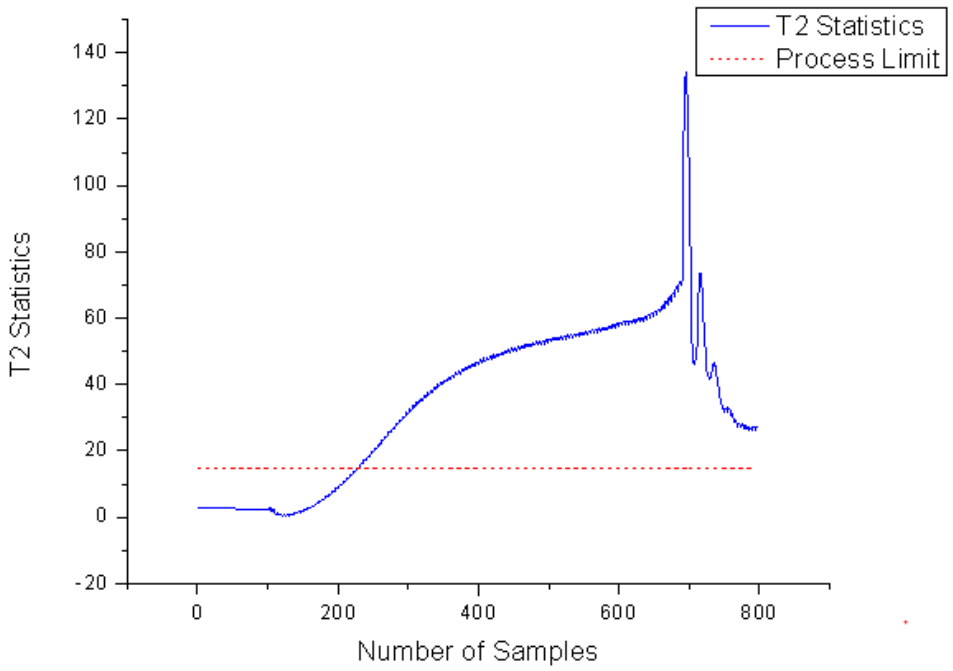

Figure 4-5. Fault detection using T^2 Statistic

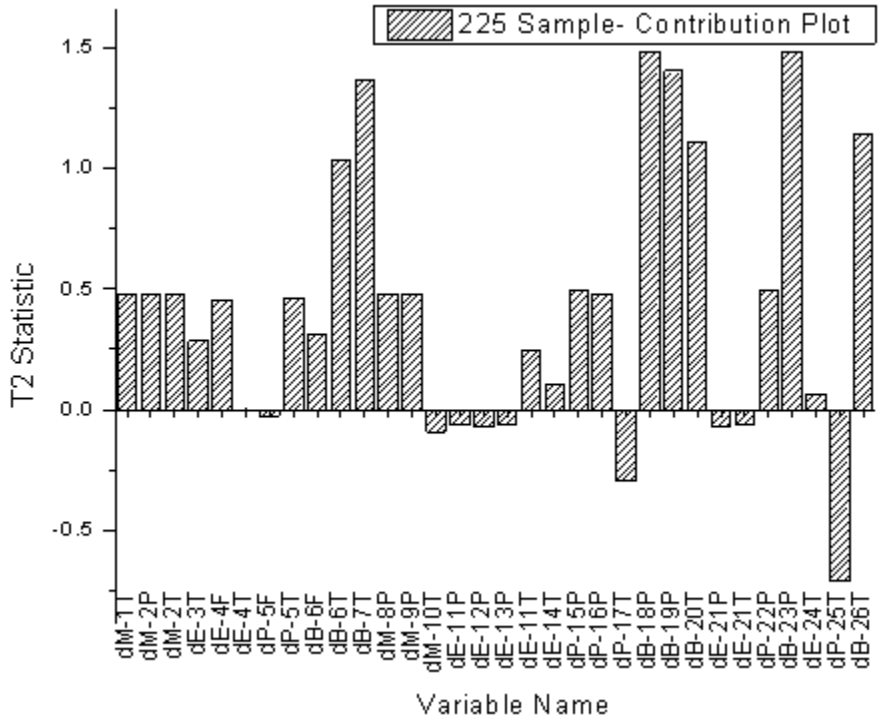

Figure 4-6. Contribution plot of variables using score matrix of fault data.

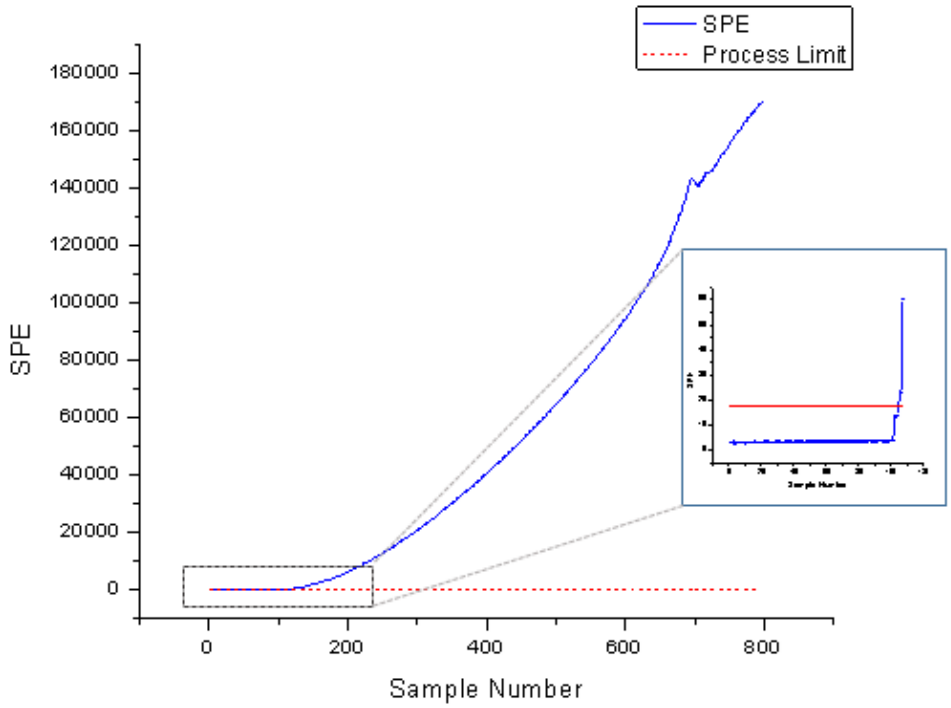

Figure 4-7. Fault detection using SPE index.

