

The copyright © of this thesis belongs to its rightful author and/or other copyright owner. Copies can be accessed and downloaded for non-commercial or learning purposes without any charge and permission. The thesis cannot be reproduced or quoted as a whole without the permission from its rightful owner. No alteration or changes in format is allowed without permission from its rightful owner.



**RESERVOIR WATER RELEASE DYNAMIC DECISION MODEL
BASED ON SPATIAL TEMPORAL PATTERN**



SURIYATI BINTI ABDUL MOKHTAR

UUM
Universiti Utara Malaysia

**MASTER OF SCIENCE (INFORMATION TECHNOLOGY)
UNIVERSITI UTARA MALAYSIA
2016**

**< INSERT PERAKUAN KERJA TESIS / DISERTASI
(CERTIFICATION OF THESIS / DISSERTATION)>**



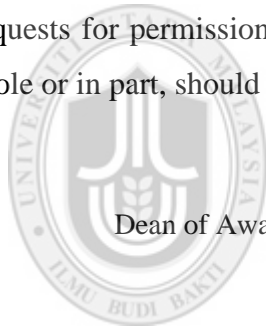
UUM

Universiti Utara Malaysia

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:



Dean of Awang Had Salleh Graduate School of Arts and Sciences

UUM College of Arts and Sciences

Universiti Utara Malaysia

06010 UUM Sintok

Abstrak

Keputusan pelepasan air bagi takungan serba guna memerlukan pakar bagi membuat keputusan dengan menghimpunkan maklumat pmutusan kompleks yang wujud secara masa nyata. Keputusan perlu mengambil kira lebih air takungan yang mencukupi bagi mengekalkan fungsi serba guna takungan dan menyediakan ruang yang mencukupi untuk hujan lebat dan aliran masuk air. Amat penting juga, pelepasan air tidak boleh melebihi tahap maksimum sungai di hilir supaya ia tidak akan menyebabkan banjir. Hujan dan paras air adalah maklumat kabur, oleh itu model keputusan memerlukan keupayaan untuk mengendalikan maklumat yang kabur. Tambahan pula, hujan yang direkodkan di lokasi yang berbeza mengambil masa berbeza untuk sampai ke dalam takungan. Situasi ini menunjukkan terdapat hubungan ruang masa yang tersembunyi di antara setiap stesen pengukur dan takungan. Oleh itu, kajian ini mencadangkan model keputusan pelepasan air takungan dinamik yang menggunakan kedua-dua maklumat ruang dan masa dalam corak input. Berdasarkan kepada corak berkenaan, model ini akan mencadangkan bila air takungan perlu dilepaskan. Model ini menggunakan Penyesuaian sistem inferens neuro-kabur (ANFIS) untuk mengendalikan dengan maklumat yang kabur. Data yang digunakan dalam kajian ini diperolehi daripada Jabatan Pengairan dan Saliran Perlis. Algoritma *Sliding Window* yang telah diubahsuai telah digunakan bagi membentuk corak masa bagi hujan, manakala maklumat ruang telah diperolehi melalui simulasi corak hujan dan paras air takungan yang telah dipetakan. Prestasi model telah diukur berdasarkan *Root Mean Square Error* (RMSE) dan *Mean Absolute Error* (MAE). Hasil daripada kajian ini menunjukkan ANFIS menghasilkan RMSE dan MAE paling rendah apabila dibandingkan dengan model *Autoregressive Integrated Moving Average* (ARIMA) dan *Backpropagation Neural Network* (BPNN). Model ini boleh digunakan oleh operator takungan bagi membantu pembuatan keputusan dan menyokong operator takungan baharu sewaktu ketiadaan operator yang berpengalaman.

Kata Kunci: Penyesuaian sistem inferens neuro-kabur (ANFIS), Kepintaran pengkomputan (CI), Pembuatan keputusan dinamik, Pembuatan keputusan pelepasan air takungan, Perolehan data ruang masa.

Abstract

The multi-purpose reservoir water release decision requires an expert to make a decision by assembling complex decision information that occurred in real time. The decision needs to consider adequate reservoir water balance in order to maintain reservoir multi-purpose function and provide enough space for incoming heavy rainfall and inflow. Crucially, the water release should not exceed the downstream maximum river level so that it will not cause flood. The rainfall and water level are fuzzy information, thus the decision model needs the ability to handle the fuzzy information. Moreover, the rainfalls that are recorded at different location take different time to reach into the reservoir. This situation shows that there is spatial temporal relationship hidden in between each gauging station and the reservoir. Thus, this study proposed dynamic reservoir water release decision model that utilize both spatial and temporal information in the input pattern. Based on the patterns, the model will suggest when the reservoir water should be released. The model adopts Adaptive Neuro-Fuzzy Inference System (ANFIS) in order to deal with the fuzzy information. The data used in this study was obtained from the Perlis Department of Irrigation and Drainage. The modified Sliding Window algorithm was used to construct the rainfall temporal pattern, while the spatial information was established by simulating the mapped rainfall and reservoir water level pattern. The model performance was measured based on the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Findings from this study shows that ANFIS produces the lowest RMSE and MAE when compare to Autoregressive Integrated Moving Average (ARIMA) and Backpropagation Neural Network (BPNN) model. The model can be used by the reservoir operator to assist their decision making and support the new reservoir operator in the absence of an experience reservoir operator.

Keywords: Adaptive Neuro-Fuzzy Inference System (ANFIS), Computational intelligence (CI), Dynamic decision making, Reservoir water release decision, Spatial temporal data mining.

Acknowledgement

First of all, I am thankful to Allah for giving me a chance and blessing me to complete this study. My sincere thanks and deep appreciation to Universiti Utara Malaysia and Ministry of Higher Education for funding this study under the Long Term Research Grant Scheme (LRGS). I would like to extend my thanks to the Department of Irrigation and Drainage Malaysia (DID) for supplying hydrology and reservoir operational data. I am also grateful to the education funding received from Ministry of Higher Education through MyBrain15 program.

First and foremost, I sincerely thank my parents, to my late father Abdul Mokhtar Ahmad and my mother Siwah Mat Long. Thanks for your support and understanding. Without both of you, I will have never been who I am. To the most important person in my life, my husband, Mursyidul 'Azim Mazlan, thank you for the love, support, sacrifices and for being with me all the time.

I am very grateful for my supervisors, En Wan Hussain Wan Ishak and Prof Dr Norita Md Norwawi, for the patience, understanding, encouragement and guidance from the beginning till the end. Special thanks to LRGS group 1 members, Prof Dr Ku Ruhana Ku Mahamud, Dr Nooraini Yusoff, Miss Aniza Mohd Din, En Zorkeflee Abu Hasan and other lecturers at Universiti Utara Malaysia who encouraged me to complete this study, and also have been my mentors when I met with difficulty during this period.

My gratitude also extends to friends who gave continuous support and encouragement: Kak Sha, Nurul, Mai, Ayu, Afdhal, Farez, Mat Yi, Sabri, Faiz, Ekin and many others.

Finally yet importantly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this study especially my family, lecturers and friends. Thank you again.

Table of Contents

Permission to Use.....	i
Abstrak	ii
Abstract	iii
Acknowledgement.....	iv
List of Tables.....	viii
List of Figures	x
List of Appendices	xii
List of Abbreviations.....	xiii
CHAPTER ONE INTRODUCTION	1
1.1 Research Background.....	1
1.2 Problem Statement	6
1.3 Research Questions	8
1.4 Research Objective.....	9
1.5 Research Scope	9
1.6 Research Significance	10
1.7 Organization of the Thesis	11
CHAPTER TWO LITERATURE REVIEW	12
2.1 Decision Making in Emergency Situation	12
2.2 Modelling Reservoir Water Release Decision	14
2.2.1 Reservoir Operation	14
2.2.2 Reservoir Water Release Decision.....	17
2.2.3 Reservoir Operating Policy.....	18
2.3 Computational Intelligence Approaches in Reservoir Operation and Decision Modelling	27
2.3.1 Artificial Neural Network	27
2.3.2 Fuzzy Logic	29
2.3.3 Genetic Algorithm	30
2.3.4 Hybrid Computational Intelligence Model	31
2.3.4.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)	32
2.3.4.2 Other Hybrid Computational Intelligence Model.....	35

2.3.5 Comparison of Computational Intelligence Techniques.....	36
2.4 Spatial Temporal Data Mining in Reservoir Operation	39
2.4.1 Temporal Data Mining.....	40
2.4.2 Spatial Data Mining	42
2.4.3 Spatial Temporal Data Mining.....	43
2.5 Summary	44
CHAPTER THREE RESEARCH FRAMEWORK AND METHODOLOGY .45	
3.1 Research Framework.....	45
3.2 Research Methodology.....	47
3.2.1 Spatial Temporal Data Mining.....	47
3.2.2 Development of Reservoir Water Release Dynamic Decision Model.....	54
3.3 Summary	58
CHAPTER FOUR SPATIAL TEMPORAL DATA MINING59	
4.1 Spatial Temporal Data Mining.....	59
4.1.1 Data Pre-processing	59
4.1.2 Spatial Temporal Pattern Formation using Sliding Window	63
4.2 Spatial Temporal Pattern Extraction Using Neural Network.....	67
4.2.1 Experimental Results on Spatial Temporal Pattern Extraction.....	68
4.2.2 Spatial Temporal Pattern of Rainfall	76
4.3 Spatial Temporal Pattern Extraction Procedure	77
4.4 Summary	79
CHAPTER FIVE PROPOSED RESERVOIR WATER RELEASE DYNAMIC DECISION MODEL BASED ON SPATIAL TEMPORAL PATTERN.....80	
5.1 Development of Reservoir Water Release Dynamic Decision Model.....	80
5.1.1 Constructing Fuzzy Membership Function.....	80
5.1.2 ANFIS Model Development	86
5.2 Development of Reservoir Water Release Dynamic Decision Model using ANFIS	89
5.3 Performance Evaluation	96
5.4 Reservoir Water Release Dynamic Decision Model.....	98
5.5 Summary	103

