CONTROL CHART PATTERNS RECOGNITION WITH CONSTRAINED DATA

RAZIEH HAGHIGHATI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Mechanical Engineering)

School of Mechanical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

APRIL 2019

DEDICATION

Dedicated to

To my beloved family

ACKNOWLEDGEMENT

All Praise to ALLAH S.W.T the Almighty, for giving me the blessing, the strength, the chance and endurance to complete this research.

I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Adnan Bin Hassan, for his generous time, invaluable guidance, enthusiastic encouragement and insightful comments throughout the journey. He taught me innumerable lessons on the academic research as well as life and patiently guided me towards my goals.

I would like to convey my appreciation to all the staff in the School of Mechanical Engineering, for kind support and help on the technical and the administrative aspect of the study. I would like to acknowledge the generous funding of this work provided by the Malaysia International Scholarship, Ministry of Higher Education, Malaysia (KPT.B.600-18/3JLD.5 (24)).

I am grateful to my amazing family members for the love, support, and consistent encouragement I have gotten over the years. I could not have finished this study without full support of my beloved father, Mr. Manouchehr. My special thanks go to my beloved mother, Ms. Nahid for her endless love, patience and prayers. I undoubtedly could not have done this without her sacrifice. And thanks to my brother, Mahdi who bore extra responsibilities during my absence and provided moral support.

Finally, I wish to give my heartfelt thanks to my husband, dr. Abdolatif, whose unconditional kindness, encouragements, prayers and academic supports enabled me to reach this milestone.

ABSTRACT

Recognition and classification of non-random patterns of manufacturing process data can provide clues to the possible causes that contributed to the product defects. Early detection of abnormal process patterns, particularly in highly precise and rapid automated manufacturing is necessary to avoid wastage and catastrophic failures. Towards this end, various control chart patterns recognition (CCPR) methods have been proposed by researchers. Most of the existing control chart patterns recognizers assumed that data is fully available and complete. However, in reality, process data streams may be constrained due to missing, imbalanced or inadequate data acquisition and measurement problems, erroneous entries and technical failure during data acquisition process. The aim of this study is to investigate and develop an effective recognition scheme capable of handling constrained control chart patterns. Various scenarios of data constraints involving missing rates, missing mechanisms, dataset size and imbalance rate were investigated. The proposed scheme comprises the following key components: (i) characterization of input data stream, (ii) imputation and feature extraction, and (iii) alternative recognition schemes. The proposed scheme was developed and tested to recognize the constrained patterns, namely, random, increasing/decreasing trend, upward/downward shift and cyclic patterns. The effect of design parameters on the recognition performance was examined. The Exponentially-Weighted Moving Average (EWMA) imputation, oversampling and Fuzzy Information Decomposition (FID) were investigated. This research revealed that some constraints in the dataset can eventually change the distribution and violate the normality assumption. The performance of alternative designs was compared by mean square error, percentage of correct recognition, confusion matrix, average run length (ARL), t-test, sensitivity, specificity and G-mean. The results demonstrated that the scheme with an ANN-fuzzy recognizer trained using FID-treated constrained patterns significantly reduce false alarms and has better discriminative ability. The proposed scheme was verified and validated through comparative studies with published works. This research can be further extended by investigating an adaptive fuzzy router to assign incoming input data stream to an appropriate scheme that matches complexity in the constrained data streams, amongst others.

ABSTRAK

Pengecaman dan klasifikasi terhadap corak data proses pembuatan yang tidak rawak boleh memberi petunjuk terhadap faktor yang mungkin mengakibatkan kecacatan pada produk. Pengesanan awal terhadap corak proses tidak normal terutamanya bagi pembuatan berautomatik cepat dan persis tinggi adalah perlu bagi mengelakkan pembaziran dan kegagalan bencana. Oleh itu, pelbagai kaedah pengecaman corak carta kawalan (CCPR) telah dicadangkan oleh penyelidik. Kebanyakan pengecam corak carta kawalan sedia ada mengandaikan bahawa data adalah tersedia dan lengkap sepenuhnya. Walau bagaimanapun realitinya, aliran data proses mungkin terhalang disebabkan oleh kehilangan, tidak seimbang atau tidak mencukupi disebabkan masalah perolehan dan pengukuran data, kesilapan memasukkan data dan kegagalan teknikal semasa proses perolehan data. Tujuan kajian ini adalah untuk menyiasat lanjut dan membangunkan satu skema pengecaman corak carta kawalan terkekang yang efektif. Pelbagai senario kekangan data yang melibatkan kadar kehilangan, mekanisme kehilangan, saiz set data dan kadar ketidakseimbangan telah dikaji. Skema yang dicadangkan terdiri daripada komponen-komponen utama yang berikut: (i) pencirian aliran data masukan, (ii) imputasi dan pengekstrakan ciri, dan (iii) skema pengecaman alternatif. Skema yang dicadangkan telah dibangunkan dan diuji untuk mengenal corak yang terkekang, iaitu rawak, kecenderongan meningkat/menurun, peralihan ke atas/bawah dan corak kitaran. Kesan rekabentuk terhadap prestasi pengecaman telah diperiksa. Kaedah imputasi purata pergerakan berfungsi berpotensi (EWMA), persampelan lebih dan Penguraian Maklumat Kabur (FID) disiasat. Kajian ini mendedahkan bahawa beberapa kekangan dalam dataset boleh akhirnya telah mengubah taburan dan menyalahi andaian normal. Prestasi rekabentuk alternatif dibandingkan dengan mean square error, peratusan pengecaman tepat, matriks kekeliruan, purata panjang larian (ARL), ujian-t, kepekaan, kekhususan dan min-G. Keputusan menunjukkan bahawa skema dengan pengecam ANN-kabur yang dilatih dengan corak terkekang yang dipulihkan oleh FID berjaya mengurangkan penggera palsu dan mempunyai keupayaan diskriminatif yang lebih baik. Skema yang dicadangkan telah disahkan melalui kajian perbandingan dengan kajian-kajian yang telah diterbitkan. Penyelidikan ini boleh diperluaskan lagi antaranya ialah dengan menyiasat penghala kabur adaptif untuk menetapkan laluan aliran data masukan yang sepadan dengan kerumitan data terkekang.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANOVA	-	Analysis of Variance
ANN	-	Artificial Neural Network
ARL	-	Average Run Length
CCP	-	Control Chart Pattern
CCPR	-	Control Chart Pattern Recognition
CYC	-	Cyclic Pattern
DS	-	Dataset Size
DT	-	Decreasing Trend Pattern
EWMA	-	Exponentially Weighted Moving Average
FID	-	Fuzzy Information Decomposition
G-MEAN	-	Geometric Mean
IID	-	Independent and Identically Distributed
IR	-	Imbalance Rate
IT	-	Increasing Trend Pattern
MAR	-	Missing at Random
MCAR	-	Missing Completely at Random
MF	-	Membership Function
ML	-	Machine Learning
MLP	-	Multilayer Perceptron
MM	-	Missing Mechanism
MR	-	Missing Rate
MSV	-	Mean Square Error
NOR	-	Normal Pattern
NMAR	-	Not Missing at Random
SPC	-	Statistical Process Monitoring
SVM	-	Support Vector Machine
US	-	Upward Shift Pattern

LIST OF SYMBOLS

b	-	Baseline noise level
c	-	Amplitude of cyclic pattern
C	-	Size of observation window
h	-	Magnitude of shift pattern
H	-	Production horizon finite length
s	-	Gradient of trend pattern
t	-	Time of sampling
T	-	Period of cyclic pattern
Y	-	Pattern variable
α	-	Smoothing factor
μ	-	Mean
σ	-	Standard Deviation
P_0	-	A series of subgroup averages sampled
P_t	-	Unstable pattern
X_k	-	Sample of n observations
$\overline{x_i}$	-	Subgroup average
$\underline{x_i}$	-	Missing at least one observation

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of the Problem

In today's global competitive marketplace, manufacturers strive to deliver the highest quality products at the lower production cost to gain an edge in the competition. With an appropriate quality control system, companies can maintain and continually improve the quality of products and processes and reduce the production cost. For the reason, manufacturing processes should be monitored effectively to avoid deviations in the production process and benefit from consistent quality.