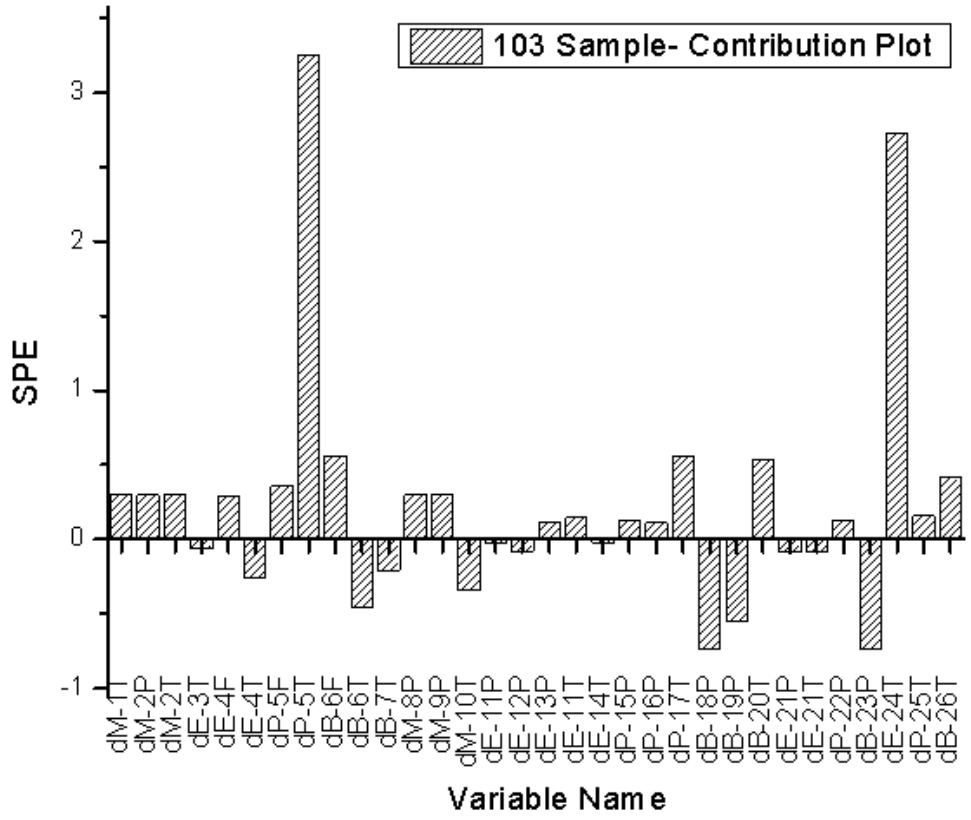

Figure 4-8. Contribution plot of variables using residual matrix of fault data.

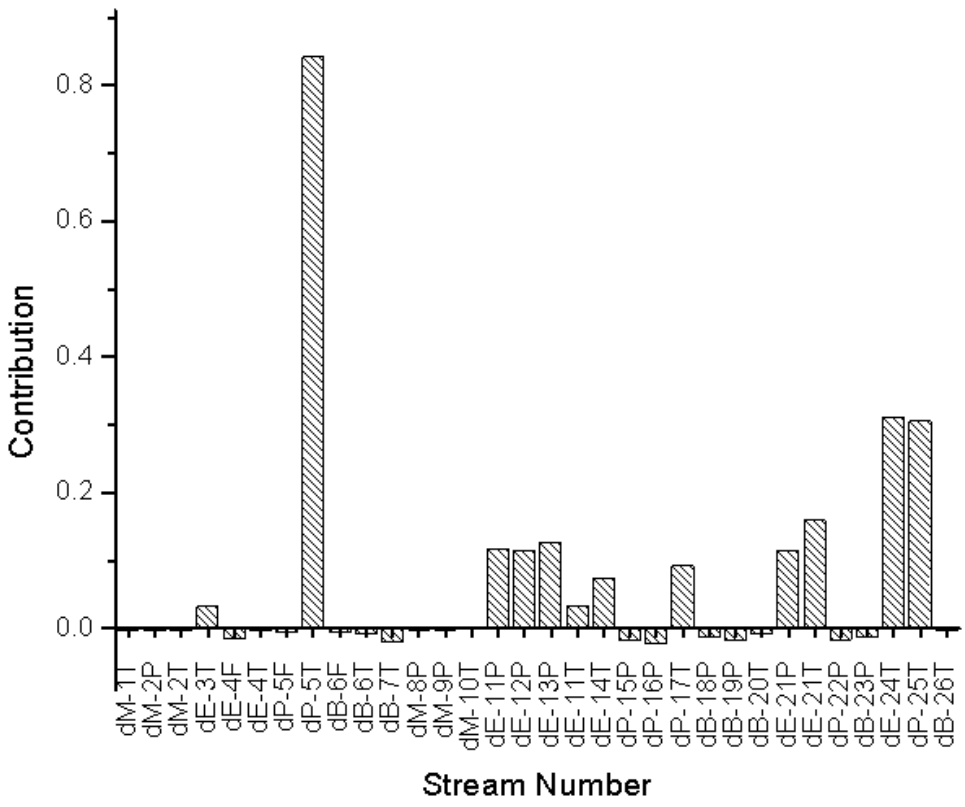

Figure 4-9. Contribution of variables in a fault case scenario.

Additionally, the developed algorithm followed the absolute descending order functions that re-arranged the variables according to their magnitude of contribution. The fault propagation path obtained in this case is represented in Figure 4-10 where the x-axis represents the hierarchy of variables affected by the developed fault. It can be seen from results that variables (dP-5T, dE-24T, dP-25T, dE-21T, dE-13P and dE-11P) associated with the de-ethanizer and de-propanizer columns are readily affect by the fault. On the other hand, the upstream variables (dM-1T, dM-2P, dM-9P, dM-8P, dM-2T, dM-10T) associated with the de-methanizer column are least affected by the fault. Furthermore, the variables with the higher and least contributions towards a fault are automatically rearranged at the left and right of the graph, respectively. The variable showing the highest contribution (dP-5T) is most likely to be the fault variable, whereas, the variable showing the least contribution (dM-10T) is considered as the least affected variable. Moreover, the difference between Figure 4-8 and Fig 4-9 can be clearly seen in terms of difference of fault magnitude among variables. Figure 4-8 highlights the only first few affected variables (dP-5T, dE-24T) that could be used to identify the root cause of fault. On the other hand, Fig 9 clearly showed an amplification in fault magnitude of all the affected variables (dP-5T, dE-24T, dP-25T, dE-21T, dE-13P, dE-11P, dE-21P, dE-12P and so on) depending on their contribution towards a fault that makes it easy to determine the fault propagation path.

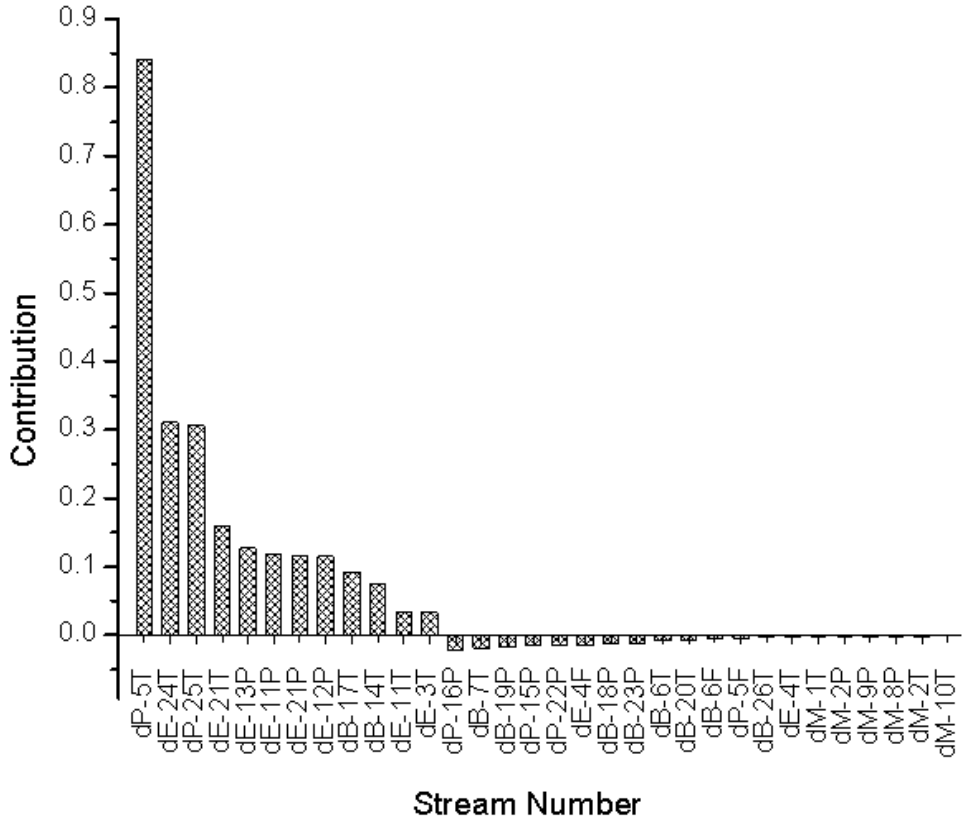

Figure 4-10. Fault propagation path for de-ethanizer column malfunction.