CHAPTER SIX CONCLUSION AND FUTURE WORK	104
6.1 Contribution of the Study	104
6.2 Future Work	105
6.2.1 Spatial Temporal Data Mining.....	105
6.2.2 Decision Variables Involved of the Modelling Techniques	106
6.3 Summary	107
REFERENCES	108



UUM

Universiti Utara Malaysia

List of Tables

Table 2.1 Application of Hedging Rules in Reservoir Operation	23
Table 2.2 Comparison of Computational Intelligence Techniques.....	37
Table 3.1 Five Categories of Rainfall Measurement	48
Table 3.2 Four Categories of Flood Stage	48
Table 3.3 Example of Reservoir Operational Data	50
Table 3.4 Example Input and Output of NN Model.....	53
Table 4.1 Mean Value for Flood Stage	60
Table 4.2 Fill in Missing Value	61
Table 4.3 Normalize Value for Rainfall Measurement.....	61
Table 4.4 Normalize Value for Flood Stage	62
Table 4.5 Data Normalization Using Min-Max Method.....	62
Table 4.6 Output Class Using BCD	63
Table 4.7 Temporal Data Representation.....	65
Table 4.8 Data set and the Number of Instances based on the Window Size.....	67
Table 4.9 Result for PADANG BESAR	69
Table 4.10 Neural Network Parameters for PADANG BESAR.....	70
Table 4.11 Results for TASOH.....	70
Table 4.12 Neural Network Parameters for TASOH.....	71
Table 4.13 Results for LUBUK SIREH.....	71
Table 4.14 Neural Network Parameters for LUBUK SIREH	72
Table 4.15 Results for KAKI BUKIT	73
Table 4.16 Neural Network Parameters for KAKI BUKIT	73
Table 4.17 Results for WANG KELIAN.....	74
Table 4.18 Neural Network Parameters for WANG KELIAN	75
Table 4.19 Summary of the Findings.....	75
Table 4.20 Summary of Network Specification.....	76
Table 5.1 Segmentation of x into two arbitrary classes (from raw data)	82
Table 5.2 Calculation of Primary Threshold Value (PRI)	83
Table 5.3 Calculation Threshold Value of NG Side	83
Table 5.4 Calculation Threshold Value of PO Side.....	84

Table 5.5 Calculation Threshold Value of PO Side.....	84
Table 5.6 Example of Training Set for ANFIS Model 1.....	88
Table 5.7 ANFIS Parameters	89
Table 5.8 Results for ANFIS Model	91
Table 5.9 ANFIS Model 1 Result	92
Table 5.10 ANFIS Model 2 Result	94
Table 5.11 Performance Evaluation Result.....	97



List of Figures

Figure 2.1. Classifications of Storage Zones of the Reservoir.....	15
Figure 2.2. Conceptual Model of Reservoir System	16
Figure 2.3. Standard Operating Policies (SOP)	19
Figure 2.4. Rule Curves for Multi-Purpose Reservoir	21
Figure 2.5. Types of Hedging Rules	22
Figure 2.6. Steps in KDD process	40
Figure 3.1. Research Framework	46
Figure 3.2. Data Pre-processing	49
Figure 3.3. Location of Timah Tasoh Reservoir and Five Gauging Stations	52
Figure 3.4. Example of Window Slice of 2.....	54
Figure 3.5. ANFIS Architecture.....	56
Figure 3.6. Process Flow for Development of Reservoir Water Release Dynamic Decision Model	57
Figure 4.1. Temporal Data Mining Classification	64
Figure 4.2. Pseudo-code for the Sliding Window	66
Figure 4.3. Pseudo-code for NN Training.....	68
Figure 4.4. Steps in Spatial Temporal Pattern Extraction Procedure.....	78
Figure 5.1. Illustration of Threshold Value Idea.....	82
Figure 5.2. Membership Function Graph for Flood Stage	85
Figure 5.3. Performance Comparison between Hybrid and Backpropagation Model	90
Figure 5.4. ANFIS Model Structure.....	93
Figure 5.5. Training Error & Checking error for Model 1.....	93
Figure 5.6. Rule Viewer for Model 1	94
Figure 5.7. Training error & Checking error for Model 2	95
Figure 5.8. Rule Viewer for Model 1	95
Figure 5.9. BPNN Model using Matlab tool	96
Figure 5.10. ARIMA Model using Matlab tool	97
Figure 5.11. RMSE and MAE.....	98
Figure 5.12. Five Layers of ANFIS Model.....	99

Figure 5.13. Input and Output Clustering 101
Figure 5.14. The Architecture of Reservoir Water Release Dynamic Decision Model
..... 102



List of Appendices

Appendix A Sample of Reservoir Operational Data (Year 2007)	118
Appendix B Sample of Hydrological Data (Year 2009).....	121



List of Abbreviations

ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
CI	Computational Intelligence
CBR	Case Based Reasoning
CDM	Classical Decision Making
GA	Genetic Algorithm
MFs	Membership Function
NDM	Naturalistic Decision Making
NN	Neural Network
RPD	Recognition Primed Decision
SW	Sliding Window



UUM
Universiti Utara Malaysia

CHAPTER ONE

INTRODUCTION

This chapter provides an overview of this study, which includes research background, problem statement, research objectives, scope of study, research significance and organization of the thesis.

1.1 Research Background

Emergency situation is one of the complex situations that require fast and accurate decisions, since the decision is very crucial to save human lives. Emergency situations can be identified by a few characteristics such as dynamic (Philips-Wren, 2009), complex (Norwawi, 2004) and action dependent (Feigh and Pritchett, 2006). Naturally, decisions that are made by people during these situations are based on instances and their past experiences. According to this view, the concept of new theory for understanding how people make decisions are important to illustrate that decisions which have been made do not only depend on a set of alternatives, but also based on their experiences. A decision is defined as the choice of one among a number of alternatives (Bohanec, 2001; Beach & Mitchell, 1978; Hersh, 1999). Naturalistic decision theory is one of the decision making approaches that defined how humans naturally make decisions in urgency and complex environments (Klein & Klinger, 1991). Decision makers have applied this theory during emergency situation such as fire fighting (Hersh, 1999). In emergency situations, typically, decisions are made under dynamic situations and also can be referred as dynamic decision making. Dynamic decision making can be defined as a series of decision that occurs in situations that change over time, where the future decision depends on

the earlier decision, and the state of the problem changes as a consequence of earlier actions of the decision makers (Brehmer, 1992; Busemeyer, 1999).

Flood is one of the emergency situations that need serious attention from the authorities involved. It is one of the natural disasters that strike due to heavy rainfall and extreme weather conditions. The impacts of flood are huge ranging from physical and structural damages to loss of life. In Malaysia, flood is one of major serious threats as the country faces two series of monsoons every year that bring heavy rain. Moreover, many areas in this country are geographically prone to flooding. A flood mitigation program has been executed in many countries to reduce its risk and impact. Flood mitigation can be divided into two approaches that are non-structural and structural approach (Tucci, 2002). Structural approach is related to the physical control of the flood, while non-structural approach is a procedure-based control. Physical control can be classified into developments of constructions or devices such as dams, planting forest and dikes, whereas procedure-based control can be described as flood forecasting and flood zoning activities.

The reservoir is one of the structural flood mitigation approaches that aim to store water during heavy rainfall in order to maintain safe discharges at the downstream areas (Smith and Ward, 1998). The reservoir can be defined as a natural or artificial lake, or pond, where water is collected and stored which are used for multi-purpose operation such as water supply, flood control, hydropower generation, and other purposes like agricultural and recreation (Wan Ishak et al., 2011c, Dictionary.com, 2012). The reservoir is constructed based on the types of reservoir systems that

consist of single and multi-reservoir systems (USACE, 1991). The single reservoir system represents the single operation of the reservoir, and the multi-reservoir system involves more than one reservoirs system operating at the same time.

Reservoir water release decision is an important action that needs to be accomplished by experienced reservoir operators due to extreme events or emergency situations. The reservoir capacity needs to be maintained in order to prevent downstream flood and to reduce water shortage problems in the future. During both situations, the gate opening decision is crucial as a late decision might cause overflow at the downstream river and might affect the dam's structure. However, different reservoirs have different purposes and multiple objective, thus requires different operation rules (Wan Ishak et al., 2012). Reservoir operation rule, or also known as operation policy, can be defined as a series of guidelines which provide guidance for reservoir operators to determine release decisions (Rittima, 2009).