Statistical process control (SPC) techniques are widely used in production industries to monitor processes and improve quality. One of the most important SPC tools is control chart which is used to differentiate between common cause and special cause of process variation. Measured quality characteristics are plotted in the control charts. A control charting procedure consists of getting a sample of n measures of a quality characteristic of a product at fixed time intervals and plotting on a graph a statistic computed through the n measures vs. a control interval delimited by two control limits. Control charts for long run processes were originally proposed by Shewhart while working for Bell Laboratories in 1920s. Later, exponentially-weighted moving average (EWMA) control charts were respectively proposed to improve the detection capabilities with respect to small to moderate process shifts to the out-of-control condition. While other control charts treat rational subgroups of samples individually, the EWMA chart tracks the exponentially-weighted moving average of all prior sample means. EWMA weights samples in geometrically decreasing order so that the most recent samples are weighted most highly while the most distant samples contribute very little. If any point falls outside the control limits or a non-random pattern is seen in the data located within the control limits, the chart display an unstable process (Montgomery, 2007). Identifying the points which are located beyond control limits are not as challenging as recognition of control chart patterns (CCPs). Because CCPs are distorted by random noise in the process which is due to common cause of variation in manufacturing (Gauri & Chakraborty, 2008). Recognition of non-random patterns of process data can provide clues to the possible causes that contributed to the process unstability. Early detection of process variation and assignable cause leads to improvement in the process yield and associated time and cost. The challenge is that in the ever-increasing rapid manufacturing, production time is too short to collect sufficient and balanced data. Figure 1.1 shows the trends leading to the problem.

Moreover, data acquisition problems, erroneous entries, inapplicable measurement, disclosure restrictions and deliberate withholding of some information makes the situation more complex when a proportion of the limited available data also might be partially lost, and all together lead to severe constrained data (Abdelrahman Senoussi *et al*., 2017). Recognition of control chart patterns with missing, inadequate and imbalanced data is a challenging task.

The area of pattern recognition with constrained data is recognized as one of the 100 key scientific research fronts and the first one in terms of citations in mathematics, computer science and engineering in 2013 (King & Pendlebury, 2013). The core researches in this area consisted of methods and algorithms designed for the recovery or restoration of signals, images, and videos with sparse data source or where noise or blur must be corrected, or missing data filled in.

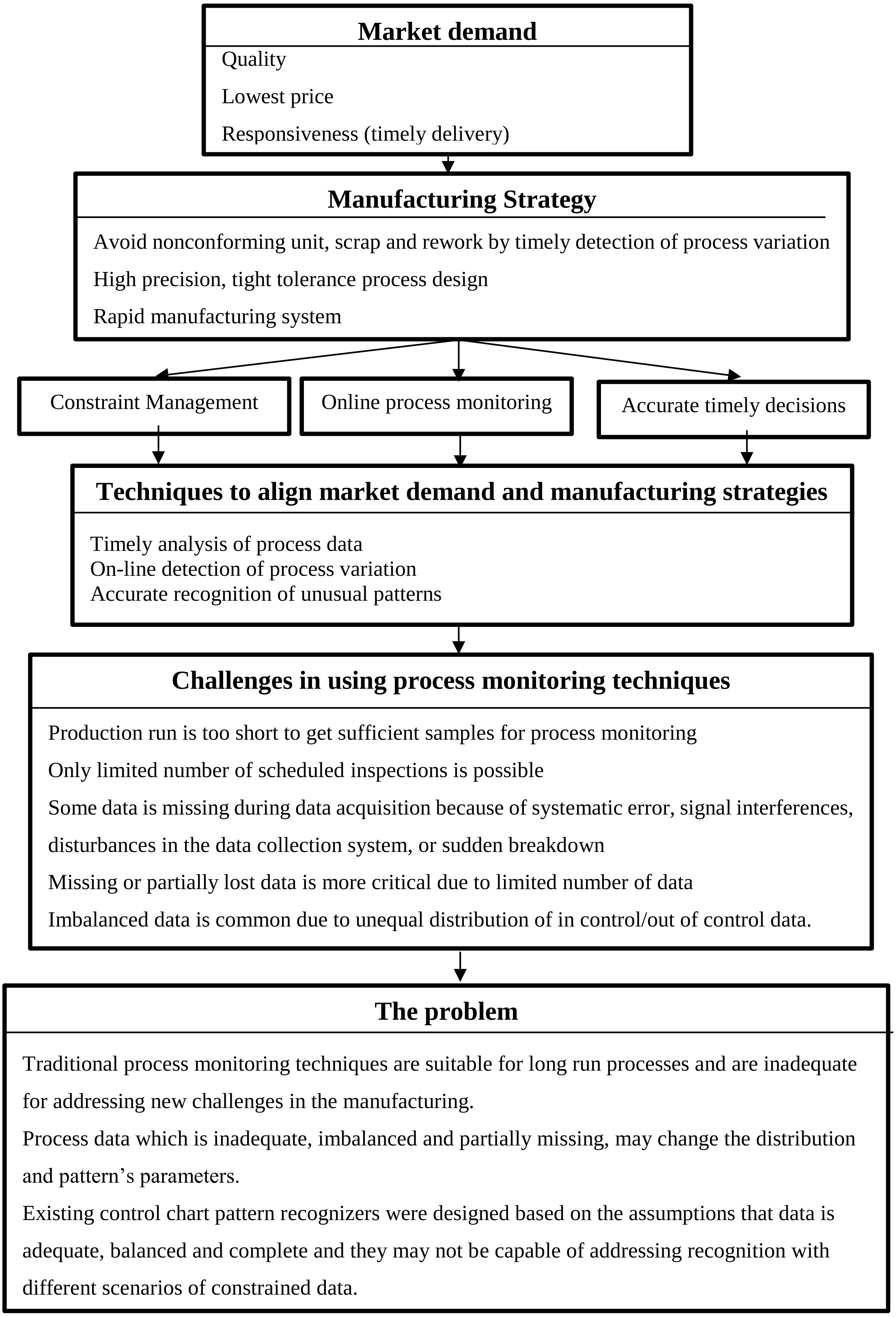

Figure 1.1 Trends leading to problem

In recent years, constrained data has received a growing attention in SPC where researchers studied construction of control charts with missing data (Madbuly *et al*., 2013; Mahmoud *et al*., 2014; Wilson, 2009), recognition of CCPs with imbalanced data (Xanthopoulos & Razzaghi, 2014), recognition of bivariate CCPs with missing data (Abdelrahman Senoussi *et al*., 2017) and a conceptual study on recognition of CCPs with constrained data (Haghighati & Hassan, 2014). Even though recognition of constrained data has been investigated in various fields of studies, our literature review was unable to find rigorous investigation on recognition of constrained control charts patterns.

1.2 Statement of the Problem

Most of the existing control chart pattern recognizers were designed based on the assumption that training data is sufficient, balanced and complete. However, in actual production processes, data may be inadequate, imbalanced and missing due to data acquisition problems, erroneous entries, inapplicable measurement, disclosure restrictions and deliberate withholding of some information. Even though constrained data issues have been investigated in various field of studies, but no rigorous study has been found on CCPR. Existing frameworks are unable to address the issues that arose with reduction of dataset size which diminishes the effectiveness of current input representation, training and recognition algorithms. To address this gap, improved recognition schemes are needed to improve recognition of constrained control chart patterns. So that, there is a need to provide improved procedures for input representation, data treatment and training of the recognizers, according to different scenarios of missing data. An effective solution should give fast and accurate recognition of process variability patterns with minimum false alarms, even when the available process data is constrained**.**

1.3 Purpose of the Research

This research aims to develop alternative recognition schemes for process data streams within several constrained data scenarios, namely missing, inadequate and imbalanced. In particular, improved procedures should be investigated with desirable performance of accurate prediction and recognition of constrained data streams within control limits of shewhart X-bar chart.