As discussed earlier in the introduction section, any single methodology is not powerful enough for the purpose of fault diagnosis, therefore, a combination of techniques are usually employed for the accurate and reliable process monitoring. To validate the accuracy of PCA model results, the fault propagation path methodology has been also cross checked using the MVGC analysis to evaluate the cause and effect relationship among variables. The outcome of the MVGC is the causal matrix that represents the cause and effect among variables. Thirty two variables have been chosen for analysis in this study. Since, it is difficult to draw a visualization of all the variables, only first ten variables (dP-5T, dE-24T, dP-25T, dE-21T, dE-13P, dE-11P, dE-21P, dE-12P, dB-17T and dB-14T) from the fault propagation path are used for causal analysis. Table 4-4 shows the causal matrix generated for the above mentioned variables.

Table 4-4. Causal matrix for de-ethanizer column malfunction

	dP-5T	dE-24T	dP-25T	dE-21T	dE-13P	dE-11P	dE-21P	dE-12P	dB-17T	dB-14T
dP-5T	-	0.0237	0.0092	0.0094	0.0057	0.0060	0.0044	0.0049	0.0049	0.0063
dE-24T	0.0682	-	0.0090	0.0055	0.0037	0.0027	0.0070	0.0032	0.0091	0.0164
dP-25T	0.0637	0.0338	-	0.0141	0.0086	0.0041	0.0043	0.0031	0.0096	0.0116
dE-21T	0.0464	0.0153	0.0190	-	0.0120	0.0055	0.0044	0.0046	0.0058	0.0083
dE-13P	0.0456	0.0084	0.0149	0.0143	-	0.0104	0.0086	0.0038	0.0065	0.0085
dE-11P	0.0373	0.0052	0.0078	0.0125	0.0171	-	0.0075	0.0037	0.0053	0.0106
dE-21P	0.0336	0.0043	0.0062	0.0043	0.0099	0.0133	-	0.0085	0.0117	0.0156
dE-12P	0.0348	0.0091	0.0065	0.0046	0.0088	0.0082	0.0126	-	0.0197	0.0117
dB-17T	0.0195	0.0048	0.0020	0.0033	0.0067	0.0086	0.0109	0.0098	-	0.0346
dB-14T	0.0187	0.0083	0.0019	0.0041	0.0023	0.0026	0.0068	0.0044	0.0202	-

The causality matrix shows the impact of one variable on other variables where each of the column variable show its effect on the all the row variables. For instance, the dP-5T variable in the first column affects all the row variables where the highest affect is received by the variable dE-24T. Similarly, all the column variables are evaluated to determine their effect on the row variables. For simplicity, table 4 shows the highlighted (grey) cells to represent the maximum causal relationship among respective column and row variables. Moreover, Figure 4-11 represents the causal relationship among variables to determine how the variation in one process variable can trigger the fault in the other variable. The bold arrows represents the sequential effect of one variable on the other variable. On the other hand, the thin arrows represents the second highest effect of the process variables on each other that can be simultaneously affected. The results showed that the hierarchy of variables obtained from MVGC analysis and the fault propagation path algorithm are in good agreement which represents the robustness of the developed algorithm.

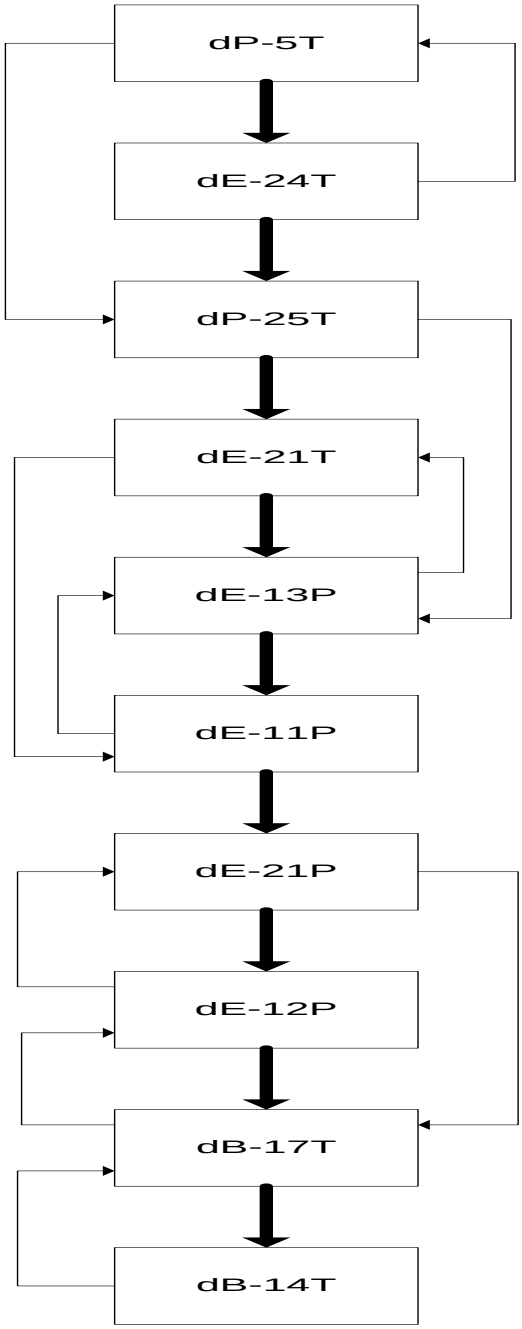

Figure 4-11. Causal flow for de-ethanizer column malfunction

4.3.2.2 Flooding in the de-propanizer column

Flooding is one of the most common problem in the distillation column operation caused by various factors including excessive vapor generation at the bottom of column, un-suitable reflux rate, increase feed flow rate and so on. (Kister et al., 1990; Kister et al., 1992) As the temperature and pressure are highly correlated at each stage due to vapor liquid equilibrium, increase in plate's temperature cause an increase in the pressure gradient too across the column. (Gorak and Schoenmakers, 2014) To analyze the flooding scenario in the distillation column, excessive vapor generation case is selected in this study by manipulating the reboiler's heat duty. The increase in reboiler heat duty increases the plate's temperature across the column which also affects the product purity. The excessive vapors generation at the bottom of the column not only entrains the liquid to the above stages causing an increased liquid hold up on the trays but also plug the downcomers with the excessive liquid. As a result of flooding, the flow rate of liquid down to the column decrease that results in an increase of ΔT and ΔP across the column. (Cheremisioff et al., 2000) In this case study, the temperature of ten stages (ST-44, ST-39, ST-28, ST-26, ST-20, ST-18, ST-16, ST-8, ST-4 and ST-1) of the de-propanizer column have been selected to generate the data. The temperatures of bottom and top stage of the distillation column are represented by ST-44 and ST-1, respectively. As done earlier, the PCA based fault propagation path detection and MVGC algorithms are used to analyze both the disturbance propagation path in the system and to determine the cause and

effect relationship among variables.