Traditionally, reservoir operates based on the standard operating procedure (SOP) which was devised by the reservoir engineer when it was built. The SOP provides reservoir engineers the guidelines on how the reservoir should be operated. However, the reservoir operation is also influenced by several factors such as sediment (Naik & Jay, 2011; Talib et al., 2012), water usage, climate change (Koutroulis et al., 2013; Raje & Mujumdar, 2010), and urbanization at reservoir upstream (Suriya & Mudgal, 2012; Du et al., 2012). These factors reduce the efficacy of the SOP and the operation rule. Eventually, reservoir operators have to apply heuristic procedures by embracing rule curves and subjective judgements

(Ngo, 2006). Hence, the rules need to be periodically re-evaluated and updated to improve the reservoir operation. In practice, reservoir operators apply their heuristic decisions on top of the current standard operation procedure (SOP). The heuristic decisions are recorded in the reservoir operating log book. Over the years, the log book has become a knowledge repository that is rich with heuristic knowledge of the reservoir operation. This knowledge might be unique as different reservoirs may require different SOPs and employ different data features.

The reservoir operational data contains dynamic characteristics that refer to the dependent decisions which are made based on past human experiences. Typically, the water that flow into the reservoir comes from the upstream river network. The gauging station is the tools that record the river water level and rainfall. The gauging stations are located at different locations and the distance between the gauging stations and the reservoir varies. This information indicates that there are spatial relationships between the gauging stations and reservoir. In addition, the distance between the gauging stations and reservoir causes delay effect, where rainfall recorded at gauging stations may take some time to reach into the reservoir. This information is called temporal information.

Both spatial and temporal information can be extracted from the reservoir operational database using spatial temporal data mining technique. Spatial temporal data mining, or referred as spatio-temporal data mining, is an emerging research area in the application of computational techniques for spatial and temporal data analysis. Temporal data mining is concerned with data mining of large sequential data sets

(Laxman and Sastry, 2006). Spatial data mining, on the other hand, focuses on discovering meaningful patterns of large spatial data sets (Shekhar et al., 2010). In theory, temporal data can be classified as data that explicitly refer to time, where spatial data carries information that defines a location or distances. The combination of these two areas has led to the introduction of a new research area called spatial temporal data mining. Spatial temporal data mining covers both activities for discovering useful information or other patterns stored in spatial temporal database. Spatial temporal data mining identifies for both spatial and temporal relationship.

In recent decades, there have been many studies which applied computational intelligence techniques for solving complex problems in hydrological and water resource management including reservoir operation covering artificial neural network (ANN), evolutionary computation, genetic algorithms (GAs), fuzzy logic (FL) etc. Computational Intelligence can be defined is a set of computational methodologies and approaches to address complexities and computational problems due to the limitations of the traditional methods (Palit & Popovic, 2006). The application of hybrid computational intelligence (CI) techniques such as combinations of ANN and other methods including fuzzy logic (Valizadeh et al., 2014; Valizadeh & El-Shafie, 2013), genetic algorithm (Pinthong et al., 2009; Liu et al., 2011) etc, have been introduced with various strengths that are able to address the weaknesses of previous methods.

The focus of this research is to develop a reservoir water release dynamic decision model by using hybrid computational intelligence technique. The spatial temporal

data mining technique has been used to extract spatial temporal pattern from reservoir operational data. This model is aimed for improving the performances of reservoir operation and management.

1.2 Problem Statement

Reservoir water release decision is one of the critical actions that take place in the reservoir operation due to emergency situations. Flood is one of the emergency situations that need serious attentions and fast decisions from the experts in order to reduce the risk of flooding and save human lives. However, this situation requires an expert to make a decision by using their experience to control particular complex situations occurring in real time environment. Decisions made under these situations are dynamic (Philips-Wren, 2009; Becerra-Fernandez et al., 2008) and action dependent (Feigh & Pritchett, 2006), which meet the characteristics of the emergency situations.

The decision on reservoir water release is crucial especially on a reservoir that is built as a multi - purpose reservoir. In addition, as a flood defence mechanism, the reservoir may serve water for other purposes such as irrigation, water supply, and recreation. Therefore, the release decision needs to consider adequate reservoir water balance in order to maintain its water supply and provide enough space for incoming inflow. The volume of water release should not exceed the safe carrying capacity of the downstream channel, so that it will not cause overflow. The exact amount of water to be released is hard to determine as the incoming inflow is fuzzy. Therefore, reservoir water release decision needs the ability to handle the fuzzy information.

Typically, reservoir water release decision is based on upstream inflow that is observed through the magnitude of the upstream rainfall and the river water level. The total volume of rainfall may come from several gauging stations and their distances to the reservoir are varied. Thus, rainfall observed at those gauging stations may take different time period to reach into the reservoir. This situation shows that there is spatial temporal relationship hidden in between each gauging station and the reservoir. Currently, there are limited studies that focus on this situation, as most of the previous studies focus on temporal relationship and the total rainfall volume (Mohan & Revesz, 2012; Wan Ishak et al., 2011b, 2012). Thus, the spatial temporal relationship needs to be considered for modelling reservoir water release decision.

One of the well-known computational intelligence techniques used for modelling reservoir water release decision is Artificial Neural Network (ANN) (Wan Ishak et al., 2012; Wan Ishak et al., 2011a; Wan Ishak et al., 2011c; Wan Ishak et al., 2011d). However, this technique suffers from poor interpretability, since it is difficult for humans to explain the practicality and logical meaning behind the learned weights of the model (Wei & Hsu, 2008; Jothiprakash & Kote, 2011; Kajornrit et al., 2013). This problem can be solved by integrating ANN with other techniques, to form a hybrid model, which combines the strength of the techniques. In this study, a hybrid computational intelligence technique, namely Adaptive Neuro Fuzzy Inference System (ANFIS) (Jang, 1993; Chang & Chang, 2006; Valizadeh & El-Shafie, 2013) is adopted to model reservoir water release decision. ANFIS is the hybrid intelligent system which combines the learning capability of fuzzy logic and artificial neural

network. The construction of ANFIS model provides an explanation on the decisions made from the knowledge embedded in the trained neural network and resolves the issues of interpretability. The application of ANFIS has been successfully used in modelling reservoir operation, forecasting and prediction (Chang & Chang, 2006; Mehta & Jain, 2009; Deka & Chandramouli, 2009; Valizadeh & El-Shafie, 2013; Valizadeh et al.,2014). Nonetheless, most of the ANFIS model which has been proposed focuses on water level forecasting and prediction. In addition, the existing models do not consider the spatial temporal relationship that exists in modelling reservoir operation problems. Hence, this study is presented by considering both spatial temporal relationships in modelling reservoir water release decision using the ANFIS.

1.3 Research Questions

The following research questions will be investigated in order to develop a reservoir water release dynamic decision model based on spatial temporal patterns:

1. Is there any spatial temporal pattern of rainfall of each gauging station?
2. How to develop reservoir water release dynamic decision model based on spatial temporal patterns?

1.4 Research Objective

The main objective of this study is to develop a reservoir water release dynamic decision model based on spatial temporal patterns. This main objective is supported by the following objectives:

- i. To identify spatial temporal pattern of rainfall of each gauging station.
- ii. To develop ANFIS model for reservoir water release decision using spatial temporal patterns of rainfall and reservoir water level.

1.5 Research Scope

The scope of this study is to develop a reservoir water release dynamic decision model based on spatial temporal patterns. The spatial temporal pattern has been extracted from reservoir operational data using spatial temporal data mining techniques. This study focuses on multipurpose reservoir specifically flood related reservoirs, and Timah Tasoh reservoir was selected as a case study due to its purpose which serves as a flood mitigation approach. Reservoir operational data and hydrological data have been collected from the Timah Tasoh dam (1999-2012) and DID which contain information of daily reservoir operation and rainfall data from five gauging stations (Padang Besar, Tasoh, Kaki Bukit, Wang Kelian and Lubuk Sirih). These rainfall stations are chosen because the locations of these stations are closest to Timah Tasoh reservoir.

1.6 Research Significance

The main significance of the study is to propose a reservoir water release dynamic decision model for reservoir operation and management. In practice, this decision model will be able to enhance performances of the reservoir system in order to support accurate water release decisions during emergency situations. This decision model has been developed by adopting a hybrid computational intelligence technique, specifically Adaptive Neuro-Fuzzy Inference System (ANFIS). This ANFIS model is examined based on spatial temporal patterns which extracted using spatial temporal data mining technique from the reservoir operational data. A spatial temporal pattern extraction procedure has been determined in order to identify spatial temporal relationships that exist in modelling reservoir water release decision.

Finally, the other outcome of this study is that it cannot stop the hydrological events such as flood, heavy rainfall, or drought from occurring, but it will reduce the risk of the events. Therefore, this risk will be prevented by addressing the signs of flood in order to ensure that the reservoir capacity is at a safe level and water discharges will not trigger a downstream flood.

1.7 Organization of the Thesis

This chapter introduced the background of the study and the research problems that need to be solved which gives direction to this study. This thesis consists of six chapters. The general ideas of the content of the following chapters are described as follows:

Chapter 2 reviews the literatures on the related study areas includes types of decision making approaches, reservoir operation, reservoir water release decision, computational intelligence approaches in reservoir operation and spatial temporal data mining.

