IMPLEMENTING PCA BASED ON FAULT DETECTION SYSTEM BASED ON SELECTED IMPORTANT VARIABLES FOR CONTINUOUS PROCESS

MOHD HUZAIFAH BIH HAMZAH

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Chemical Engineering

Faculty of Chemical and Natural Resources Engineering
UNIVERSITI MALAYSIA PAHANG

FEBRUARY 2013

IMPLEMENTING PCA BASED ON FAULT DETECTION SYSTEM BASED ON SELECTED IMPORTANT VARIABLES FOR CONTINUOUS PROCESS

ABSTRACT

Multivariate Statistical Process Control (MSPC) is known generally as an upgraded technique, from which, it was emerged as a result of reformation in conventional Statistical Process Control (SPC) method where MSPC technique has been widely used for fault detection and diagnosis. Currently, contribution plots are used in MSPC method as basic tools for fault diagnosis. This plot does not exactly diagnose the fault but it just provides greater insight into possible causes and thereby narrow down the search. Therefore, this research is conducted to introduce a new approach and method for detecting and diagnosing fault via correlation technique. The correlation coefficient is determined using multivariate analysis techniques that could use less number of newly formed variables to represent the original data variations without losing significant information, namely Principal Component Analysis (PCA). In order to solve these problems, the objective of this research is to develop new approaches, which can improve the performance of the present conventional MSPC methods. The new approaches have been developed, the Outline Analysis Approach for examining the distribution of Principal Component Analysis (PCA) score. The result from the conventional method and ne approach were compared based on their accuracy and sensitivity. Based on the results of the study, the new approaches generally performed better compared to the conventional approaches.

.

MELAKSANAKAN PCA BERDASARKAN SISTEM PENGESANAN KEROSAKAN BERDASARKAN PEMBOLEHUBAH TERPILIH PENTING UNTUK PROSES BERTERUSAN

ABSTRAK

Multivariat Kawalan Proses Statistik (MSPC) dikenali secara amnya sebagai teknik yang dinaik taraf, dimana ia telah muncul sebagai hasil reformasi dalam Kawalan Proses konvensional Statistik (SPC) kaedah di mana teknik MSPC telah digunakan secara meluas untuk mengesan kerosakan dan diagnosis. Pada masa ini, plot sumbangan digunakan dalam kaedah MSPC sebagai alat asas untuk diagnosis fault. Plot ini tidak tepat mendiagnosis kerosakan tetapi ia hanya memberikan gambaran yang lebih besar ke dalam sebab yang mungkin dan sekali gus mengecilkan carian. Oleh itu, kajian ini dijalankan untuk memperkenalkan pendekatan baru dan kaedah untuk mengesan dan mendiagnosis kerosakan melalui teknik korelasi. Pekali korelasi ditentukan menggunakan teknik analisis multivariat yang boleh menggunakan nombor yang kurang daripada pemboleh ubah yang baru ditubuhkan untuk mewakili variasi data asal tanpa kehilangan maklumat penting, iaitu Analisis Komponen Utama (PCA). Dalam usaha untuk menyelesaikan masalah ini, objektif kajian ini adalah untuk membangunkan pendekatan baru, yang boleh meningkatkan prestasi kaedah konvensional yang hadir MSPC. Pendekatan baru telah dibangunkan, Pendekatan Rangka Analisis untuk memeriksa pengagihan Analisis komponen prinsipal (PCA) Perincian. Hasil dari kaedah konvensional dan pendekatan dibandingkan berdasarkan ketepatan dan sensitiviti mereka. Berdasarkan keputusan kajian, pendekatan baru yang dilakukan umumnya lebih baik berbanding dengan pendekatan konvensional.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

α	Desired confidence percentages
$\mathbf{A}$	Square matrix
j	Variables
k	Principal component
l	Number of data
m	NOC data size
n	Number of retained principal component
p	Number of variables
$\mathbf{R}$	Correlation matrix
$\mathbf{S}$	Variance-Covariance matrix
S	NOC standard deviation
δ	Shifted value in standard deviation unit
T	Temperature of the System
τ	Time constant
τ_{D}	Derivative time constant
τ_{I}	Integral time constant
θ	Dead time
$\mathbf{V}$	Eigenvector matrix
$\mathbf{v}$	Eigenvector or loading vectors
$\bar{x}$	NOC mean
$\bar{x}_1$	Shifted mean
x_i	i th observation in the process
$\boldsymbol{X}$	Data matrix
$\boldsymbol{X}_{\boldsymbol{NOC}}$	Data matrix NOC