The data for the ten variables from de-propaznizer column is generated through dynamic simulation for both the normal and fault case scenarios. PCA model and the fault detection indices are developed using the normal (training) data and tested against the fault data to detect both the fault variable and to develop the disturbance propagation path. Figure 4-12 shows the score plot for both the normal and abnormal operation represented as blue and red points, respectively, on PC1 and PC2. Using the algorithm given in Figure 4-1, the confidence limits for the fault detection indices are developed and compared with the sample's T^2 statistics and SPE. Furthermore, the developed algorithm is tested for the flooding case to determine the fault propagation path in the de-propanizer distillation column. Figure 4-13 represents the fault propagation path in terms of change in the stage's temperature that are affected by an increase in reboiler's heat duty. For instance, the fluctuation in reboiler's heat duty more readily affects the temperature of bottom plates (ST-44, ST-39) compared to the top plates (ST-4, ST-1) of the column, the similar effect can be seen from the fault propagation path estimation results. The results from the fault propagation path algorithm also shows that the faults propagates from bottom to the top of the column where the variables showing the highest and least contribution are arranged to the left and right of the graph, respectively. Moreover, the results obtained from the fault propagation path has been cross checked with the MVGC analysis to analyze the cause and effect relationship among fault variables. The causal matrix obtained from the MVGC corresponding to eleven variables is given in Table 4-5.

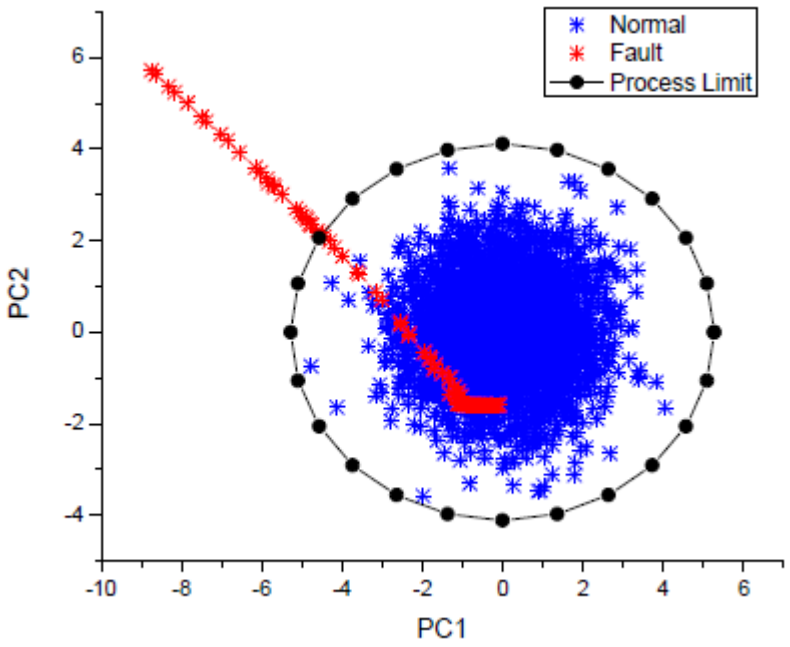

Figure 4-12. Score plot for training and fault data on principal components for flooding in de-propanizer column

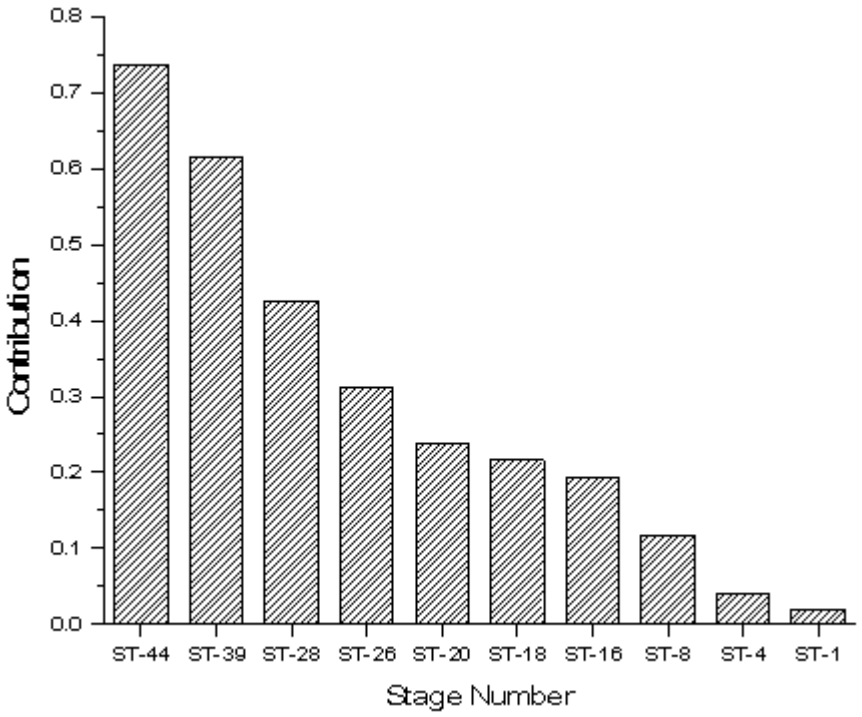

Figure 4-13. Fault propagation path estimation for flooding in de-propanizer column.

Table 4-5. Causal matrix for de-propanizer distillation column flooding

	ST-44	ST-39	ST-28	ST-26	ST-20	ST-18	ST-16	ST-8	ST-4	ST-1
ST-44	-	0.0356	0.0165	0.0075	0.0044	0.0050	0.0023	0.0034	0.0069	0.0108
ST-39	0.1159	-	0.0136	0.0050	0.0057	0.0030	0.0030	0.0037	0.0134	0.0172
ST-28	0.0488	0.0370	-	0.0077	0.0049	0.0022	0.0046	0.0026	0.0139	0.0254
ST-26	0.0672	0.0203	0.0262	-	0.0120	0.0039	0.0056	0.0048	0.0182	0.0203
ST-20	0.0481	0.0153	0.0089	0.0115	-	0.0143	0.0094	0.0043	0.0083	0.0135
ST-18	0.0384	0.0124	0.0058	0.0086	0.0212	-	0.0107	0.0060	0.0063	0.0140
ST-16	0.0357	0.0102	0.0047	0.0069	0.0119	0.0155	-	0.0126	0.0166	0.0194
ST-8	0.0338	0.0084	0.0077	0.0040	0.0104	0.0094	0.0176	-	0.0238	0.0176
ST-4	0.0178	0.0046	0.0030	0.0016	0.0068	0.0076	0.0171	0.0133	-	0.0435
ST-1	0.0159	0.0104	0.0102	0.0020	0.0014	0.0020	0.0049	0.0054	0.0248	-

It can be seen from results that each stage showed a strong causal relationship with the next stage. Where, the highlighted cells show the maximum causal effect of column variables to the corresponding row variables. For instance, ST-28 variable in the column affects the ST-26 variable in the row. Similarly, the hierarchy of variables are generated that are affected by the fault from cause and effect matrix. Figure 4-14 shows the sequential pattern of variables (bold arrows) that are affected during distillation column flooding. On the other hand, thin arrows represent the second highest causal relationship among variables that can simultaneously affect each other. Moreover, the results showed that the hierarchy of variables generated from the fault propagation path follows the similar pattern as obtained from MVGC analysis. Therefore, the developed algorithm for fault amplification can be efficiently used for determining the fault propagation path in a system.