Chapter 3 describes the research methodology adopted in this study.

Chapter 4 discusses the experimental results on the spatial temporal data mining. The detail explanation on how to extract spatial temporal pattern has been discussed. A spatial temporal pattern extraction procedure adopted in this study is also presented.

Chapter 5 presents the findings and discussion on the development of reservoir water release dynamic decision model using Adaptive Neuro-Fuzzy Inference System (ANFIS).

Chapter 6 summarizes the thesis with contributions of this study and provides ideas of future work.

CHAPTER TWO

LITERATURE REVIEW

This chapter reviews the literatures of existing research related to decision making approaches, reservoir operation, reservoir water release decision, reservoir operating policy, computational intelligence approaches in reservoir operation and decision modelling, spatial temporal data mining and a summary of the chapter.

2.1 Decision Making in Emergency Situation

Decision making becomes one of the critical processes that require faster and accurate decision during emergency situations. Every decision making process produces a final choice of possible alternatives to provide a solution to overcome the problems. Decision making can be defined as a process of finding possible solutions by selecting a choice from a set of alternatives (Azuma et al., 2006; Hersh, 1999). Modelling decision-making is classified into two main approaches based on the classical or naturalistic decision theory (Hersh, 1999).

Previously, people make decisions based on classical decision theory (Hersh, 1999) that outlines how people should make the decision. This theory concentrates on the choice of an alternative or decision event from a fixed set (Hersh, 1999). A classical decision model is focused on the idea of reasonable choice and decision maker make decisions through choosing a solution among a set of available alternatives that are satisfied with the target (Hersh, 1999; Azuma et al., 2006). As an example of classical decision theory is a rational decision making model which decisions made based on the rationality of the decision maker (Li, 2008). However, classical

decision model assumes that decision makers have mental capacities for analysis and evaluation that vastly beyond those actually overcome by people (Club Managers Association of America (CMAA), 1999). These limitations of the classical decision model provide an attempt to the decision makers to introduce a new theory called naturalistic decision making (Wan Ishak et al., 2011c).

Naturalistic decision making has a different idea. Naturalistic decision theory is proposed to model the actions and behaviour of the decision makers (Klein & Klinger, 1991). These models can be defined as how humans actually make a decision in complex real world settings such as during floods, fire fighting and other emergency situations. This model believes that the best solution comes from the actions, from observing the results. It does not attempt to derive the optimal and ideal solution (Azuma et al. 2006). This allows decision makers to predict the action by imagining it in a particular situation, look forward to important steps, and considering the most likely reaction to it (Hersh, 1999).

Klein and Klinger (1991) presented naturalistic decision making that shows how people can make effective decisions without performing analyses. They have studied command and control decision making and have generated a recognized model of naturalistic decision making. According to Lipshitz et al. (2001), this study discussed four essential characteristics of NDM, which are process orientation, situation-action matching decision rules, context-bound in formal modelling and empirical-based prescription. Process orientation refers to the input-output orientation which describes the cognitive process of decision makers. Situation- action matching

decision rules indicates the decision makers matching the action with the appropriate situations rather than making a choice. Context-bound in formal modelling refers to the experience-tied knowledge of decision makers, meanwhile empirical-based prescription illustrates on how prescriptions are derived from descriptive models of expert performance.

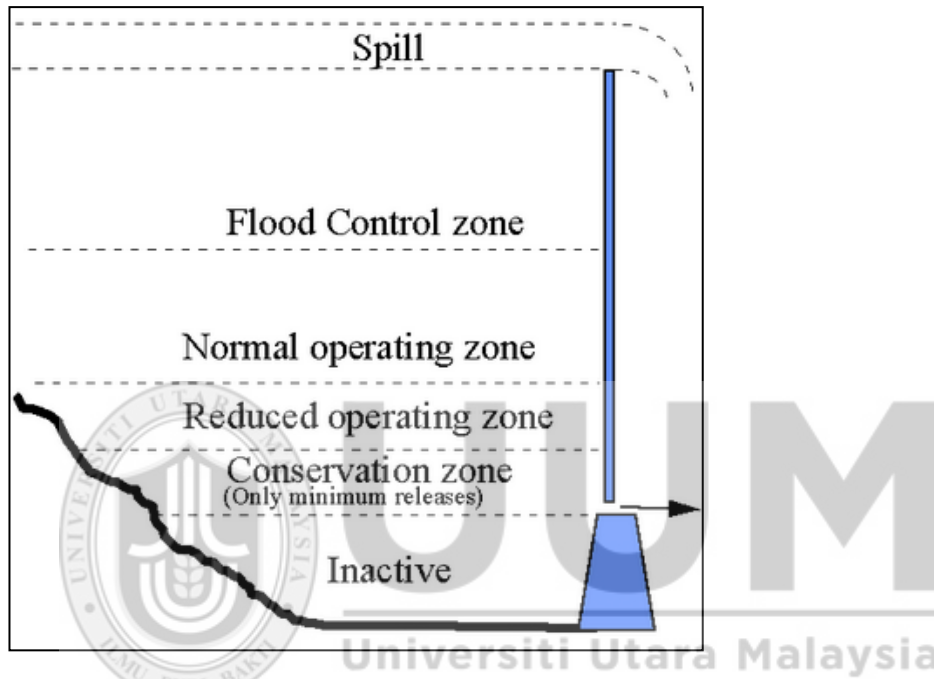
2.2 Modelling Reservoir Water Release Decision

Modelling reservoir operation is crucial in order to make decisions either to maintain or to release a reservoir capacity. Reservoir water release decision is one of the critical actions especially during emergency situations such as flood, heavy rainfall or drought. In practice, reservoir water release decision involves many decision variables and multi-purpose objectives that need to be considered. However, these release decisions are guided by a reservoir operating policy in order to determine accurate water release decision as well as to improve performances of the reservoir operation.

2.2.1 Reservoir Operation

Reservoir operation is a complex problem in water resource management which involves many hydrological variables and multi-purpose objectives due to unpredicted situations such as flood or drought. The reservoir is constructed based on the types of reservoir systems that consist of single and multi-reservoir systems (USACE, 1991). Furthermore, reservoirs can be classified according to their function and purposes. It can be classified into two types: single-purpose and multi-purpose reservoir. Generally, a single-purpose reservoir is operated for a particular operation

purpose only, such as for flood control or hydropower generation. Meanwhile, a multi-purpose reservoir is constructed for several operational purposes such as flood control, water supply, hydropower generation, recreation and irrigation. Figure 2.1 shows the example for classification of storage zones of the reservoir.

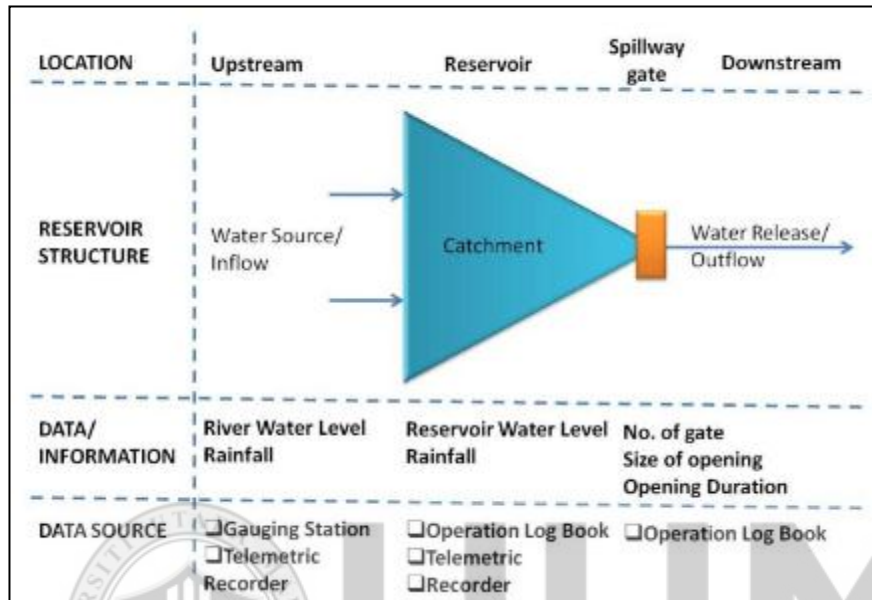


(Source: MIKE Basin, n.d)

Figure 2.1. Classifications of Storage Zones of the Reservoir

The operation and management of a reservoir system is a complicated problem due to the variations of demand for multi-purpose reservoir operations and multiple conflicting objectives. Reservoir operation involves many critical decisions which make it difficult for decision makers and reservoir operators. Oliveira and Loucks (1997) described that reservoir has a complex operational structure which involves many decision variables and multi-purpose operation as well as considerable risk and uncertainty that requires urgent decision making as a result of unpredicted

circumstances such as flood or high inflow event. Figure 2.2 shows the conceptual model of the reservoir system which consists of four components: upstream, reservoir catchment, spillway gate and downstream.