LIST OF ABBREVIATIONS

CSTR	Continuous Stirred Tank Reactor
CUSUM	Cumulative Sum Control Chart
EWMA	Exponential Weight Moving Average
MSPC	Multivariate Statistical Process Control
NIPALS	Non-iterative Partial Least Squares
NOC	Normal Operation Condition
OA	Outline Analysis
OC	Out of Control
PCA	Principal Component Analysis
PFDD	Process Fault Detection and Analysis
SPC	Statistical Process Control
SPE	Squared Prediction Errors
UCL	Upper Control Limit

CHAPTER 1

INTRODUCTION

Chemical process systems are highly sensitive to abnormal changes in operating condition. So that, to attain the maximum possible yield in chemical process, it necessary to ensure that the process is maintained around the desired limit. As a direct consequence, the accuracy and the sensitivity of the process monitoring tool is very important. The Multivariate Statistical Process Control (MSPC) method has been applied because it provided a wide range of tools to perform process monitoring and also very effective at extracting hidden information in problems with multiple correlated variables (Louwerse and Smilde, 2000). This research demonstrates the application of the MSPC method to provide a monitoring tool, which is capable of detecting and diagnosis the process fault.

1.1 Research Background of the Project

Most of the Statistical Process Control (SPC) techniques involve operations on single response variables such as weight, pH, temperature, specific gravity, concentration and pressure. This is nature because one is usually interested in a problem involving a single response. Normally, the fault in the process is sought through the usage of the SPC control chart, but in practice most of the SPC control charts are based on charting only a small number of variables, usually the final product of quality variables. These approaches are often inadequate for modern and complex process industries. For this reason, a multivariate approach is applied in the SPC realm to detect the fault condition in the large number of variables observations.

There are however, a number of occasions when more than one response variables (multivariate) are of importance to a problem, and these variables should be studied collectively in order to take advantage of the information about the relationship among the data. With the advance of process sensors and data acquisition systems, today chemical processes are becoming better instrumented.

In many cases, this instrumentation provides an abundance of data, some of which can be classified as redundant for example, the measurements are highly correlated. Multivariate method such as Principal Component Analysis (PCA) can express the essential information contained in these measurements in term of relatively small dimension of new variables without losing the previous information. By applying the MSPC, this new strategy of monitoring fault and diagnosis process operating condition can predict process degradation and equipment failure, thus it can improve the chemical plant production process using the diagnosis through this method.

Fault detection and the monitoring of process performance is an integral part for a successful operation. The MSPC chart can be used to monitor the performance of any given process. The main function of this control chart is to compare the current state of the process against the Normal Operating Condition (NOC). The 'NOC' condition exists when the process or product variables remain close to the desired values. In contrast, the Out of Control (OC) occurs when fault appears in the process. The fault of malfunction is designated when the process departs from an acceptable range of observed variables.

1.2 Problem Statement

The present conventional MSPC has several weaknesses in process fault detection and diagnosis. Some researchers in this field had commented that the MSPC is a powerful tool for data complexity reduction and fault detection in the significant fault appearance data. According to Manabu and his research partner, (2000), the current fault detection and diagnosis method via MSPC is limited to significant faults and does not point put the insignificant ones accurately. Qin (2001) also commented that the contribution chart does not have a control limit, making it difficult to determine the root cause of the abnormal operating condition.

As a summary of summary other researchers, the weakness of the conventional MSPC can be briefly concluded into three disadvantages. First of all, the complicated control charts are not “user friendly”, secondly the conventional MSPC fault detection tools are easily rise up to noisy fault signals and lastly the conventional fault diagnosis is not ready with a proper control limit, thus it cannot determine the root cause of the fault especially multiple faults. In order to improve the limitation of MSPC, this research should focus on the alternative, which can solve the disadvantages mentioned above.

1.3 Research Objective.

The main purpose of this research is to study the impact of applying various modes of normal operating condition (NOC) in terms of the number of samples and variable variations on the process monitoring performance for continuous-based process. Therefore, the main objectives of this research are:

i. To develop the conventional MSPM method based on a single NOC
ii. To implement the conventional MSPM method based on different modes of NOC.