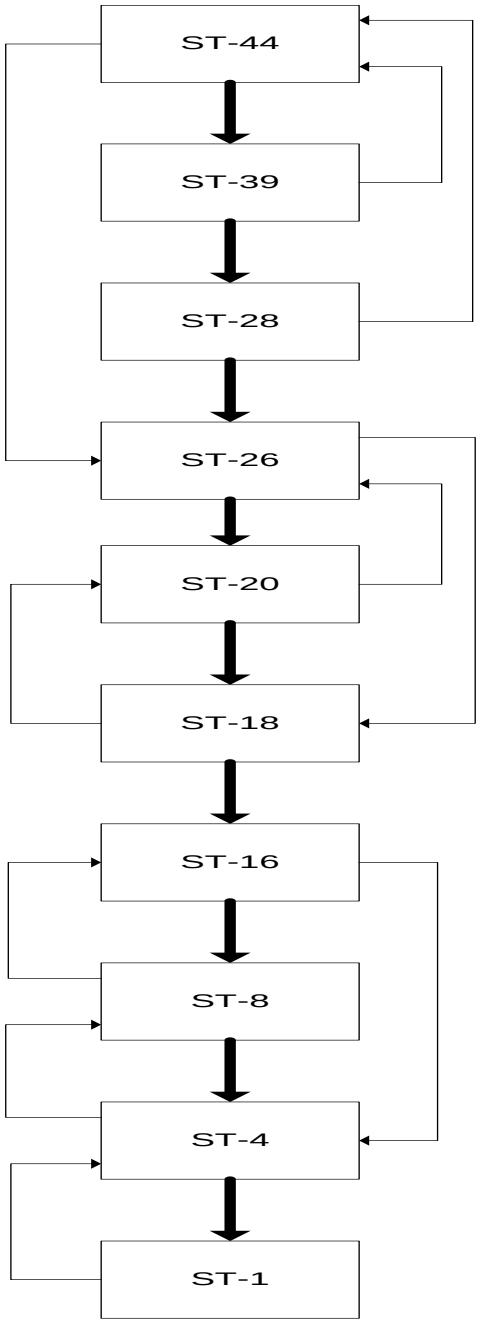

Figure 4-14. Causal flow for flooding in de-propanizer column.

4.4 Conclusion

With the increase in complexities of industrial systems, multivariate data analysis methodologies for process monitoring and fault diagnosis is getting a lot of attention. The occurrence of any fault in the process readily affect the associated variables and disturbs their normal co-relations. Hence, instant fault detection and identifying the root cause of fault is of great importance. In this study, PCA based fault amplification methodology has been developed for estimating the fault propagation path in the industrial systems. The developed algorithm projects the samples on the RS to estimate the fault directions. The RS is usually masked by the normal process variations in addition to the fault data, therefore, fault data is rescaled with the mean average and unit variance of the residuals of the normal data to maximize the fault magnitude in the resultant matrix. Furthermore, the covariance of the fault amplified RS is generated followed by the SVD analysis to generate orthogonal matrix corresponding to the largest eigenvalue which in turn represents the hierarchy of variables contributing to the certain fault. Using absolute descending order functions, the variables are arranged according to their magnitude of contribution towards a fault which in turn represent the fault propagation path. The hierarchy of variables obtained from fault propagation path is verified by using the time domain MVGC analysis that highlights the causal relationship among variables. In this study, LNG fractionation process train and distillation column operation is considered to test the developed algorithms. The comparative analysis obtained for both case studies showed that the developed fault propagation path and the causal pattern obtained from the MVGC analysis are in good agreement. Therefore, the developed algorithm can be used in various process and product development industries for not only

estimating the fault propagation directions but also for the instant fault detection.

CHAPTER 5 **Concluding Remarks**

This thesis addressed process monitoring and fault diagnosis strategies for industrial process such as early detection of a process malfunction or mode changes, increase sensitivity of process sensors and fault propagation path estimation. In order to satisfy the purpose of the study, the process monitoring and fault diagnosis methodology developed and validated LNG process in Incheon and lab-scale semiconductor etching process. In the development and validation of process monitoring and fault diagnosis algorithm to codes, various tools such as excel VBA, MATLAB® with Classification Learner, Visual C++ and Plant Analyst are conducted. In order to handle a huge amount of process operation data, MATLAB or Plant Analyst among commercial packages are more suitable than spreadsheet when monitoring method is based on PCA and Granger Causality. In order to simulate dynamic model, Aspen HYSYS also used for design LNG Plant.

In chapter 2, DWT filter was introduced to maximize the SNR ratio and remove noise source due to the measurement characteristics of the CCD array. The photon shot noise is the most dominant source of noise in OES and has different frequency characteristics from signal, so separation of noise and signal through DWT filter was effective. As a result of introducing DWT filter to OES signal, SNR ratio increased from 300 to 1300. After the noise signal associated with CCD array device characteristics such as shot noise is removed through the DWT filter, only the signal intensity due to process variation is measured. Therefore, it is possible to detect more

sensitive process condition change when compared with before implementation. Applying the variable selection algorithm to the measured OES signal, only 20 signals related to the reaction gas among 2048 signals were selected. After variable selection, we generated PCA model using 20 selected signals and testing detectability in multivariate statistical model in case of changes in source power and pressure. As a result, in the PCA model using DWT filtering and variable selection introduced in this study, the minimum detectable variation was reduced from 1.33% to less than 0.67%.

In Chapter 3, the development of the operation mode classification method according to the process operation condition enables the process monitoring with the process monitoring model having the most similar dynamic characteristics to the current process state. Therefore, compared to the existing methodology, we were able to reduce the problem of false alarms and missing faults. With the process monitoring methodology, several compressor failure were detected and prevented. Some other benefits realized from this system at LNG testbed plant installed in the Incheon include; supporting daily operation to achieve stable casting operation by providing real-time indications of process health during operation; and helping operators and process engineers better understand the process by troubleshooting operation failures using multivariate statistical tools.

In the Chapter 4, hierarchy of variables obtained from fault propagation path is verified by using the time domain MVGC analysis that highlights the causal relationship among variables. In this chapter, LNG fractionation process train and distillation column operation is considered to test the developed algorithms. The comparative analysis obtained for both case studies showed that the developed fault

propagation path and the causal pattern obtained from the MVGC analysis are in good agreement. Therefore, the developed algorithm can be used in various process and product development industries for not only estimating the fault propagation directions but also for the instant fault detection.

In addition to the existing DCS-based statistical process monitoring and control techniques, the DWT filtering technique to improve the sensitivity of the process sensor, a method of improving the performance of the PCA model through variable selection in Chapter 2, combined with the early detection by the multimode PCA, the process mode classification method in Chapter 3, and the method for finding the root cause using the fault propagation path in chapter 4, the operator or process engineer can help achieve the goal of stable process operation and improvement of process operation efficiency. Related to Industry 4.0, smart manufacturing, or the industrial Internet of Things is the technological evolution from embedded systems to digital manufacturing and production systems powered by big data and advances in technology, such as artificial intelligence, rapid automation, machinery, and additive manufacturing. Related to Industry 4.0, our research contributes to the design and implementation of process automation and efficient process monitoring systems.– productivity, ensuring consistency in quality, reducing costs, and optimizing inventory.

Nomenclature and Abbreviations

AGR: Acid gas removal

DWT: Discrete wavelet transform

GC: Granger Causality

LNG: Liquefied natural gas

MVGC: Multi variate granger causality

NG: Natural gas

NGL: Natural gas liquids

PCA: Principal component analysis

PC: Principal component

PCS: Principal component subspace

RS: Residual subspace

SPE: Squared prediction error

SVD: Singular value decomposition

Literature cited

Ahmed, U., Ha, D., Shin, S., Shaukat, N., Zahid, U., & Han, C. (2017). Estimation of disturbance propagation path using principal component analysis (PCA) and multivariate granger causality (MVGC) techniques. Industrial & Engineering Chemistry Research, submitted.