(Source: Wan Ishak et al., 2011e)

Figure 2.2. Conceptual Model of Reservoir System

In practice, reservoir operations use certain operating rules. However, each reservoir system is unique and has a different way to define its operating rules (Wan Ishak et al., 2011a). Operation policy can be defined as a series of guidelines or strategies generated for reservoir planners and operators to guide the release decisions (Rittima, 2009). Typically, operation rules are identified from the reservoir operation manual when it first operated (Wan Ishak et al., 2012). Conventionally, reservoir operation is operated based on the judgments, heuristic procedures or past experiences of the reservoir operator. This provides guidelines in making water release decision according to the hydrological conditions, current reservoir storage level, reservoir purpose etc. (Ngo, 2006).

2.2.2 Reservoir Water Release Decision

Reservoir water release decision is one of the challenging tasks for the reservoir operator since the decision deals with many complicated decision variables and multi-purpose operation. During flood and drought seasons, reservoir water release decisions are guided by the reservoir operating policy that helps reservoir operators to determine accurate time for water release decision. The release decision consists of actions to determine the amount of water to be retained and to be discharged from a reservoir under various conditions (Wurbs, 1993). Jain and Singh (2003) described that the decision making on water releases relies on the reservoir storage level, demands, inflows, time, previous release etc.

Gate operation is one of the main significant functions of the reservoir that must be considered for water release decisions. Gated spillways are designed as hydraulic controls that provide a unique relationship between the reservoir water level and the discharge (Afsar and Salehi, 2011). The decision on the gate opening is made based on the reservoir water level which changes over time due to volume of rainfall. The decision to open the spillway gate is very critical because any mistake by the reservoir operator will cause downstream flood which gives impact to the human life and property (Hassin et al., 2006).

The operation of gate opening involves typical information that needs to be considered before making any decisions regarding water release, such as the number of the gate opening, size of the opening, and duration of the opening (Wan Ishak et al., 2011a). The amount of water that can be discharged through spillway gates

depends on the reservoir operating policies. Thus, the spillway gate must be regulated according to hydrological conditions and needs to be controlled in order to support the reservoir operation.

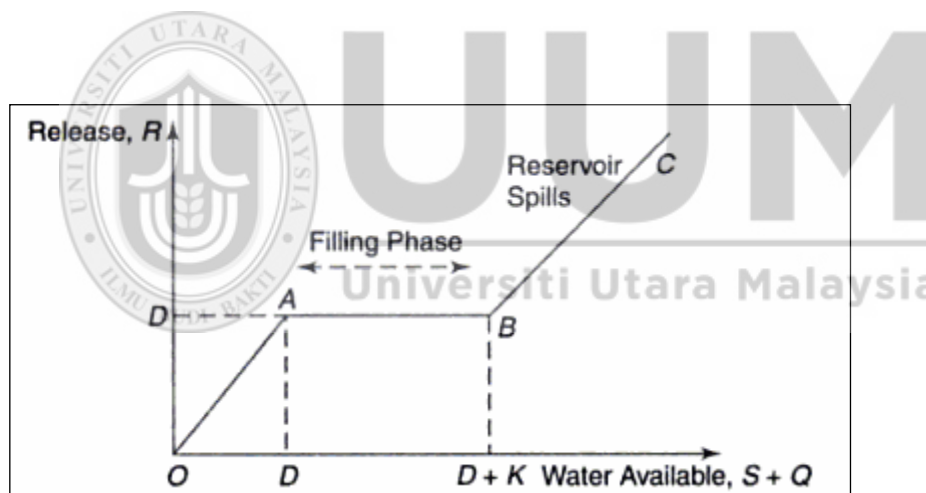
The reservoir operator needs to determine the right time before making a decision to release water through the spillway gate to the downstream. In practice, experts use their experience or knowledge to make a decision either to open or close the gate, but they need to develop the skills by understanding the basic structure and the current system of the reservoir operation before using that knowledge (Hassin et al., 2006).

2.2.3 Reservoir Operating Policy

Reservoir operating policy is one of the most important procedures in the reservoir operation. These policies are used to guide reservoir operators in making water release decisions. In general, reservoir operating rules are determined at the early stage when it's operated. However, different reservoirs require different operating rules (Ngo 2006; Husain 2012). There are numerous types of reservoir operating policies have been previously suggested and discussed:

i. Standard Operating Policy (SOP)

Standard operating policy (SOP) is designed based on an idea, which is to discharge as much water to fulfil the target delivery based on the capacity that can be provided by a reservoir (Tu et al. 2008). SOP is the simplest and most often-used reservoir policy that releases water as close to meet the demand as possible, and does not preserve water for future requirements. If sufficient water is not available to meet the demand, the reservoir is emptied (Neelakantan and Pundarikanthan, 1999). SOP is the optimal operating policy with an objective to minimize the total deficit over the time horizon. Figure 2.3 shows the Standard Operating Policies (SOP) that has been applied for reservoir operation and management.



(Source: Majumdar & Mujumdar, 2005)

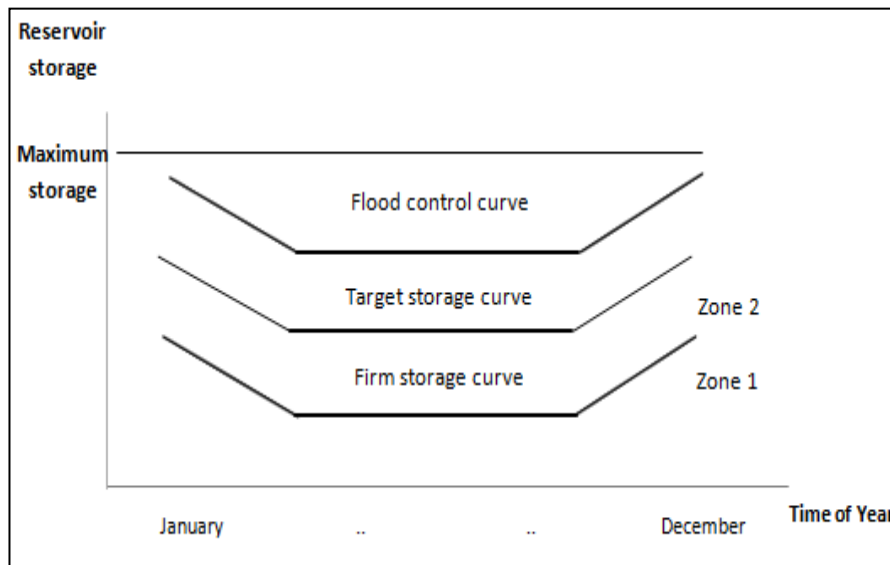
Figure 2.3. Standard Operating Policies (SOP)

ii. Rule Curves

Rule curves are one of the basic operating policies for managing and guiding reservoir operation. The term "rule curve" is used to refer to the elevations which define ideal (desirable or target) storage volumes and provides a mechanism for release rules to be specified as a function of storage content (USACE, 1991). However, this rule requires to be updated when drought occurs (Hoshimoto et al., 1982). These policies traditionally are derived through simulation techniques (Chang et al., 2005).

Chang et al. (2005) proposed operating rule curves for managing multi-purpose reservoir operation by using genetic algorithms. The two genetic algorithms (GAs) have been investigated to prove the effectiveness of both coded for optimizing the operating rule curves. The result indicates that genetic algorithms provide better performances for deriving a new rule curve in terms of water release compared to current M-5 operating rule curves.

According to Tu et al. (2003), this study described that reservoir rule curves consist of three storage curves, which are known as flood control curve (the upper curve), target storage curve (the middle curve), and firm storage curve (the lower curve). These curves are defined when the reservoirs start operation. Figure 2.4 shows the storage curves for multi-purpose reservoir. These rule curves are used as a trigger function of the beginning reservoir storage level to start hedging for managing multi-purpose reservoir operations.



(Source: Tu et al., 2003)

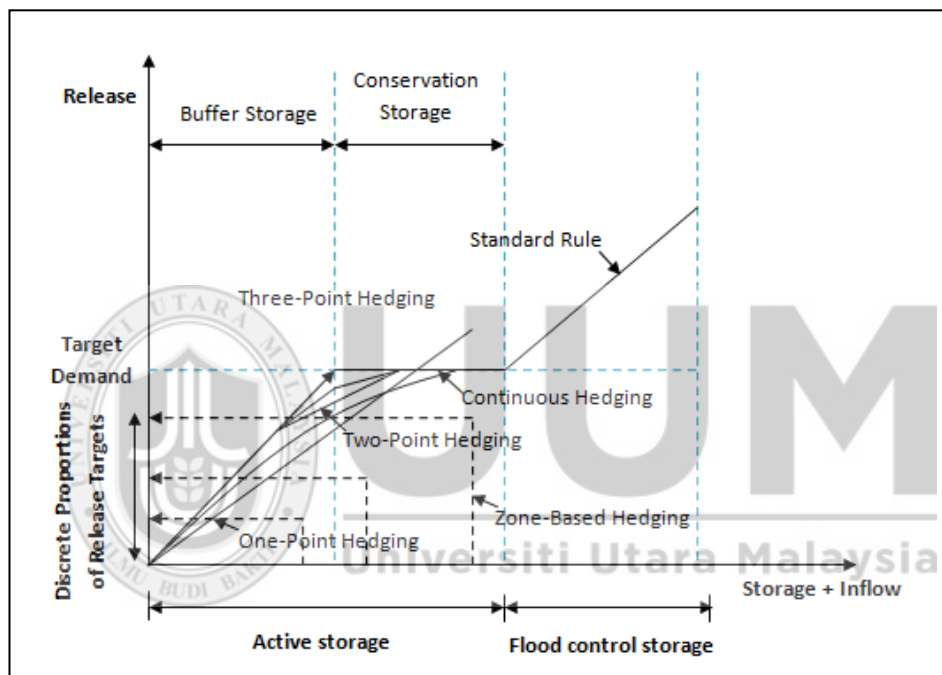
Figure 2.4. Rule Curves for Multi-Purpose Reservoir

iii. Hedging Rules

Hedging rule is one of the reservoir operating policies that are designed to trigger the water release and reserve water for future use (You, 2008). These policies allow a minimum deficit in the current delivery to reduce the probability of water shortage problems in the future. Based on this principle, hedging rules are used to discharge water based on the water demand and maintain it for the next use (Draper & Lund, 2004).