1.4 Scope of research.

Scope of propose study are on the development of PCA-based fault detection system based on various modes of NOC models for continuous-based process. There are three main scope will be investigated using MATLAB.

i. The conventional MSPM method will be develop based on single NOC mode. The linear PCA algorithm is used for reducing the multivariate data dimensions.
ii. The MSPM will be run traditionally by implementing different mode, which in this research is on two modes. According to Zhao et al. (2004),in spite of the success of applying PCA based MSPM tools to process data for detecting abnormal situations, when these tools are applied to a process with multiple operating modes, many missing and false alarms appear even when the process itself under other steady-state nominal operating conditions.
iii. As all data have been obtained, it will be analyze further with two multivariate control charts namely Hotelling's T^2 and Squared Prediction Errors (SPE) statistic for the fault detection operation.

1.5 Rational and Significance

In this research, effort mainly concentrates on breaking through the current limitation and the further application of MSPC on a multivariate continuous chemical process. The main contributions of this research are:

i. Application of MSPC tools on the fault detection and diagnosis.
ii. An Eigenvalue-eigenvector PCA approach had been used for developing Principals Components model.

- Modified Process Fault Detection and Diagnosis, mechanisms are also developed based on the Outline Analysis.

1.6 Research Contribution.

In this research, effort mainly concentrates on breaking through the current limitation and the further application of MSPC on a multivariable continuous chemical process. The main contributions of this research are:

i. Application of MSPC tools on the fault detection and diagnosis for reactor in a chemical plant.
ii. An Eigenvalue-eigenvector PCA approach had been used for developing Principals Components model instead of the conventional algorithm.
iii. Modified Process Fault Detection and Diagnosis, PFDD mechanisms are also developed based on the Outline Analysis.

1.7 Organization of The Thesis

This thesis contains five chapters: introduction, literature review, research methodology, result analysis and discussion, and conclusion as recommendations. The first chapter of this thesis mainly present about the introduction of the research projects, which consists of the research background, problem statement, research objectives and scopes.

Second chapter, covers the literature review. This chapter presents the development of Process Fault Detection and Diagnosis and MSPC methods.

In the following chapter, the methodology for the research project will be proposed. The propose methodologies are described and present step by step.

Chapter four mainly focuses on results analysis and discussion. The suggested fault detection and diagnosis results are presented and compared to the results obtained by means of conventional approach.

Finally this thesis wrap up with the conclusion and recommendations for future researches.

CHAPTER 2

LITERATURE REVIEW

This chapter focuses on two main literature parts that had been reviewed. Firstly, this chapter introduces the important concept and application of Process Fault Detection Diagnosis (PFDD). Besides, it also addresses the development of PFDD research progress. In succession, the second part presents the overview of process monitoring via statistical approach such as some key elements in MSPC method and progress in MSPC research.

2.1 Process Fault Detection and Diagnosis

As the central component of abnormal event management PFDD involves interpreting the current condition or status of the plant by giving sensor reading and process knowledge. Early detection and diagnosis of the process faults, the plant is still operating in controllable condition can help avoid undesired event progression and reduce the amount of productivity loss. This can be achieve by timely detection of an abnormal event or fault, diagnosing its root of the fault and then taking appropriate supervisory control decisions and actions to bring the process back to the normal, safe and acceptable operating state (Venka et al, 2002). Thus, PFDD becomes an important an important aspect of operating a process plant. Not only it is important form the point of quality but also for the maintenance and safety viewpoints, where providing human operators the assistance in this most pressing area of need.

To improve the performance of the fault detection and diagnosis, it’s a challenge for the research in this area to understand the fault, which came from the failure and malfunction of the process. The definition of fault depends on the characteristics of the process chosen for measurement or calculation, their acceptable range and the accuracy of the statistic used for classification of a potential fault.(David and David, 1994). In practice, the word fault can be categorized into three groups (Venka et al, 2002), they are:

i. Equipment failures and degradation.
ii. Process structure change.
iii. Parameter drifts.