1.4 Objectives of the Study

The research objectives are listed as below:

1. To formulate an effective imputation algorithm that can address missing data in control chart patterns.
2. To develop recognition schemes to address recognizer training with inadequate, imbalanced and missing data.
3. To assess characteristics of constrained data scenarios in alternative recognition schemes.

1.5 Scope and Key Assumptions

The scope of study covers the followings:

1. Manufacturing process data were assumed to be univariate and X-bar Shewhart control chart was used to investigate process variation.
2. Standard synthesized and published reference manufacturing data was used for modelling and validation.
3. Constrained data is limited to three types of constraints, namely, missing data, inadequate data and imbalanced data.

4. Solution approach for pattern recognition is limited to soft computing techniques.

1.6 Significance of Research

This study is significant and important from theoretical and practical aspects. This research is rationalized and motivated from the following perspectives:

From theoretical perspective, developing improved recognition schemes for constrained data, can address shortcoming of existing frameworks. Existing recognition frameworks are lacking in addressing challenging issues in training and recognition of data streams that are missing, inadequate and imbalanced. Therefore, it is important to design and develop improved schemes with relevant input representation and training procedures for different scenarios of constrained data.

From practical perspective, in the era of rapid manufacturing where the production time is too short to collect sufficient data, effective techniques are required to facilitate accurate process monitoring and recognition, in early phases of process deviation and deterioration. Improved procedures and techniques should be able to recognize and predict process unstability when sufficient representative process data is not available due to various reasons.

1.7 Operational Framework for Research Activities

The recognition schemes were developed and tested in four phases as it is shown in Figure 1.2. First, a method was designed for simulation and characterization of datasets with missing data. The missing data factors that can change the distribution and normality of data were identified in this step. In the second phase, various imputation techniques were applied to estimate the missing data. The performance of imputation techniques was measured by mean square error (MSE) in which estimation error of selected statistical properties from the actual complete dataset was compared in various techniques. An imputation technique with the least estimation error was selected and further investigated with various missing data properties. In addition, a minimal feature set was selected to represent the imputed dataset with the lowest deviation of feature values from complete dataset features. Then the minimal feature set was used to train the MLP-based network to recognize six possible patterns, namely, random, increasing/decreasing trend, shift up/down and cyclic patterns from the process data streams with missing data. The MLP network performance was measured in several missing data scenarios with/without imputation and features.

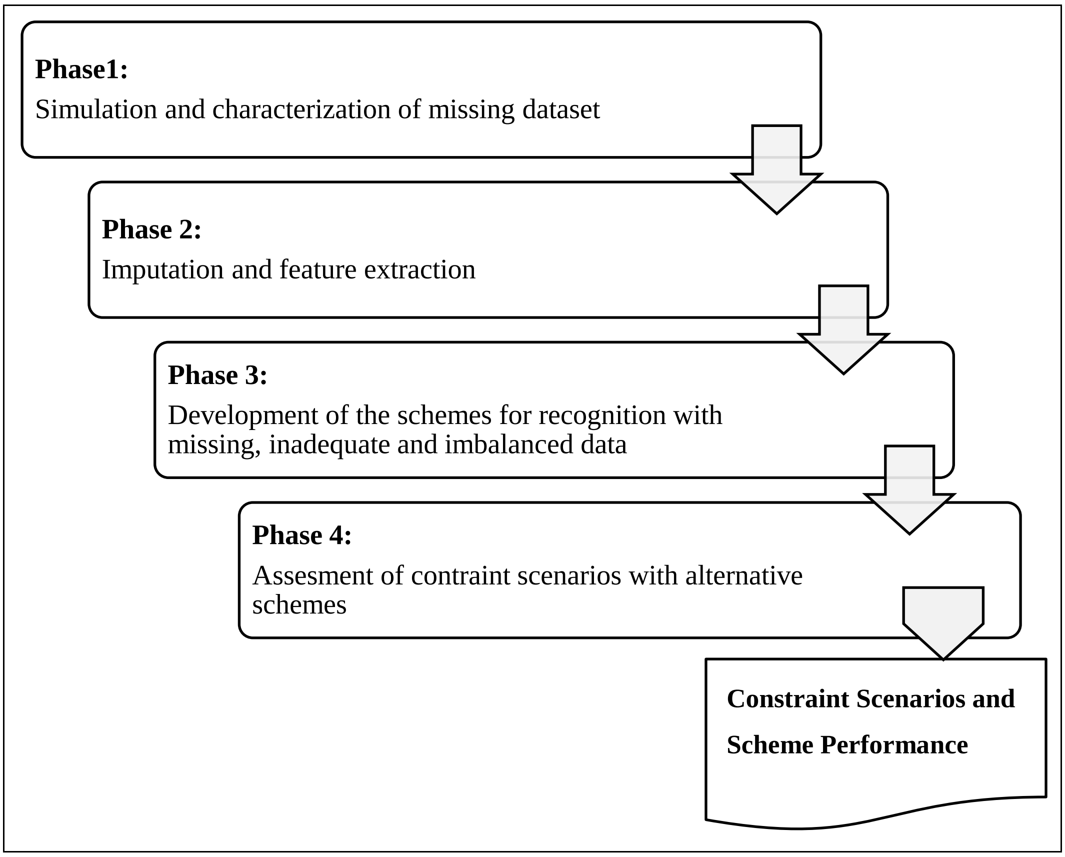

Figure 1.2 Operational framework for research activities

In the third phase, the feature set was updated to suit the recognition of streaming data. The performance of the recognizer trained with features extracted from partially developed CCPs was measured with or without missing data. Then the effect of training the MLP recognizer with balanced and imbalanced dataset was examined. Also, size of training dataset was varied and the simultaneous effect of imbalanced and inadequate dataset in training and recognition performance was studied. To improve recognition performance, an oversampling technique and use of a hybrid ANN-fuzzy recognizer was proposed to address training with missing, inadequate and imbalanced data. Three schemes from the solution approaches were developed and their performance were compared after training with various dataset size and imbalance rates. Finally, in the fourth phase, assessment of constraint scenarios with alternative schemes was studied and the performance in the recommended scheme was compared with alternative schemes in the literature. Each phase serves a building block towards achieving the research objectives.

1.8 Definition of Terms

1. Constrained data

Constrained data in this research refers to process data streams which are missing, inadequate or imbalanced.

2. Constraint Control chart patterns (CCPs)

Control chart patterns were introduced in Western Electric Company (1958). Six basic CCPs are (a) normal, (b) cyclic, (c) increasing trend, (d) decreasing trend, (e) upward shift, and (f) downward shift patterns.

3. Imputation

In statistics, imputation is the process of replacing missing data with plausible substituted values.

4. Input representation

It is a process of transforming raw input data into a new format for representation of input into the recognizers. It involves pre-processing activities such as normalisation, standardization, feature extraction and feature selection.

5. Missing Mechanism

Missing mechanism group data by investigating whether the data is missing at random or not.

6. Missing at random

When the missingness is independent of the missing variable but it can be traced or predicted by other variables in the process, the missing mechanism is called missing at random (MAR).

7. Missing completely at random

When the probability that a value is missing is independent from the variable itself and any other variable in the process or external influence, the missing mechanism is called missing completely at random (MCAR).

8. Not missing at random

When there is a relationship between the value in the variable which is missing with the other observed values in the same variable, the missing mechanism is called not missing at random (NMAR).