Many previous studies have been made by numerous researchers on the development of hedging rules for improving reservoir system operation (Tu et al., 2003, 2008; Draper & Lund, 2004; Shiau & Lee, 2005; You & Cai, 2007, 2008a, 2008b; Rittima, 2009; Shiau, 2009, You & Cai, 2009; Guo et al., 2012; Taghian et al., 2013). Table 2.1 summarizes some of the related studies to the application of hedging rules in

reservoir operation and management. Based on the literatures stated in Table 2.1, it can be concluded here that most of the studies applied the implementation of hedging rule for improving reservoir operation which has been done using a numerical/mathematical model. Draper and Lund (2004) presented the common forms of hedging policy which are described in the following manners as shown in Figure 2.5 (Rittima, 2009):



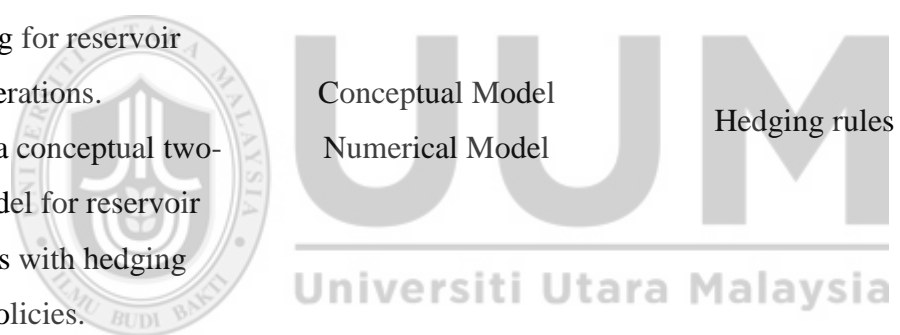
(Source: Rittima, 2009)

Figure 2.5. Types of Hedging Rules

Table 2.1

Application of Hedging Rules in Reservoir Operation

Authors/ Title	Objective	Model	Method	Suggestion/ Remarks
Tu et al. 2003	To develop a mixed integer linear programming model that considers both traditional rule curves and hedging rules to manage and operate multi-purpose, multi-reservoir systems.	Mathematical Model	MILP Branch-and-bound algorithm	The reservoir rule curves and the hedging rules provide operating guidelines that can be followed to minimize the impacts of a drought.
Tu et al. 2008	To devise new hedging rules for an existing multi-reservoir system	Optimization Model Mathematical Model	Transformation technique Mixed Integer Nonlinear Programming Model to Mixed-Integer Quadratic Programming Model	The reservoir hedging rules are recommended to be re-evaluated and updated to improve reservoir systems.
Rittima, 2009	To simulate reservoir operations using a hedging policy	Simulation Model	Simulink Toolbox in Matlab	Optimization techniques should be used in order to

				determine the most adequate hedging pattern for real-life reservoir systems.
You & Cai, 2008a	<p>This theoretical analysis will be used to improve numerical modelling for reservoir operations.</p> <p>To develop a conceptual two-period model for reservoir operations with hedging policies.</p>	 <p>Conceptual Model Numerical Model</p>	Hedging rules	<p>Because of the limitation of theoretical derivation, several analyses are left without an explicit result. For example: bounds of water availability for hedging.</p>
You & Cai, 2008b	<p>To develop a new numerical method that incorporates hedging rules derived from theoretical analysis.</p>	Numerical model	Hedging rules	<p>This study is limited to a two-period model. The predicted inflows are I.I.D with foresight limited to the next period.</p> <p>The model is not for</p>

intra-year reservoir operations, which may take account hedging season to season, month to month or with even shorter time interval.

The accuracy of future inflows has greater significant effects on HR II than HR I because it depends on future shortage conditions to determine hedging. However, the results are limited to the assumption of uniformly distributed monthly demand. Further studies are needed to explore the

To derive optimal hedging rules for minimizing short-term and long-term shortage characteristics for a water-supply reservoir.

Shiau and Lee, 2005



UUM

Universiti Utara Malaysia

Optimization Model

Mathematical

Hedging rule I

Hedging rule II

Compromise

Programming

				optimal hedging rules for the variable monthly demand.
Guo et al., 2012	To derive an operating rule for multi-reservoir systems by combining parametric rule with the hedging rule to avoid severe short supply during drought.	Optimization and Simulation Model	Particle swarm optimization	The proposed rule can make a radical reduction of decision variables for optimizing reservoir system.
Taghian et al., 2013	To improve reservoir performances by developing a hybrid model for optimizing both conventional rule curve and hedging rule.	Hybrid Model	Genetic Algorithms Linear Programming and Simulation Program	The result shows the hybrid model will improve the reservoir performances compared to applying the rule curve alone.

2.3 Computational Intelligence Approaches in Reservoir Operation and Decision Modelling

Computational intelligence (CI) can be defined as computational methodologies and approaches that are designed based on AI-based methods including artificial neural networks (ANN), genetic algorithms, fuzzy logic etc. In the last several years, the popularity of CI techniques has been increasing in various water resources problems, including reservoir operations, as these techniques can deal with many complexities compared to the use of traditional methods. Generally, CI techniques may be able to solve the problems of the uncertainty of hydrological data and the complexity of the physical characteristic of a reservoir system. The application of computational intelligence techniques has been reviewed to be related to the reservoir operation and decision modelling.

2.3.1 Artificial Neural Network

Artificial Neural Networks (ANN) is one of the popular artificial intelligence methods that have the ability to handle non-linear, dynamic and noisy data effectively. It represents a computational mechanism based on a simplified mathematical model of the perceptrons (neurons). Artificial Neural Network has been widely used for modelling reservoir operations especially in terms of deriving operating policies to guide reservoir operators in making water release decisions (Raman & Chandramouli, 1996; Neelakantan & Pundarikanthan, 2000; Cancelliere et al., 2002; Ramani Bai et al., 2007; Sharifi et al., 2005; Wan Ishak et al., 2011a; 2012). Raman and Chandramouli (1996) proposed reservoir operating policies by using neural network for improving the reservoir operation. The performances of Dynamic Programming (DP) model, Stochastic Dynamic Programming (SDP)

model, Standard Operating Policy (SOP) model and Multiple Linear Regression Procedure (DPR) model are compared, and it has been found that neural network based on dynamic programming algorithm provides better performance compared to other models.

According to Neelakantan and Pundarikanthan (2000), a simulation-optimization model based on neural network has been developed for water supply reservoir systems. This study involves three stages of general framework for developing a simulation-optimization model. The result of this study indicates this model performs well compared to the conventional simulation-optimization model. On the other hand, Ramani Bai et al. (2007) proposed a model based on neural network with four different multi-layer feed forward neural network (MLFFNN) modes for dam operation and control by representing the training set to the network in the following four different data forms, which are HANN, ANN, IANN, and EANN. The findings show that data representation based on the explicit method of using optimized values (EANN) yield better performance compared to the other representations. This study recommends that NN should be used with data mining rather than optimization alone.

According to Cancelliere et al. (2002), artificial neural networks with dynamic programming techniques were used for deriving irrigation reservoir operating rules. The results of this study show that the operating rules improve performances of the system operation in both normal and drought periods. On the other hand, Wan Ishak et al. (2011a; 2012) applied Neural Network to model reservoir water release decisions. Temporal data mining technique has been successfully used for extracting

temporal patterns from reservoir operational data. NN model was developed to classify the rules of reservoir water release. The finding of this study indicates the performance of NN depicted acceptable performance.

2.3.2 Fuzzy Logic

Fuzzy Logic is another method of artificial intelligence that has the ability to deal with uncertainty in the reservoir operation due to various hydrological condition, complex structure and regulating policy. The following studies have been reviewed on the application of fuzzy logic for modelling reservoir operations (Panigrahi & Mujumdar, 2000; Dubrovin et al., 2002).

Panigrahi and Mujumdar (2000) suggested fuzzy rule based system for improving single reservoir systems. The fuzzy rule based operation model is developed and compared with SDP operation for release decisions (Panigrahi & Mujumdar, 2000). The result of this study shows that the fuzzy rule based operation and SDP operation provide results in different values of the performance index by using two performance criteria: reliability and resiliency. This result is derived because of the different discretisation schemes and the way the two policies are used in operation. However, the application of this model may remain limited to the single reservoir system problems.