Baek, K. H., Jung, Y., Min, G. J., Kang, C., Cho, H. K., & Moon, J. T. (2005). Chamber maintenance and fault detection technique for a gate etch process via self-excited electron resonance spectroscopy, *Journal of Vacuum Science and Technology B 23*, (1), 125-129.

Baek, K. H., Edgar, T. F., Song, K., Choi, G., Cho, H. K., & Han, C. (2014). An effective procedure for sensor variable selection and utilization in plasma etching for semiconductor manufacturing, *Computers & Chemical Engineering 61*, (11), 20-29.

Barnett, L. and A.K. Seth, (2014). The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. Journal of neuroscience methods, 223, 50-68.

Bezergianni, S. and A. Kalogianni, (2008). Application of principal component analysis for monitoring and disturbance detection of a hydrotreating process. Industrial & Engineering Chemistry Research, 47(18), 6972-6982.

Booth, J. P., Braithwaite, N. St. J., Goodyear, A., & Barroy, P. (2000). Measurements of characteristic transients of planar electrostatic probes in cold plasmas, *Review*

of Scientific Instruments 71, (7), 2722-2727.

Chen, Y. M, Hao, X. L., Zhang, G. Q., & Wang, S. (2006). Flow meter fault isolation in building central chilling systems using wavelet analysis, *Energy Conversion and Management 47,* (13-14), 1700-1710.

Cheremisinoff, N.P., (2000) Handbook of chemical processing equipment. Butterworth-Heinemann.

Chiang, L. H., Russell, E. L., & Braatz, R. D., 2000. Fault detection and diagnosis in industrial systems. Springer Science & Business Media.

Choi, S. W., Martin, E. B., Morris, A. J., Lee, I. B., 2006. Adaptive multivariate statistical process control for monitoring time-varying processes. Ind. Eng. Chem. Res. 45, 3108-3118.

Ding, H., Liu, J. H., & Shen, Z. R. (2003). Drift reduction of gas sensor by wavelet and principle component analysis, *Sensors and Actuators B: Chemical 96,* (1-2), 354-363.

Dunia, R., S.J. Qin, T.F. Edgar, and T.J. McAvoy, (1996). Identification of faulty sensors using principal component analysis. AIChE Journal, 42(10), 2797-2812.

Ferrer, A., (2007). Multivariate statistical process control based on principal component analysis (MSPC-PCA): Some reflections and a case study in an autobody assembly process. Quality Engineering, 19(4), 311-325.

Gaman, C., Keville, B., Zhang, Y., Holohan, A. M., Daniels, S., & Turner, M. M. (2015). Real-time multivariable control of SF_6/O_2/Ar plasmas. In 32nd *International Conference on Phenomena in Ionized Gases*, Iasi, Romania, 155-

156.

Ge, Z., Song, Z., & Gao, F., (2013). Review of recent research on data-based process monitoring. Industrial & Engineering Chemistry Research, 52(10), 3543-3562.

Goodlin, B. E. (2002). *Multivariate endpoint detection of plasma etching processes, Ph. D. Thesis*. MIT Libraries.

Gorak, A. and H. Schoenmakers, (2014). Distillation: Operation and Applications, Academic Press.

Granger, C.W., (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society, 424-438.

Granger, C.W.,(1988). Some recent development in a concept of causality. Journal of econometrics, 39(1), 199-211.

Ha, D., Park, D., Koo, J., Baek, K. H., & Han, C. (2016). Improvement of principal component analysis modeling for plasma etch processes through discrete wavelet transform and automatic variable selection. Computers & Chemical Engineering, 94, 362-369.

Ha, D., Ahmed, U., Lee, C.J., Baek, K. H., & Han, C. (2017). Multi-mode operation of principal component analysis with k-nearest neighbor algorithm to monitor compressors for liquefied natural gas mixed refrigerant processes. Computers & Chemical Engineering, submitted.

He, Q. P., Wang, J., 2010. Large-scale Semiconductor process fault detection using a fast pattern recognition-based method. IEEE Trans. Semicond. Manuf. 23, 194-

200.

Hong, J.J. and J. Zhang. (2010) Progressive PCA modeling for enhanced fault diagnosis in a batch process. In Control Automation and Systems (ICCAS), 2010 International Conference on. IEEE.

Hong, J.J., J. Zhang, and J. Morris, (2011). Fault localization in batch processes through progressive principal component analysis modeling. Industrial & Engineering Chemistry Research, 50(13), 8153-8162.

Hong, J.J., J. Zhang, and J. Morris, (2014). Progressive multi-block modelling for enhanced fault isolation in batch processes. Journal of Process Control, 24(1), 13-26.

Hwang, D. H., Han, C., (1999). Real-time monitoring for a process with multiple operating modes. Control Eng. Pract. 7, 891-902.

Jackson, J.E. and G.S. Mudholkar, (1979). Control procedures for residuals associated with principal component analysis. Technometrics, 21(3), 341-349.

Jin, H. D., Lee, Y. H., Lee, G., Han C., 2006. Robust recursive principal component analysis modeling for adaptive monitoring. Ind. Eng. Chem. Res. 45, 696-703.

Kahng, A. B. (2010). Scaling: More than Moore's law. *IEEE Design & Test of Computers*, 27(3), 86-87.

Kano, M., Hasebe, S., Hashimoto, I., & Ohno, H., 2001. A new multivariate statistical process monitoring method using principal component analysis. Computers & chemical engineering, 25(7), 1103-1113.

Keville, B., Zhang, Y., Gaman, C., Holohan, A. M., Daniels, S., & Turner, M. M.

(2013). Real-time control of electron density in a capacitively coupled plasma. *Journal of Vacuum Science & Technology A*, 31(3), 031302.

Kister, H.Z., (1990). Distillation operations, McGraw-Hill.

Kister, H.Z., (1992). Distillation design. Vol. 223, McGraw-Hill New York.

Klick, M., Rehak, W., & Kammeyer, M. (1997). Plasma diagnostics in RF discharges using nonlinear and resonance effects, *Japanese Journal of Applied Physics Part 1 36*, (7B), 4625-4631.

Kosanovich, K.A., K.S. Dahl, and M.J. Piovoso, (1996). Improved process understanding using multiway principal component analysis. Industrial & engineering chemistry research, 35(1), 138-146.

Kourti, T., (2002). Process analysis and abnormal situation detection: From theory to practice. IEEE Control Systems Magazine, 10-25.

Kresta, J.V., J.F. MacGregor, and T.E. Marlin (1991). Multivariate statistical monitoring of process operating performance. The Canadian Journal of Chemical Engineering, 69(1), 35-47.

Lada, E.K., Lu, J.C., & Wilson, J.R. (2002). A Wavelet-Based Procedure for Process Shift Detection, *IEEE Transactions on Semiconductor Manufacturing 15*, (1) 79-90.

Landells, K. and Z. RaWi, (2012) Abnormal Event Detection Using Principal Component Analysis. U.S. Patent No. 8,121,817B2.

Landman, R. and S. Jämsä-Jounela, (2013). Data-Driven Causal Analysis and its application on a Large-scale Board Machine. 2013, Master's thesis, Aalto

University, Degree programme for Master of Science in Chemical Technology, Espoo, Aalto University, Espoo.

Landman, R., J. Kortela, Q. Sun, and S.-L. Jämsä-Jounela, (2014). Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity. Computers & Chemical Engineering, 71, 446-456.