9. Online monitoring

Online monitoring refers to a continuous tracking of process data streams to identify whether a process has deviated from a normal operating condition.

10. Online recognition

Online recognition is a process of recognising an unknown CCP and to assign it to one of the prescribed classes.

11. Inadequate data

Majority of supervised recognizers rely on a vast volume of data in order to learn all the possible stable and unstable process patterns. However, in short-run production and start-up processes, the historical dataset size is too small to train the recognizer. This issue is called inadequate data problem in this research and it is one of the data constraints.

12. Imbalanced data

One of the most common problems faced in real world databases is imbalanced data. The reason is that often one pattern particularly the one which represent the process in its stable state has more historical sample in comparison to the other patterns which represent the process in its unstable status. Performance of the recognizer which is trained with such dataset, is usually biased towards the majority class.

13. Scheme

Scheme is a recognition and constrained data treatment strategy which is formulated to address recognition of CCPs.

14. Baseline Scheme

Baseline scheme refers to CCPR scheme that is developed based on a ANN pattern recognizer and EWMA imputation which is capable of treatment of missing data.

15. Enhanced Schemes

Enhanced Schemes refer to recognition with hybrid ANN and fuzzy either with/without constrained data treatment using fuzzy information decomposition.

1.9 Structure of the Thesis

This thesis is organized into seven chapters. Chapter 1 provides an introduction to the research. Chapter 2 discusses background and review of relevant literature that is used to formulate research problem. Chapter 3 serves as the research methodology. Chapter 4 presents an investigation toward developing a baseline scheme for recognition of CCPs with missing data. Chapter 5 is an enhancement of the proposed baseline scheme for recognition of CCPs which can address inadequate and imbalanced data. Research findings are discussed in the Chapter 6 and the final chapter concludes this research. Figure 1.3 shows the structure of this thesis.

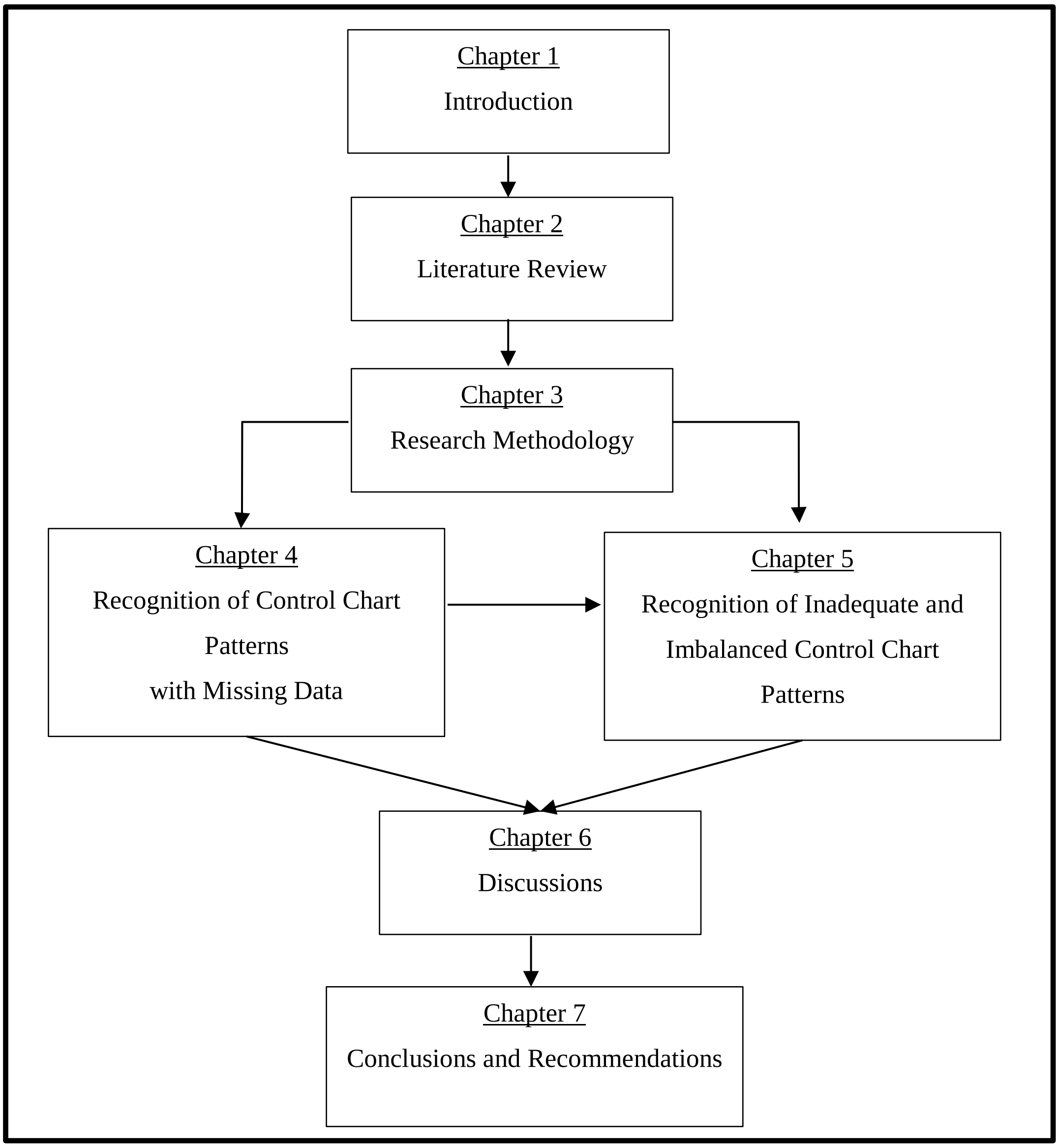

Figure 1.3 Structure of the thesis

REFERENCES

Abdelrahman Senoussi, A. M., Masood, I., Abdol Rahman, M. N., & Hassan, M. F. (2017). ‘Recognition of control chart patterns with incomplete samples’ IOP Conference Series: Materials Science and Engineering, 226(012098) .

Acuna, E., & Rodriguez, C. (2004). *The treatment of missing values and its effect on classifier accuracy*, in *Classification, Clustering, and Data Mining Applications*. Berlin: Springer, pp. 639-647.

Alaeddini, A., & Dogan, I. (2011). ‘Using Bayesian networks for root cause analysis in statistical process control’, *Expert Systems with Applications, 38*(9), 11230-11243.

Alcock, R. J., & Manolopoulos, Y. (1999). Time-Series Queries Employing a Feature-Based Approach. *7th Hellenic Conference on Informatics*, August 27-29, Greece.

AlGhanim, A. (1997) ‘An unsupervised learning neural algorithm for identifying process behavior on control charts and a comparison with supervised learning approaches’, *Computers & Industrial Engineering, 32*(3), 627-639.

Arteaga, F., & Ferrer, A. (2005) ‘Framework for regression-based missing data imputation methods in on-line MSPC’, *Journal of Chemometrics, 19*(8), 439-447.

Askarian, M., Benitez, R., Graells, M., & Zarghami, R. (2016) ‘Data-based fault detection in chemical processes: Managing records with operator intervention and uncertain labels’ *Expert Systems with Applications, 63*, 35-48.

Bag, M., Gauri, S. K., & Chakraborty, S. (2012) ‘An expert system for control chart pattern recognition’ *International Journal of Advanced Manufacturing Technology, 62*(1-4), 291-301.

Balakrishnan, N., Bersimis, S., & Koutras, M. V. (2009) ‘Run and Frequency Quota Rules in Process Monitoring and Acceptance Sampling’ , *Journal of Quality Technology, 41*(1), 66-81.

Bhadra, T., & Bandyopadhyay, S. (2015) ‘Unsupervised feature selection using an improved version of Differential Evolution’, *Expert Systems with Applications, 42*(8), 4042-4053.