On the other hand, Dubrovin et al. (2002) presented a real-time fuzzy rule-based model for improving multi-purpose reservoir system. The application of a new method for fuzzy inference, total fuzzy similarity was applied and compared to a

traditional method. This model consists of two sub models which are reference water level model and release model. The results show that the model performs well but it did not consider expert judgment in the most unpredicted circumstances (Dubrovin et al., 2002). The several weaknesses of the model have been highlighted in this study, among them is the inability to observe the situations of forward and backward at several time steps.

2.3.3 Genetic Algorithm

Genetic algorithm becomes one of the popular techniques over other techniques because of its ability to overcome the complexity problems and search for optimum solution. The application of Genetic Algorithms (GA) has been used widely to derive reservoir operating policies for managing reservoir operations (Ahmed & Sarma, 2005; Jothiprakash & Shanthi, 2006; Momtahn & Darianie, 2007; Chen et al., 2007; Jothiprakash et al., 2011).

A Genetic Algorithm (GA) model was presented by Ahmed and Sarma (2005) for deriving an operating policy for multi-purpose reservoir. The GA model has been proposed to derive optimal operating policy, and the performances of the model are evaluated with stochastic dynamic programming (SDP). Momtahn and Darianie (2007) proposed the direct search approach for deriving optimal operational policies. Moreover, a GA optimization method was applied and compared to the conventional models concerning ESO and ISO approaches. The result shows that the proposed models performed better compared to ISO models in various forms of policies, in

which the optimization for the objective functions can be easily conducted with this approach (Momtahn & Darianie, 2007).

Jothiprakash and Shanthi (2006) proposed a genetic algorithm model for deriving the optimal operating policies for multi-purpose reservoir operations. The objective function has been used along with operating rule curves in order to minimize the annual sum of squared deviation of desired releases. The result shows the application of the model performs better in the real-time reservoir operation. Jothiprakash et al. (2011) have developed a genetic algorithm (GA) and a backward moving stochastic dynamic programming (SDP) model to derive optimal operating policies for a multi-reservoir system. This model has been developed to reduce the annual sum of squared deviation of desired target releases. As a result of this study, the GA provided better performances for finding operating policies for improving multi-reservoir systems as compared to SDP models.

2.3.4 Hybrid Computational Intelligence Model

In recent years, the limitations in the current optimization and simulation models to solve the complex problems in reservoir operations has led to the introduction of hybrid techniques for improving the current reservoir operation systems. The application of the hybrid techniques has been widely used for modelling complex reservoir operation and prediction

2.3.4.1 Adaptive Neuro-Fuzzy Inference System (ANFIS)

Fundamentally, ANFIS can be defined as a network model representation of the Sugeno-type fuzzy system whose membership function parameters are adjusted using either a back propagation algorithm alone, or in combination with a least square type of method (Jang, 1993). A first order Sugeno-type fuzzy model has rules as the following (Cruz & Mestrado, 2009):

Rule 1: If x is A_1 and y is B_1 , then $f_1 = p_1x + q_1y + r_1$

Rule 2: If x is A_2 and y is B_2 , then $f_2 = p_2x + q_2y + r_2$

Neuro- fuzzy systems have been developed as a result of the integration of two techniques: neural networks and fuzzy logic. The strengths of both neural network and fuzzy logic have been captured into a single framework by using neuro-fuzzy system (Mehta & Jain, 2009). The application of ANFIS has been used widely in prediction and modelling reservoir operation. Chang and Chang (2006) presented a neuro-fuzzy hybrid approach to develop a system for water level prediction during flood. Prediction models for reservoir management have been developed using the Adaptive Network Based Fuzzy Inference System (ANFIS). In this study, two ANFIS models have been developed: Model 1 with human decision as input variables, Model 2 without human decision as input variables. The result demonstrates that the application of ANFIS can be used successfully to forecast reservoir water level for the next three hours. However, this study focuses on prediction of reservoir water level and this might be significant different from modelling reservoir water release decision because of the variables involves of modelling technique.

The Neuro-fuzzy technique has been proposed by Mehta and Jain (2009) to develop an operation policy which is used for multi-purpose reservoirs and can be easily implemented by operators. Three models of Fuzzy Rule Based (FRB) have been developed and tested. Adaptive Neuro-Fuzzy Interactive System (ANFIS) and Fuzzy Mamdani (FM) have been used to compute the release from all developed models. The result shows that the ANFIS-cluster provides the best performances compared to FM, but FM is more user friendly.

On the other hand, the application of hybrid model was successfully applied to modelling reservoir operation during critical situations (Deka & Chandramouli, 2009). Deka and Chandramouli (2009) proposed the fuzzy neural network (FNN) application which combines the ability of two techniques: artificial neural networks and the fuzzy logic. In this study, the proposed modelling approach is initiated for finding a reservoir operating policy. The performance comparison of DP, DPNN and DPFNN were evaluated. The result of this study indicates that DPFNN model is more efficient in order to find a non-linear relationship among different variables. Nonetheless, the fuzzy rules do not give any explanation about the relationship between the high flow and other environmental variables such as rainfall, watershed and etc that are considered to be important.

Recently, Valizadeh and El-Shafie (2013) proposed multiple input fuzzifications in ANFIS to predict the dam level using different types of membership functions. The current model was compared with traditional ANFIS by fuzzifying each type of input in the modelling system to demonstrate the behaviour of the different membership

functions for each input. Three different time lags (R_{t0L1} , R_{t1L1} , and R_{t2L2}) were observed with two different MFs for the two inputs, rainfall and dam level from previous days. Based on three statistical evaluations, the model that employed different types of MF's performs better than other models, especially for time lag. Nevertheless, this study does not consider the spatial temporal relationship between dam level and rainfall distribution for water level forecasting.

Valizadeh et al. (2014) described a different ANFIS model in order to apply different types of MFs for each type of input for water level forecasting. This study was applied for two case studies: Klang Gates Dam and Rantau Panjang Station. Two different ANFIS models have been proposed to observe the performance of the new fuzzifying pattern in the ANFIS model with different types of membership function. Based on the results, the model which applied the generalized bell-shaped MF and the Gaussian MF provided better results in the two case studies compared to the model that used the generalized bell-shaped MF for the two types of inputs. Overall, this study proved the ability of the model for applying different types of membership function in the ANFIS model. However, this study focuses on reservoir water level forecasting and as the input variables involves is not considering the spatial temporal relationship exists for developing the ANFIS model.

2.3.4.2 Other Hybrid Computational Intelligence Model

Pinthong et al. (2009) presented a hybrid application which combines genetic algorithms and neurofuzzy system for improving water management system for multi-purpose reservoir operations. The results demonstrate the releases predicted by genetic algorithms and neurofuzzy model provide higher reliability for multi-purpose operations compared to actual operations.

Luo et al. (2015) proposed a hybrid multi-objective optimization (MO) approach named MO-PSO–EDA which combination of particle swarm optimization (PSO) algorithm and the estimation of distribution algorithm (EDA) for solving the MO – RFCO (Reservoir Flood Control Operation) problem. Three sub-populations and probability models have been developed to solve complex multi-objective optimization problem. The results illustrate MO-PSO–EDA model which combines two methods provides better performance compared with its two variations MO-PSO and MO-EDA. However, the study highlights the needs of comprehensive decision making by considering the accurate estimation of flood damages in the future.

2.3.5 Comparison of Computational Intelligence Techniques

The comparisons of computational intelligence techniques such as Artificial Neural Network (ANN), Fuzzy Logic (FL), Genetic Algorithm (GA) and Hybrid Computational Intelligence Model are summarized in Table 2.2. The applications of these techniques are compared according to their purpose, application, strength and weaknesses. To conclude here, a hybrid computational intelligence model provides better performances in terms of model development compared to stand-alone techniques.



Table 2.2

Comparison of Computational Intelligence Techniques

Techniques / Feature	Artificial Neural Network (ANN)	Fuzzy Logic (FL)	Genetic Algorithm (GA)	Hybrid CI Model
Purpose	To handle dynamic, non-linear and noisy data	To deal with uncertainty and inaccuracies	To overcome complexity problems and search for optimum solution	To overcome the limitations of stand-alone methods
Application	Forecasting Prediction Simulation	Expert System Embedded System	Optimization Simulation	Forecasting Optimization Simulation Expert System
Strength	Provides a novel solution to solve complex real-time problems (Dawson & Wilby, 2001). The development of the model does not require mathematical formulation (Chaves & Chang, 2008).	Simplicity and flexibility (Dubrovin & Turunen, 2002). Handle problems with impreciseness and uncertainty (Rani & Moreira, 2010).	A powerful capability to deal with a non-linear system global optimization (Chaves & Chang, 2008). GA provides ability to solve non-linear and	Combines the learning ability of the both techniques (Deka & Chandramouli, 2009). Enhance the reliability and the accuracy in terms of model development (Deka &