Lane, S., Martin, E. B., Kooijmans, R., Morris, A. J., (2001). Performance monitoring of a multi-product semi-batch process. J. Process Control 11, 1-11.

Lee, J. M., Yoo, C., & Lee, I. B., 2004. Fault detection of batch processes using multiway kernel principal component analysis. Computers & chemical engineering, 28(9), 1837-1847.

Lee, M. H., Jang, S. H., & Chung, C. W. (2007). Floating probe for electron temperature and ion density measurement applicable to processing plasmas, *Journal of Applied Physics 101*, (3), 033305.

Li, G., S.J. Qin, and T. Yuan, (2016). Data-driven root cause diagnosis of faults in process industries. Chemometrics and Intelligent Laboratory Systems, 159, 1-11.

Li, W., Yue, H. H., Valle-Cervantes, S., & Qin, S. J., (2000). Recursive PCA for adaptive process monitoring. Journal of process control, 10(5), 471-486.

Litvak, H. E. (1996). End point control via optical emission spectroscopy. Journal of Vacuum Science & Technology B, 14(1), 516-520.

Lynn, S. A., MacGearailt, N., & Ringwood, J. V. (2012). Real-time virtual metrology and control for plasma etch, *Journal of Process Control 22*, 666-676.

Ma, B., McLoone, S., Ringwood, J., & MacGearailt, N. (2010). An analysis of noise

on optical emission spectroscopy measurements, In *Signals and System Conference (ISSC 2010)*, Cork, Ireland, 186-191.

MacGregor, J. F., & Kourti, T., (1995). Statistical process control of multivariate processes. Control Engineering Practice, 3(3), 403-414.

Mallat, S. (1989). Multifrequency channel decompositions of images and wavelet models, *IEEE Transactions on Acoustics, Speech and Signal Processing 37*, (12), 2091-2110.

Miletic, I., S. Quinn, M. Dudzic, V. Vaculik, and M. Champagne (2004). An industrial perspective on implementing on-line applications of multivariate statistics. Journal of Process Control, 14(8), 821-836

Mujica, L., J. Rodellar, A. Fernandez, and A. Guemes, (2010). Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. Structural Health Monitoring, 1475921710388972.

Oh, C., Ryoo, H., Lee, H., Kim, S. Y., Yi, H. J., & Hahn, J. W. (2010). Spatially resolvable optical emission spectrometer for analyzing density uniformity of semiconductor process plasma, *Review of Scientific Instruments 81*, (10), 103109.

Park, S., Jeong, S., Jang, Y., Ryu, S., Roh, H. J., & Kim, G. H. (2015). Enhancement of the virtual metrology performance for plasma-assisted oxide etching process by using plasma information (PI) parameters, *IEEE Transactions on Semiconductor Manufacturing 28*, (3), 241-246.

Penha, R. and J.W. Hines, (2001). Using principal component analysis modeling to monitor temperature sensors in a nuclear research reactor. in Proceedings of the

2001 Maintenance and Reliability Conference (MARCON 2001), Knoxville, TN. Citeseer.

Joe Qin, S., (1998). Recursive PLS algorithms for adaptive data modeling. Computers & Chemical Engineering, 22(4), 503-514.

Joe Qin, S., (2003). Statistical process monitoring: basics and beyond. Journal of chemometrics, 17(8-9), 480-502.

Joe Qin, S., (2012). Survey on data-driven industrial process monitoring and diagnosis. Annual Reviews in Control, 36(2), 220-234.

Raich, A. and A. Cinar (1996). Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. AIChE Journal, 42(4), 995-1009.

Ramm, P., Klumpp, A., Weber, J., & Taklo, M. M. (2010). 3D System-on-Chip technologies for More than Moore systems. Microsystem technologies, 16(7), 1051-1055.

Rangan, S., Spanos, C., & Poolla, K. (1997). Modeling and filtering of optical emission spectroscopy data for plasma etching systems, In *Proceedings of the American Control Conference, Albuquerque*, New Mexco, 627-628.

Seth, A.K., (2010). A MATLAB toolbox for Granger causal connectivity analysis. Journal of neuroscience methods, 186(2), 262-273

Sobolewski, M. A. (2006). Real-time, noninvasive monitoring of ion energy and ion current at a wafer surface during plasma etching, *Journal of Vacuum Science and Technology A 24*, (5), 1892-1905.

Tan, S., Wang, F., Peng, J., Chang, Y., & Wang, S., 2011. Multimode process

monitoring based on mode identification. Industrial & Engineering Chemistry Research, 51(1), 374-388.

Tong, H. and C.M. Crowe, (1995) Detection of gross erros in data reconciliation by principal component analysis. AIChE Journal, 41(7), 1712-1722

Valle, S., S.J. Qin, and M.J. Piovoso. (2001). Extracting fault subspaces for fault identification of a polyester film process. In American Control Conference, 2001. Proceedings of the 2001. IEEE.

De Veaux, R.D., L.H. Ungar, and J.M. Vinson (1994). Statistical approaches to fault analysis in multivariate process control. in American Control Conference, 1994. IEEE.

Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N., & Yin, K., 2003. A review of process fault detection and diagnosis: Part III: Process history based methods. Computers & chemical engineering, 27(3), 327-346.

Villegas, T., M.J. Fuente, and M. Rodríguez (2010). Principal component analysis for fault detection and diagnosis. Experience with a pilot plant. In CIMMACS'10 Proceedings of the 9th WSEAS international conference on computational intelligence, man-machine systems and cybernetics.

Waldrop, M. M. (2016). The chips are down for Moore's law, *Nature News 530*, (7589), 145-147.

Wang, D., & Romagnoli, J. A. (2005). Robust multi-scale principal components analysis with applications to process monitoring, *Journal of Process Control 15*, (8), 869-882.

Wang, X., Kruger, U., & Irwin, G. W., 2005. Process monitoring approach using fast moving window PCA. Industrial & Engineering Chemistry Research, 44(15), 5691-5702.

White, D. A, Boning, D., Butler, S. W., & Barna, G. G. (1997). Spatial characterization of wafer state using principal component analysis of optical emission spectra in plasma etch, *IEEE Transactions on Semiconductor Manufacturing 10,* (1), 52-61.

White, D., Goodlin, B., Gower, A., Boning, D., Chen, H., Sawin, H., & Dalton, T. (2000). Low open-area endpoint detection using a PCA-based T^2 statistics and Q statistic on optical emission spectroscopy measurements, *IEEE Transactions on Semiconductor Manufacturing 13,* (2), 193-207.

Xu, X., Xiao, F., & Wang, S. (2008). Enhanced chiller sensor fault detection, diagnosis and estimation using wavelet analysis and principal component analysis methods, *Applied Thermal Engineering 28*, (2-3), 226-237.

Yoo, C. K., Villez, K., Lee, I. B., Rosén, C., Vanrolleghem, P. A., 2007. Multi-model statistical process monitoring and diagnosis of a sequencing batch reactor. Biotechnol. Bioeng. 96, 687-701.

Yoon, S. and J.F. MacGregor, (2003). Principal-component analysis of multiscale data for process monitoring and fault diagnosis. AIChE Journal, 50(11), 2891-2903.

Yuan, T. and S.J. Qin, (2014). Root cause diagnosis of plant-wide oscillations using

Granger causality. Journal of Process Control, 24(2), 450-459.

Yuan, T., G. Li, Z. Zhang, and S.J. Qin. (2016). Deep causal mining for plant-wide oscillations with multilevel Granger causality analysis. In American Control Conference (ACC), 2016. American Automatic Control Council (AACC).