Brown, M. L., & Kros, J. F. (2003) 'Data mining and the impact of missing data' , *Industrial Management & Data Systems, 103*(8), 611-621.

Camacho, F. L., Torres, R., & Pollan, R. R. (2015) Classification of Antimicrobial Peptides with Imbalanced Datasets. In E. Romero, N. Lepore, J. D. GarciaArteaga, & J. Brieva (Eds.), *11th International Symposium on Medical Information Processing and Analysis* (Vol. 9681).

Capizzi, G. (2015) 'Recent Advances in Process Monitoring: Nonparametric and Variable-Selection Methods for Phase I and Phase II', *Quality Engineering, 27*(1), 44-67.

Capizzi, G., & Masarotto, G. (2012) 'An enhanced control chart for start-up processes and short runs' , *Quality Technology and Quantitative Management, 9*(2), 189-202.

Castagliola, P., Achouri, A., Taleb, H., Celano, G., & Psarakis, S. (2013) 'Monitoring the Coefficient of Variation Using Control Charts with Run Rules' , *Quality Technology and Quantitative Management*, *10*(1), 75-94.

Castillo, E. D., & Montgomery, D. C. (1994) 'Short-run statistical process control: Q-chart enhancements and alternative methods', *Quality and Reliability Engineering International, 10*(2), 87-97.

Celano, G., Castagliola, P., Fichera, S., & Nenes, G. (2013) 'Performance of t control charts in short runs with unknown shift sizes' , *Computers & Industrial Engineering, 64*(1), 56-68.

Celano, G., Castagliola, P., & Trovato, E. (2012) 'The economic performance of a CUSUM t control chart for monitoring short production runs' *Quality Technology and Quantitative Management*, *9*(4), 329-354.

Chandra, M. J. (2001). *Statistical quality control.* 1st edn (June 21, 2001): CRC Press.

Chen, J., & Liang, Y. (2016) 'Development of fuzzy logic-based statistical process control chart pattern recognition system' , *International Journal of Advanced Manufacturing Technology, 86*(1-4), 1011-1026.

Chen, Y. (2015). *The UCR Time Series Classification Archive.* Available at: www.cs.ucr.edu/~eamonn/time_series_data/.

Chen, Z., Lu, S., & Lam, S. (2007) 'A hybrid system for SPC concurrent pattern recognition' , *Advanced Engineering Informatics, 21*(3), 303-310.

Cheng, C.-S. (1989) Group technology and expert systems concepts applied to statistical process control in small-batch manufacturing. PhD Thesis, Arizona State University Tempe, Arizona.

Cheng, C. S., Huang, K. K., & Chen, P. W. (2015) 'Recognition of control chart patterns using a neural network-based pattern recognizer with features extracted from correlation analysis' , *Pattern Analysis and Applications, 18*(1), 75-86.

Cheng, C. S., & Hubele, N. F. (1996) 'A pattern recognition algorithm for an (x)over-bar control chart' , *Iie Transactions, 28*(3), 215-224.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. 3rd edn. Londen: lawrence erlbaum associates, publishers.

Dialameh, M., & Jahromi, M. Z. (2017) 'A general feature-weighting function for classification problems' , *Expert Systems with Applications*, *72*, 177-188.

Dua, D. (2017). *UCI Machine Learning Repository*. University of California, Irvine, School of Information and Computer Sciences, Available at http://archive.ics.uci.edu/ml.

Ebrahimzadeh, A., Ranaee, V. (2009). *A Hybrid Intelligent Technique for Recognition of Control Chart Patterns*. New York: IEEE.

El-Midany, T. T., El-Baz, M. A., & Abd-Elwahed, M. S. (2010) 'A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks' , *Expert Systems with Applications, 37*(2), 1035-1042.

Farhangfar, A., Kurgan, L., & Dy, J. (2008) 'Impact of imputation of missing values on classification error for discrete data' , *Pattern Recognition, 41*(12), 3692-3705.

Furutani, H., Yamamoto, K., Ogura, H., & Kitazoe, Y. (1988) 'New control chart for multivariate data with missing values' , *Computers and Biomedical Research, 21*(1), 1-8.

Garcia-Diaz, J. C. (2007). 'The effective variance control chart for monitoring the dispersion process with missing data' , *European Journal of Industrial Engineering*, *1*(1), 40-55.

García-Laencina, P. J., Sancho-Gómez, J.-L., & Figueiras-Vidal, A. R. (2010) 'Pattern classification with missing data: a review' , *Neural Computing and Applications, 19*(2), 263-282.

Garjani, M., Noorossana, R., & Saghaei, A. (2010) 'A neural network-based control scheme for monitoring start-up processes and short runs' , *The International Journal of Advanced Manufacturing Technology, 51*(9), 1023-1032.

Garson, G. (2012). *Missing values analysis and data imputation*. Asheboro, NC: Statistical Publishing Associates.

Gauri, S. K. (2010) 'Control chart pattern recognition using feature-based learning vector quantization' , *International Journal of Advanced Manufacturing Technology, 48*(9-12), 1061-1073.

Gauri, S. K. (2012) 'Improved feature-based test statistic for assessing suitability of the preliminary samples for constructing control limits of (X)over-bar chart' , *International Journal of Advanced Manufacturing Technology, 58*(9-12), 1171-1187.

Gauri, S. K., & Chakraborty, S. (2006) 'Feature-based recognition of control chart patterns' , *Computers & Industrial Engineering, 51*(4), 726-742.

Gauri, S. K., & Chakraborty, S. (2008) 'Improved recognition of control chart patterns using artificial neural networks' , *International Journal of Advanced Manufacturing Technology, 36*(11-12), 1191-1201.

Gauri, S. K., & Chakraborty, S. (2009) 'Recognition of control chart patterns using improved selection of features' , *Computers & Industrial Engineering, 56*(4), 1577-1588.

Ge, Z., Song, Z., & Gao, F. (2013a). 'Review of Recent Research on Data-Based Process Monitoring' , *Industrial & Engineering Chemistry Research, 52*(10), 3543-3562.

Ge, Z. Q., Song, Z. H., & Gao, F. R. (2013b) 'Self-Training Statistical Quality Prediction of Batch Processes with Limited Quality Data' , *Industrial & Engineering Chemistry Research, 52*(2), 979-984.

Gebremeskel, G. B., Yi, C., Wang, C., & He, Z. (2015) 'Critical analysis of smart environment sensor data behavior pattern based on sequential data mining techniques' , *Industrial Management & Data Systems, 115*(6), 1151-1178.

Grzenda, M., Bustillo, A., & Zawistowski, P. (2012) ‘A soft computing system using intelligent imputation strategies for roughness prediction in deep drilling’ , *Journal of Intelligent Manufacturing, 23*(5), 1733-1743.

Gu, N., Cao, Z., Xie, L., Creighton, D., Tan, M., & Nahavandi, S. (2013) ‘Identification of concurrent control chart patterns with singular spectrum analysis and learning vector quantization’ , *Journal of Intelligent Manufacturing, 24*(6), 1241-1252.

Guh, R.-S. (2004) ‘Optimizing feedforward neural networks for control chart pattern recognition through genetic algorithms’ , *International Journal of Pattern Recognition and Artificial Intelligence, 18*(02), 75-99.

Guh, R. S. (2003) ‘Integrating artificial intelligence into on-line statistical process control’ , *Quality and Reliability Engineering International, 19*(1), 1-20.

Guh, R. S., & Tannock, J. D. T. (1999) ‘Recognition of control chart concurrent patterns using a neural network approach’ , *International Journal of Production Research, 37*(8), 1743-1765.

Guh, R. S., Zorriassatine, F., Tannock, J. D. T., & O'Brien, C. (1999) ‘On-line control chart pattern detection and discrimination - a neural network approach’ , *Artificial Intelligence in Engineering, 13*(4), 413-425.