The details of the description of the fault will be discuss later on chapter III, which will explain the classes of fault, and the generation o faults in data simulations. A lot of PFDD systems have been developed. Each method is used to detect and diagnosis fault in different approaches. The common characteristics of this method are presented in section 2.1.1.

2.1.1 Characteristics of Process Fault Detection and Diagnosis System (PFDD).

PFDD normally works together with a process control system. Process control is a system, which maintains desired conditions in a physical system by adjusting selected manipulated variables in the process. Figure 2.1 is the schematic diagram of difference between feedback control system and PFDD system.

Obvious, process control deals with the error signal that is measured by measuring device. Control gives order to final control elements, which manipulate the input parameter and makes sure the process always maintain at desired state.

Comparably, traditional process control system given current process condition where it do not take interest to state variables and parameters estimation, resulting a situation whereby it can neither detect faults nor carry out where the fault come from.

In order to detect and diagnose fault in the process, PFDD system have to implement some additional measuring hardware and work accomplish with fault analytical method by using historical data from the process records respectively. Figure 2.1 also show that the process monitoring, fault detection and fault diagnosis are strongly linked.

In the view of plant safety and process performance, any abnormal condition in process operation should be detected at an early stage and the cause should be figured out. These steps are important and the diagnosis of the faults should be done while faults are still minimal to be neglected that mean not harmful to the process and the process also can still be covered and maintained.

The following are desirable characteristics that a PFDD system must ne fulfilled to become more effective. (Dash and Venkat, 2000)

i. Early detection and diagnosis.

   Further analysis and action can be taken after accurate detection and diagnosis are signaled.

ii. Insolubility.

   Insolubility refers to the ability of the system to discriminate between different failures.

iii. Insignificant fault identifiability.

   The system should be able to recognize occurrence of insignificant fault and not misclassify them as normal operation.

iv. Multiple fault identifiability.

   The PFDD system should be able to detect and diagnose multiple faults.

v. Explanation facility.

   PFDD system should able to identify the source of the fault and explain on reason about the cause and effect relationship in a process.