Yue, H. H., Qin, S. J., Markle, R. J., Nauert, C., & Gatto, M. (2000). Fault detection of plasma etchers using optical emission spectra, *IEEE Transactions on Semiconductor Manufacturing 13*, (3), 374-385.

Yue, H. H., Qin, S. J., Wiseman, J., & Toprac, A. (2001). Plasma etching endpoint detection using multiple wavelengths for small open-area wafers. *Journal of Vacuum Science & Technology A*, 19(1), 66-75.

Yue, H.H. and S.J. Qin, (2001). Reconstruction-based fault identification using a combined index. Industrial & engineering chemistry research, 40(20), 4403-4414.

Zhang, Y., & Li, S., 2013. Modeling and monitoring between-mode transition of multimodes processes. IEEE Transactions on Industrial Informatics, 9(4), 2248-2255.

Zhao, S. J., Zhang, J., Xu, Y. M., 2004. Monitoring of processes with multiple operating modes through multiple principle component analysis models. Ind. Eng. Chem. Res. 43, 7025-7035.

Zuppa, M., Distante, C., Persaud, K. C., & Siciliano, P. (2007). Recovery of drifting sensor response by means of DWT analysis, *Sensors and Actuators B: Chemical 120*, (2), 411-416.

Abstract in Korean (요 약)

현대의 산업 공정은 수 많은 단위 공정의 조합으로 이루어진 복잡한 장치 산업이며, 유량, 온도, 압력, 농도 및 조성 등의 수 많은 공정 변수들은 강한 선형적 또는 비선형적인 상관관계를 가지고 있다. 컴퓨터 성능의 향상과 공정 제어 시스템의 발전으로 인하여 공정에서 발생하는 수많은 데이터 및 공정 정보들을 단지 몇 명의 보드 작업자와 현장 작업자들이 관리할 수 있게 되었다. 그렇지만, 한 명의 작업자가 관리해야 하는 공정과 장치가 증가하였기 때문에, 작업자가 전체적인 공정 정보를 인지하지 못한 채로 발생하게 되는 공정 이상 또는 장치 이상이 발생할 수 있는 위험성 또한 존재한다. 이러한 문제를 해결하기 위해 변수 선별법, 잡음 제거, 공정 운전 모드 분류 및 모드 변경 감지를 통해 모니터링 성능을 향상시키는 주성분분석법 모델링 절차를 제안하였다. 그리고 공정 이상 진단 및 인과관계 분석 기법을 도입하였다. 공정 변수 간의 인과관계 행렬을 계산하였으며 이를 통해 예상치 못한 공정 상태의 변화에 대한 원인을 분석하였다. 제안된 방법론은 인천에 있는 액화천연가스 공정과 플라즈마 식각 장치의 플라즈마 상태 모니터링을 대상으로 적용, 검증하였다.

Chapter 2에서는 광진단계가 가지고 있는 잡음을 제거하기 위한 신호처리 기법의 적용과 모니터링 민감도를 향상시키는 다변량 통계 기법을

제시하였다. 플라즈마 센서 중 광진단계는 널리 사용되고 있으며 광진단계의 고차원성은 주성분분석법과 같은 다변량 통계기법을 필요로 한다. 그렇지만 주성분분석법은 통계적인 계산 과정에서 대상 공정이 가지고 있는 물리적인 의미를 평가 절하할 수 있으며 광진단계의 전하 결합 소자(CCD)의 고유 잡음은 주성분모델의 성능을 저하시킬 수 있다. 그러므로 물리적으로 중요한 변수를 미리 선별하고 광진단계 기반 플라즈마 데이터를 모델링학기 전에 잡음 신호를 필터링 하는 것이 필요하다. 이를 위하여 이번 chapter에서는 플라즈마에서 물리적으로 의미 있는 파장을 선별하는 기법과 CCD array에서 잡음 신호를 필터링하기 위한 이산 웨이블릿 변환 기법을 제시하였다. 기존의 conventional PCA 모델을 적용하였을 때에는 테스트 한 모든 공정 이상 상황에 대해 감지하지 못하였지만 이번 chapter에서 도입한 PCA 모델에서는 0.67%의 소스 전력 변화와 같은 극히 적은 공정의 변화도 감지할 수 있었다.

Chapter 3에서는 액화천연가스 혼합 냉매 공정의 성능을 향상시키고 공정 가동 정지를 방지하기 위한 운전 모드 식별 및 다중 모드 PCA 방법론의 적용 방법을 제시하였다. 액화천연가스 혼합 냉매 공정은 주로 천연가스를 액화하기 위한 공정이다. 냉매 압축용 압축기는 고속 회전 부품과 결합되어 고압 조건에서 운전된다. 그러나 압축기의 오작동으로 인

해 공정 가동 중단이 발생하고, 장비에 치명적인 손상을 야기시키며 안전에 관한 문제가 발생할 수 있다. 기존의 방법론에서는 공정이 단일 운전 모드에서 운전된다고 가정하므로 공정의 오작동과 운전 모드의 변경을 구별하기가 쉽지 않다. 따라서 k−nearest neighbor algorithm (k−NN)를 이용하여 운전 모드를 분류하고 이를 다중 운전 모드용 주성분 분석법 (MPCA)로 공정 모니터링 및 이상 탐지를 수행하였다. 제안된 방법의 성능 검증을 위하여 실제 LNG 플랜트 운전 데이터를 활용하였으며 검증 결과 기존의 PCA 방법론에 비해 보다 정확하고 빠른 이상 감지 능력을 가짐을 확인할 수 있었다.

Chapter 4에서는 PCA 기반의 이상 증폭 알고리즘을 사용하여 시스템에서의 이상의 원인과 이상 전파 경로를 모두 탐지할 수 있는 방법론을 제시하였다. 개발된 알고리즘은 공정 샘플을 잔여 부분 공간 (RS)에 투사하여 외란 전파 경로를 결정한다. 일반적으로 공정 이상 데이터의 RS는 이상의 크기를 증폭하기 위해 최소화되어야 하는 정상적인 공정 내에서의 변화와 겹쳐진다. 증폭된 이상을 포함하는 RS는 공분산 행렬로 변환되고, 그 다음에는 최대 고유치에 대응하는 이상 방향 행렬을 차례로 생성하는 특이 값 분해(SVD) 분석이 수행된다. 이상이 발생한 변수는 절대 내림차순 함수를 사용하여 이상 전파 경로를 나타내는 이상에 대한 기여도에 따라 다시 배열된다. 또한 다변량 Granger Causality 알고리즘

을 사용하여 개발된 알고리즘에서 얻은 변수 간의 인과 관계를 분석할 수 있다. 두 가지 방법론은 LNG Fractionation 공정의 distillation column 운전 케이스에 대해 검증하였으며 이상 시나리오는 이상 방향을 계산한다고 가정하였다. 이상 전파 경로 알고리즘에서 얻어진 변수들의 계층 구조는 MVGC 알고리즘과 잘 일치함을 알 수 있다. 그러므로 이상 증폭 방법론은 이상 전파 경로뿐만 아니라 이상의 근본 원인을 확인하기 위하여 사용될 수 있다.

위의 적용 결과를 토대로 제시된 다변량 통계기법 기반 분석 방법은 분산 제어 시스템이 적용된 다양한 공정에서 공정 모니터링 및 이상 진단을 위한 유용한 정보를 제공함으로써 생산성 향상, 안전성 향상 등을 이룰 수 있음을 알 수 있다.

주요어: 공정 모니터링, 이상진단, 운전 모드 식별, 그레인저 인과관계, 주성분분석법, k−NN 분류 기법

학번: 2013−30990

성명: 하 대 근