Gutierrez, H. d. l. T., & Pham, D. T. (2016) ‘Estimation and generation of training patterns for Control Chart Pattern Recognition’ , *Computers & Industrial Engineering*.

Hachicha, W., & Ghorbel, A. (2012) ‘A survey of control-chart pattern-recognition literature (1991-2010) based on a new conceptual classification scheme’ , *Computers & Industrial Engineering, 63*(1), 204-222.

Haghighati, R., & Hassan, A. (2014) ‘A conceptual methodology for recognition of constrained control chart patterns’ , *Advanced Materials Research* (Vol. 845, pp. 696-700).

Haghtalab, S., Xanthopoulos, P., & Madani, K. (2015) ‘A robust unsupervised consensus control chart pattern recognition framework’ , *Expert Systems with Applications, 42*(19), 6767-6776.

Hassan, A. (2002) *On-Line Recognition of Developing Control Chart Patterns*. PhD thesis. Faculty of Mechanical Engineering, Universiti Teknologi Malaysia. Skudai.

Hassan, A. (2011) ‘An improved scheme for online recognition of control chart patterns’ , *International Journal of Computer Aided Engineering and Technology, 3*(3-4), 309-321.

Hassan, A., Shariff Nabi Baksh, M., Shaharoun, A. M., & Jamaluddin, H. (2003) ‘Improved SPC chart pattern recognition using statistical features’ , *International Journal of Production Research, 41*(7), 1587-1603.

He, H. B., Chen, S., Man, H., Desai, S., & Quoraishee, S. (2010). *Imbalanced Learning for Pattern Recognition: An Empirical Study*, in E. M. Carapezza (Ed.), *Unmanned-Unattended Sensors and Sensor Networks Vii* (Vol. 7833). Toulouse, France: Proceedings of SPIE-The International Society for Optical Engineering.

He, Q. P., & Wang, J. (2011) ‘Statistics Pattern Analysis: A New Process Monitoring Framework and its Application to Semiconductor Batch Processes’ , *Aiche Journal, 57*(1), 107-121.

He, Z., Wang, Z., Tsung, F., & Shang, Y. (2016) ‘A Control Scheme for Autocorrelated Bivariate Binomial Data’, *Computers & Industrial Engineering*. Volume 98 Issue C, 350-359.

Hillier, F. S. (1962). *X chart control limits based on a small number of subgroups*: Applied Mathematics and Statistics Laboratories, California: Stanford University.

Ho, L. L., & Trindade, A. L. G. (2009) ‘Economic design of an chart for short-run production’ , *International Journal of Production Economics, 120*(2), 613-624.

Homenda, W., & Lesinski, W. (2014). *Decision Trees and Their Families in Imbalanced Pattern Recognition: Recognition with and without Rejection*, in Saeed, K. & Snasel, V. (Eds.), *Computer Information Systems and Industrial Management, Cisim 2014* .Vol. 8838, pp. 219-230.

Hwarng, H. B., & Hubele, N. F. (1993) ‘Back-propagation pattern recognizers for X control charts: methodology and performance’ , *Computers and Industrial Engineering, 24*(2), 219-235.

Jackson, Q., & Landgrebe, D. A. (2001) ‘An adaptive classifier design for high-dimensional data analysis with a limited training data set’ , *IEEE Transactions on Geoscience and Remote Sensing, 39*(12), 2664-2679.

Jacques, J., Taillard, J., Delerue, D., Dhaenens, C., & Jourdan, L. (2015) ‘Conception of a dominance-based multi-objective local search in the context of classification rule mining in large and imbalanced data sets’ , *Applied Soft Computing, 34*(0), 705-720.

Jensen, W. A., Birch, J. B., & Woodall, W. H. (2008) ‘Monitoring correlation within linear profiles using mixed models’ , *Journal of Quality Technology, 40*(2), 167-183.

Jones-Farmer, L. A., Ezell, J. D., & Hazen, B. T. (2014) ‘Applying control chart methods to enhance data quality’, *Technometrics, 56*(1), 29-41.

Kao, L. J., Lee, T. S., & Lu, C. J. (2016) ‘A multi-stage control chart pattern recognition scheme based on independent component analysis and support vector machine’ , *Journal of Intelligent Manufacturing, 27*(3), 653-664.

Kazemi, M. S., Kazemi, K., Yaghoobi, M. A., & Bazargan, H. (2016) ‘A hybrid method for estimating the process change point using support vector machine and fuzzy statistical clustering’ , *Applied Soft Computing, 40*, 507-516.

Khajehzadeh, A., & Asady, M. (2015) ‘Recognition of Control Chart Patterns Using adaptive neuro-fuzzy inference system and Efficient Features’ , *International Journal of Scientific & Engineering Research, 6*(9).

Khormali, A., & Addeh, J. (2016) ‘A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine’ , *Isa Transactions, 63*, 256-264.

Kim, S. B., Jitpitaklert, W., Chen, V. C. P., Lee, J., & Park, S.-K. (2013) ‘Data mining model adjustment control charts for cascade processes’ , *European Journal of Industrial Engineering, 7*(4), 442-455.

Kim, S. K., & Kirchner, E. A. (2016) ‘Handling Few Training Data: Classifier Transfer Between Different Types of Error-Related Potentials’ , *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24*(3), 320-332.

King, C., & Pendlebury, D. A. (2013). *Research Fronts 2013*. Thomson reuters, Available at http://nldxb.njfu.edu.cn/Upload/Park/0e8006c0-e367-445e-8078-d07681cfa339.pdf.

Kolter, J. Z., & Maloof, M. A. (2007) ‘Dynamic weighted majority: An ensemble method for drifting concepts’ , *Journal of Machine Learning Research, 8*(Dec), 2755-2790.

Koutras, M. V., Bersimis, S., & Antzoulakos, D. L. (2006) ‘Improving the performance of the chi-square control chart via runs rules’ , *Methodology and Computing in Applied Probability, 8*(3), 409-426.

Koutras, M. V., Bersimis, S., & Maravelakis, P. E. (2007) ‘Statistical process control using shewhart control charts with supplementary runs rules’ , *Methodology and Computing in Applied Probability, 9*(2), 207-224.

Krawczyk, B., Jelen, L., Krzyzak, A., & Fevens, T. (2014a) *One-Class Classification Decomposition for Imbalanced Classification of Breast Cancer Malignancy Data*, in Rutkowski, L. , Korytkowski, M., Scherer, R., Tadeusiewicz, L. A. Zadeh, L. Z. & Zurada, L.Z. (Eds.), *Artificial Intelligence and Soft Computing Icaisc 2014, Pt I* .Vol. 8467, pp. 539-550.

Krawczyk, B., Wozniak, M., & Schaefer, G. (2014b) ‘Cost-sensitive decision tree ensembles for effective imbalanced classification’ , *Applied Soft Computing, 14*, 554-562.

Kubat, M., & Matwin, S. (1997) Addressing the curse of imbalanced training sets: one-sided selection. *14th international conference on machine learning*. Ontario, Canada, 179-186.

Lambert, D., & Liu, C. H. (2006) ‘Adaptive thresholds: Monitoring streams of network counts’ *Journal of the American statistical Association, 101*(473), 78-88.

Lesany, S. A., Koochakzadeh, A., & Ghomi, S. (2014) ‘Recognition and classification of single and concurrent unnatural patterns in control charts via neural networks and fitted line of samples’ , *International Journal of Production Research, 52*(6), 1771-1786.

Liao, T. W. (2005) ‘Clustering of time series data—a survey’, *Pattern Recognition, 38*(11), 1857-1874.

Little, J. A. (1988) ‘A Test of Missing Completely at Random for Multivariate Data with Missing Values’ , *Journal of the American statistical Association, 83*(404), 1198-1202.

Little, R., & Rubin, D. B. (2002) *Statistical analysis with missing data*. 2nd edn. New York: John Wiley.