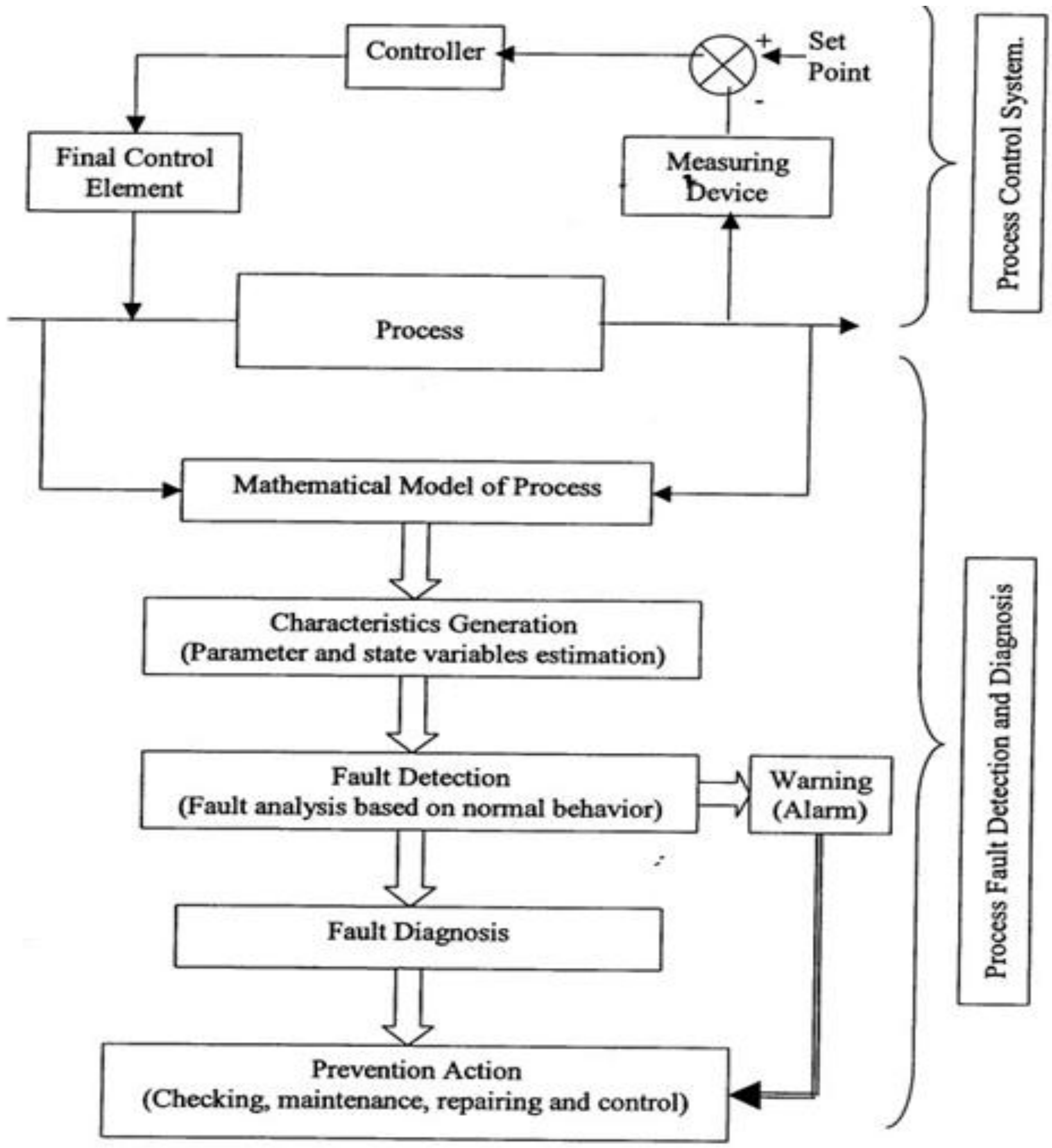

Figure 2.1: Schematic Diagram of Difference between Feedback Control System and PFDD system.

2.1.2 Classification of PFDD System.

A lot of PFDD systems have been developed and a short exposition on the various PFDD method are now present. All of these PFDD systems have their own strategy, methods, assumption, theoretical background and objective. These PFDD systems are often applied in different process operations and handled by different operators. Due to the above reason, variation in PFDD is a good philosophy to design those PFDD systems as tools that fit into the way people work when they perform a fault detection and diagnosis, rather than trying to fit man power to the system. Dash and Vankat (2000) had summarized a detailed classification of PFDD, which is shown in Figure 2.2. The classification is categorized in term of quantitative or qualitative methods and process model based or process historical based.

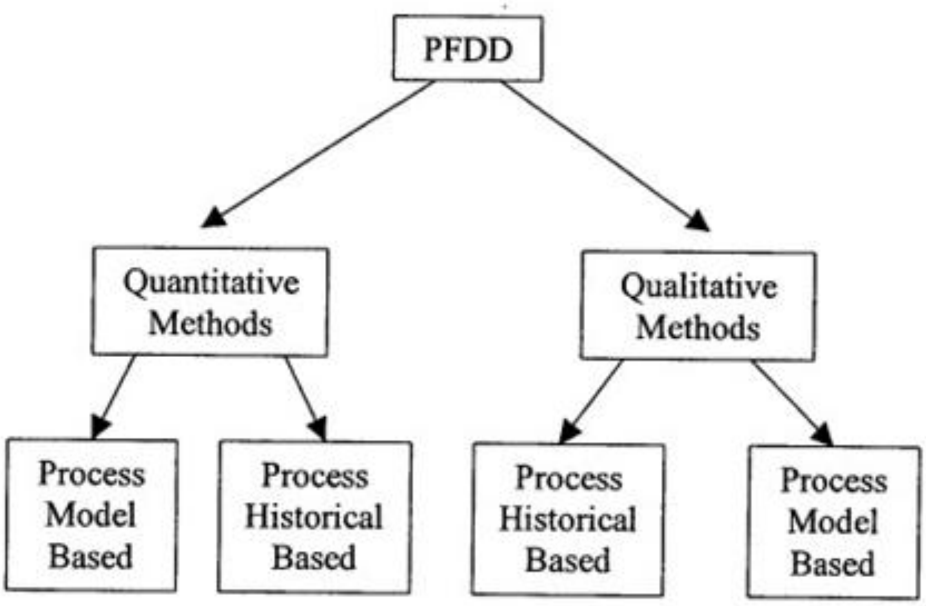

Figure 2.2 Classification of PFDD approach.