Liu, S., Zhang, J., Xiang, Y., & Zhou, W. (2017) ‘Fuzzy-Based Information Decomposition for Incomplete and Imbalanced Data Learning’ , *IEEE Transactions on Fuzzy Systems, 25*(6), 1476-1490.

Lomax, R.G., & Hahs-Vaughn, D.L. (2013). *Introduction to Statistical Concepts*. 3rd edn. Routledge.

Lu, C.-J., Shao, Y. E., & Li, P.-H. (2011) ‘Mixture control chart patterns recognition using independent component analysis and support vector machine’ , *Neurocomputing, 74*(11), 1908-1914.

Lucas, J. M., & Saccucci, M. S. (1990) ‘Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements’ , *Technometrics, 32*(1), 1-12.

Madbouly, A., Gragert, L., Freeman, J., Leahy, N., Gourraud, P. A., Hollenbach, J. A., Maiers, M. (2014) ‘Validation of statistical imputation of allele-level multilocus phased genotypes from ambiguous HLA assignments’ , *Tissue Antigens, 84*(3), 285-292.

Madbuly, D. F., Maravelakis, P. E., & Mahmoud, M. A. (2013) ‘The Effect of Methods for Handling Missing Values on the Performance of the MEWMA Control Chart’ , *Communications in Statistics-Simulation and Computation, 42*(6), 1437-1454.

Mahmoud, M. A., Saleh, N. A., & Madbuly, D. F. (2014) ‘Phase I Analysis of Individual Observations with Missing Data’ , *Quality and Reliability Engineering International, 30*(4), 559-569.

Malarvizhi, M. R., & Thanamani, D. A. S. (2012) ‘K-Nearest Neighbor in Missing Data Imputation’ , *International Journal of Engineering Research and Development*, 5(1), 05-07.

Mao, W. T., Wang, J. W., Wang, L. Y. (2015). Online Sequential Classification of Imbalanced Data by Combining Extreme Learning Machine and improved SMOTE Algorithm *2015 International Joint Conference on Neural Networks*. Jul 12-17. Ireland, 2161-4393.

Masud, M. M., Woolam, C., Gao, J., Khan, L., Han, J. W., Hamlen, K. W., & Oza, N. C. (2012) ‘Facing the reality of data stream classification: coping with scarcity of labeled data’ , *Knowledge and Information Systems, 33*(1), 213-244.

Matsumoto, M., & Nishimura, T. (1998) ‘Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator’ , *ACM Transactions on Modeling and Computer Simulation (TOMACS), 8*(1), 3-30.

Montgomery, D. C. (2007). *Introduction to statistical quality control*. 7th edn. John Wiley & Sons.

Nasr, E. S. A., Al-Mubaid, H. (2009). Mining Process Control Data Using Machine Learning. *2009 international conference on computers and industrial engineering.* Jul 06-09. Troyes, France, 1434-1439.

Nelson, L. S. (1984). Column: Technical aids: The Shewhart control chart–tests for special causes. *Journal of Quality Technology, 16*(4).

Ni, J. C., Li, L., Qiao, F., & Wu, Q. D. (2014) ‘A GS-MPSO-WKNN method for missing data imputation in wireless sensor networks monitoring manufacturing conditions’ , *Transactions of the Institute of Measurement and Control, 36*(8), 1083-1092.

Noorossana, R., Toosheghanian, M., & Gazaneh, F. M. (2013) ‘Using genetic algorithm and response surface methodology for statistically constrained optimization of VSI X-bar control charts under multiple assignable causes and non-normality’ , *The International Journal of Advanced Manufacturing Technology*, *67*(9-12), 2325-2342.

Olgun, M. O., & Ozdemir, G. (2012) ‘Control chart pattern recognition using statistical-feature based bayes classifier’ , *Journal of the Faculty of Engineering and Architecture of Gazi University, 27*(2), 303-311.

Pacella, M., Semeraro, Q., & Anglani, A. (2004) ‘Adaptive Resonance Theory-based neural algorithms for manufacturing process quality control’ , *International Journal of Production Research, 42*(21), 4581-4607.

Pasini, A. (2015) ‘Artificial neural networks for small dataset analysis’ , *Journal of Thoracic Disease, 7*(5), 953-960.

Pelegrina, G. D., Duarte, L. T., & Jutten, C. (2016) ‘Blind source separation and feature extraction in concurrent control charts pattern recognition: Novel analyses and a comparison of different methods’ , *Computers & Industrial Engineering*, *92*, 105-114.

Perez-Rodriguez, J., de Haro-Garcia, A., & Garcia-Pedrajas, N. (2011) Instance selection for class imbalanced problems by means of selecting instances more than once, in Lozano, J. A. Gamez, J. A. & Moreno, J.A. (Eds.). *Advances in Artificial Intelligence* .Vol. 7023, pp. 104-113.

Pham, D. T., & Chan, A. B. (1998). Control chart pattern recognition using a new type of self-organizing neural network. *Proceedings of the Institution of Mechanical Engineers Part I-Journal of Systems and Control Engineering, 212*(I2), 115-127.

Pham, D. T., & Chan, A. B. (2001) 'Unsupervised adaptive resonance theory neural networks for control chart pattern recognition' , *Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture, 215*(1), 59-67.

Pham, D. T., & Oztemel, E. (1994) 'Control chart pattern-recognition using learning vector quantization networks' , *International Journal of Production Research, 32*(3), 721-729.

Pham, D. T., & Sagiroglu, S. (2001) 'Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms' , *International Journal of Machine Tools and Manufacture*, 41, 419–430.

Pham, D. T., & Wani, M. A. (1997) 'Feature-based control chart pattern recognition' *, International Journal of Production Research, 35*(7), 1875-1890.

Preda, C., Duhamel, A., Picavet, M., & Kechadi, T. (2005) *Tools for Statistical Analysis with Missing Data: Application to a Large Medical Database*, in R. Engelbrecht, R, Geissbuhler, A. Lovis, C. & Mihalas, G. (Eds.), *Connecting Medical Informatics and Bio-Informatics* Vol. 116, pp. 181-186. Amsterdam: Ios Press.

Quanz, B. (2012) *Learning with Low-Quality Data: Multi-View Semi-Supervised Learning with Missing Views.*PhD Thesis, University of Kansas, Lawrence.

Quesenberry, C. P. (1991) 'SPC Q charts for start-up processes and short or long runs', *Journal of Quality Technology, 23*(3), 213-224.

Radtke, P. V. W., Granger, E., Sabourin, R., Gorodnichy, D. (2012). Adaptive Selection of Ensembles for Imbalanced Class Distributions. *21st International Conference on Pattern Recognition*. Nov 11-15. Tsukuba, Japan. 2980-2984.

Ranaee, V., & Ebrahimzadeh, A. (2013) 'Control chart pattern recognition using neural networks and efficient features: a comparative study' , *Pattern Analysis and Applications, 16*(3), 321-332.

Ranaee, V., Ebrahimzadeh, A., & Ghaderi, R. (2010) 'Application of the PSO-SVM model for recognition of control chart patterns' , *Isa Transactions, 49*(4), 577-586.

Ranjan, S., Yu, C., Zhang, C., Kelly, F., & Hansen, J. H. L. (2016). Language recognition using deep neural networks with very limited training data. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 20-25 March. 5830-5834.

Rässler, S., Rubin, D. B., & Zell, E. R. (2013) 'Imputation' , *Wiley Interdisciplinary Reviews: Computational Statistics, 5*(1), 20-29.

Romao, X., Delgado, R., & Costa, A. (2010) 'An empirical power comparison of univariate goodness-of-fit tests for normality', *Journal of Statistical Computation and Simulation, 80*(5), 545-591.

Rooney, J. J., & Heuvel, L. N. V. (2004) 'Root cause analysis for beginners' , *Quality progress, 37*(7), 45-56.

Salehi, M., Kazemzadeh, R. B., & Salmasnia, A. (2012) 'On line detection of mean and variance shift using neural networks and support vector machine in multivariate processes' , *Applied Soft Computing, 12*(9), 2973-2984.

Sandhan, T., Choi, J. Y. (2014) Handling imbalanced datasets by partially guided hybrid sampling for pattern recognition. *22nd International Conference on Pattern Recognition.* Aug 24-28. Stockholm, Sweden, pp. 1449-1453.

Sarakit, P., Theeramunkong, T., Haruechaiyasak, C. (2015) Improving Emotion Classification in Imbalanced YouTube Dataset Using SMOTE Algorithm. *2nd International Conference on Advanced Informatics: Concepts, Theory and Applications Icaicta*. Aug 19-22, 2015. Chonburi, Thailand.

Scheffer, J. (2002) 'Dealing with missing data', *Research Letters in the Information and Mathematical Sciences, 3*, 153-160.

Shewhart, W. (1925) 'The application of statistics as an aid in maintaining quality of a manufactured product', *Journal of the American statistical Association, 20*(152), 546-548.

Shewhart, W. A., & Deming, W. E. (1939). *Statistical method from the viewpoint of quality control.* Washington, D.C: Courier Corporation.

Silva-Ramírez, E.-L., Pino-Mejías, R., & López-Coello, M. (2015) 'Single imputation with multilayer perceptron and multiple imputation combining multilayer perceptron and k-nearest neighbours for monotone patterns', *Applied Soft Computing, 29*(0), 65-74.

Silvert, W., & Baptist, M. (2000). *Can Neuronal Networks be Used in Data-Poor Situations?*, in Lek, S. & Guégan, J. F. (Eds.), *Artificial Neuronal Networks*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 241-248.

Simões, F. D., Leoni, R. C., Machado, M. A. G., & Costa, A. F. B. (2016) 'Synthetic charts to control bivariate processes with autocorrelated data' , *Computers & Industrial Engineering, 97*, 15-25.

Sullivan, D., & Andridge, R. (2015) 'A hot deck imputation procedure for multiply imputing nonignorable missing data: The proxy pattern-mixture hot deck' , *Computational Statistics & Data Analysis, 82*, 173-185.

Sun, Y. M., Kamel, M. S., Wong, A. K. C., & Wang, Y. (2007) 'Cost-sensitive boosting for classification of imbalanced data' , *Pattern Recognition, 40*(12), 3358-3378.

Suyundikov, A., Stevens, J. R., Corcoran, C., Herrick, J., Wolff, R. K., & Slattery, M. L. (2015) 'Accounting for Dependence Induced by Weighted KNN Imputation in Paired Samples, Motivated by a Colorectal Cancer Study' , *Plos One*, *10*(4).

Swift, J. A. (1987) *Development of a knowledge-based expert system for control-chart pattern recognition and analysis.* PhD thesis. Oklahoma State University.US.

Teoh, W. L., & Khoo, M. B. C. (2012). A preliminary study on the double sampling X-bar chart with unknown parameters. *International Conference on Innovation Management and Technology Research (ICIMTR).*

Veiga, P., Mendes, L., & Lourenço, L. (2015). A retrospective view of statistical quality control research and identification of emerging trends: a bibliometric analysis. *Quality and Quantity*.

Vivaracho-Pascual, C., Simon-Hurtado, A. (2010). Improving ANN Performance for Imbalanced Data Sets by Means of the NTIL Technique *2010 International Joint Conference on Neural Networks Ijcnn 2010*. 21-22 May. Malacca, Malaysia. 80-84.

Wasserman, G. S. (1994) 'Short run spc using dynamic control chart', *Computers & Industrial Engineering, 27*(1–4), 353-356.

Waterhouse, M., Smith, I., Assareh, H., & Mengersen, K. (2010) ‘Implementation of multivariate control charts in a clinical setting’ , *International Journal for Quality in Health Care, 22*(5), 408-414.

Waterhouse, M. A., Moser, C., Brighouse, R. D., Foster, K. A., Smith, I. R., & Mengersen, K. (2013) ‘Data Quality Improvement in Clinical Databases Using Statistical Quality Control: Review and Case Study’ , *Therapeutic Innovation & Regulatory Science, 47*(1), 70-81.

Western Electric, C. (1956). *Statistical quality control handbook*: Indianapolis: Western Electric Co.

Wilson, S. R. (2009). *Control Charts with Missing Observations.* PhD Dissertation, Virginia Polytechnic Institute, Blacksburg.

Wu, S., Wu, B. (2006). Wavelet neural network-based control chart patterns recognition. *6th World Congress on Intelligent Control and Automation*. Jun 21-23. Dalian, China, 9718-9721.

Xanthopoulos, P., & Razzaghi, T. (2014) ‘A weighted support vector machine method for control chart pattern recognition’ , *Computers & Industrial Engineering, 70*, 134-149.

Xie, L., Gu, N., Li, D., Cao, Z., Tan, M., & Nahavandi, S. (2013) ‘Concurrent control chart patterns recognition with singular spectrum analysis and support vector machine’, *Computers & Industrial Engineering, 64*(1), 280-289.

Yang, S.F. (2000) ‘Statistical process control for short run manufacturing systems’ , *Process Control and Quality, 11*(5), 433-439.

Yang, W.A., & Zhou, W. (2015) ‘Autoregressive coefficient-invariant control chart pattern recognition in autocorrelated manufacturing processes using neural network ensemble’, *Journal of Intelligent Manufacturing, 26*(6), 1161-1180.

Yeh, C.W., Li, D.-C., & Zhang, Y.-R. (2012) ‘Estimation of a data-collection maturity model to detect manufacturing change’ , *Expert Systems with Applications, 39*(8), 7093-7101.

Zaman, M., & Hassan, A. (2018) ‘Improved statistical features-based control chart patterns recognition using ANFIS with fuzzy clustering’ , *Neural Computing and Applications*, 1-15.

Zan, T., Wang, M., & Fei, R. (2010) *Pattern Recognition for Control Charts Using AR Spectrum and Fuzzy ARTMAP Neural Network* in Jiang Z. & Zhang C. L.

(Eds.), *Manufacturing Science and Engineering, Pts 1-5*. Vol. 97-101, pp. 3696-3702.

Zhang, Q., Rahman, A., & D'este, C. (2013). Impute vs. Ignore: Missing values for prediction. *International Joint Conference on Neural Networks (IJCNN):*IEEE, 1-8.

Zhang, Y., Luo, B. (2008). Parallel classifiers ensemble with hierarchical machine learning for imbalanced classes. 7th *International Conference on Machine Learning and Cybernetics*. Jul 12-15. Kunming, China, 94-99.

Zheng, Z., Wu, X., & Srihari, R. (2004) 'Feature selection for text categorization on imbalanced data' , *ACM Sigkdd Explorations Newsletter, 6*(1), 80-89.

Zhou, X., Jiang, P., & Wang, X. (2018) 'Recognition of control chart patterns using fuzzy SVM with a hybrid kernel function' , *Journal of Intelligent Manufacturing*, *29*(1), 51-67.

List of Publications and Papers Presented

To date, this research has contributed three publications as follows:

Journal with Impact Factor

1. Haghighati, R., & Hassan, A. (2018). Recognition performance of imputed control chart patterns using exponential weighted moving average European Journal of Industrial Engineering, Vol. 12, No. 5, pp.637-660. **(Q3, IF: 1.085)**

Indexed Conference Proceedings

1. Haghighati, R., & Hassan, A. (2014). A conceptual methodology for recognition of constrained control chart patterns *Advanced Materials Research* (Vol. 845, pp. 696-700).

2. Haghighati, R., & Hassan, A. (2019). Feature extraction in control chart patterns with missing data *Journal of Physics. Ser. 1150*. 012013.