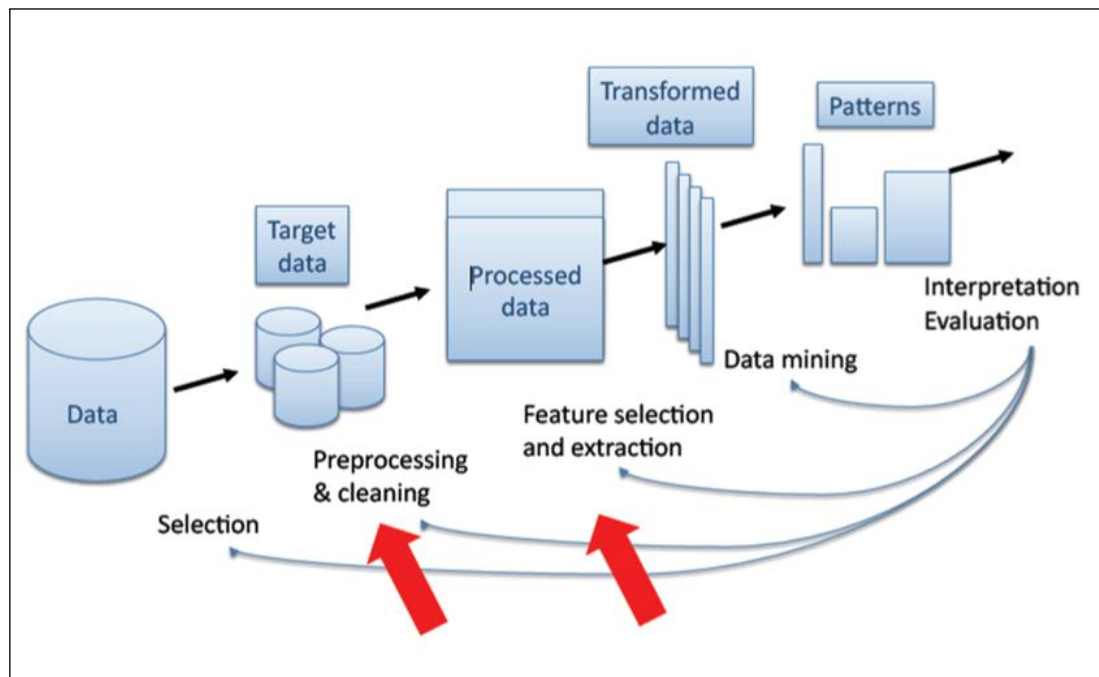
	ANN has ability to handle non-linear relationships of data efficiently (Galavi & Shui, 2012).		non-convex problems (Chaves & Chang, 2008).	Chandramouli, 2009).
Weaknesses	<p>Suffers from the black box problems (Deka & Chandramouli, 2005).</p> <p>Time consuming for training data (Chaves & Chang, 2008).</p> <p>Overfitting problem during NN training (Othman & Naseri, 2011).</p>	<p>Suffers from a lack of solid mathematical foundations (Dubrovin & Turunen, 2002).</p> <p>Suffers from a “curse of dimensionality” (Russell & Campbell, 1996).</p>	<p>The difficulty of explicitly accounting for constraints and maintaining feasible solutions in the population (Chaves & Chang, 2008).</p>	N/A

2.4 Spatial Temporal Data Mining in Reservoir Operation

Real time reservoir operation is a complex problem that involves many critical decisions and always changes due to unpredicted circumstances, such as climate change, extreme events, hydrological conditions etc. These situations are strongly related to hydrological variables where the values dynamically change through time and space. The concept of time explicitly refers to the temporal data where space represents spatial data and spatial relationship involved in the reservoir operation. Extracting the time and space information in the reservoir data can be done using data mining techniques.

Data mining is a multi-disciplinary field which involves various research areas including machine learning, database technology, statistics, information retrieval, neural networks, pattern recognition, artificial intelligence, high-performance computing, knowledge-based systems, and data visualization (Han et al., 2006). Data mining is a part of the overall process of KDD (Knowledge Discovery in Database).

In other words, it can be referred to a particular step in this process. In theory, data mining can be defined as a process for extracting useful patterns from data (Fayyad et al., 1996; Han et al., 2006). Figure 2.6 shows the steps in KDD process.



(Source: Fayyad et al., 1996)

Figure 2.6. Steps in KDD process

2.4.1 Temporal Data Mining

Temporal data mining is one of the popular topics in data mining research. In general, temporal data mining is an important extension of the data mining techniques and it can be defined as a process of extracting implicit, potentially useful patterns from the temporal database (Lin et al., 2002; Shekhar et al., 2010; Shahnawaz et al., 2011; Wan-Ishak et al., 2012). Basically, temporal data mining refers to the process for mining time series and sequential data sets. In practice, temporal data mining focuses on the analysis of temporal data for discovering temporal patterns from temporal data.

Temporal data can be classified into five categories that consists of static data, sequences, time stamped, time series and fully temporal (Roddick & Spiliopoulou,

2002; Shahnawaz et al., 2011). Static data refers to the data that has zero temporality where sequence data is events or transaction in a particular order. Time stamped, time series and fully temporal data refer to data that has time related information.

The application of temporal data mining in modelling reservoir operation has been successfully applied in previous studies by the following researchers (Wan Ishak et al., 2011b, 2012; Mohan & Revesz, 2012). According to Wan Ishak et al. (2011b, 2012), the application of temporal data mining and neural network has been proposed for modelling reservoir water release decision. Sliding window technique was used to segment temporal pattern from the reservoir operational data. This study demonstrates temporal data mining has been successfully applied in modelling reservoir water release decision. However, this study only focuses on temporal relationship of reservoir data which does not consider spatial relationship that exists in modelling reservoir water release decision. Mohan and Revesz (2012) presented the application of temporal data mining for handling reservoir data with uncertainty. The experiments have been done using different types of data classification methods in order to compare their performances in the case of uncertain data. The results of this study show the uncertain temporal data mining provided better performances compared to regular data mining. However, this study emphasizes a more comprehensive modelling of reservoir systems needs to consider with their spatial relationships.

2.4.2 Spatial Data Mining

On the other hand, spatial data mining is the other process to discover meaningful and interesting information from large spatial datasets (Shekar et al., 2005; Geetha et al., 2008; Shekar et al., 2010). The process of extracting patterns of spatial data sets is more difficult compared to traditional numeric and categorical data. In addition, spatial data analyses are crucial due to the complexity and characteristics of spatial data types, such as line, point, region, polygons etc. These types of data which are usually handled by geographic information system (GIS) are used for spatial data analysis.

Spatial data mining has become one of the popular study areas among the researchers in numerous applications. According to Wang (2012), geographic information system (GIS) and data mining methods have been applied for reservoir characterization and horizontal well-placement guidance acquisition. This study has been proposed to identify reservoir properties with spatial correlation. On the other hand, Manikandan and Srinivasan (2013) presented spatial data mining and Prim's algorithm for mining spatially co-located moving objects. The R-TREE for mining the patterns of spatial co-location has been applied for minimizing computation time. The results of this study show the proposed technique performs better in terms of memory space and time.

2.4.3 Spatial Temporal Data Mining

Spatial temporal data mining, or also known as spatio-temporal data mining, is an integration between spatial and temporal data mining that focuses on finding spatial and temporal relationships or other patterns from spatio-temporal data (Kalyani & Chaturvedi, 2012). It also can be considered as activities that integrate both spatial and temporal concepts. Spatial temporal data mining defines both spatial and temporal relationships. Two important attributes of spatial temporal data mining are location and time, in which is location refers to spatial relationship and time refers to temporal relationship of data.

In addition, the relationships of spatial and temporal are implicitly defined, that is, they are not explicitly encoded in a database. These relations must be extracted from the data and it is necessary to discover meaningful patterns in the data sets. The process of extracting the interesting patterns from spatial temporal database is more complex in order to consider both spatial and temporal relations. According to Rashid and Hossain (2012), they described the challenging issues of spatio-temporal data mining. The ability for analyzing the huge amount of data is still inadequate and is needed to adapt data mining tools. The application of spatio-temporal data mining has been widely discussed in reservoir operation and management. The need for considering both spatial and temporal relationships has been highlighted in these previous works (Chang & Chang, 2006; Mohan & Revesz, 2012).

2.5 Summary

Decision making in reservoir operation is a very crucial especially in emergency situations. Naturalistic decision theory is one of the decision making approaches that illustrates how people made a decision based on their instances or past experiences. This theory illustrates on how NDM take place in the emergency situation in order to support decision making in the reservoir operation. Decision making regarding water release is one of the challenging tasks in reservoir operation since the decision involves many decision variables and multi-purpose operations. A literature on computational intelligence on modelling reservoir operation and management has also been reviewed.



UUM
Universiti Utara Malaysia

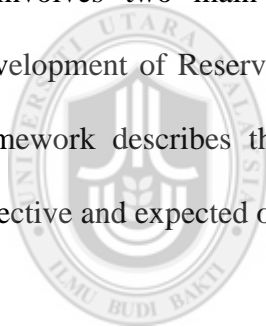
CHAPTER THREE

RESEARCH FRAMEWORK AND METHODOLOGY

This chapter elaborates the research framework and research methodology adopted in this study. This research methodology elaborates the sequence of activities and expected outcomes of each phase.

3.1 Research Framework

In this section, the research framework which has been implemented in this study will be discussed. Figure 3.1 illustrate the research framework adopted in this study. It involves two main phases, which are Spatial Temporal Data Mining and Development of Reservoir Water Release Dynamic Decision Model. This research framework describes the method or techniques used for achieving the research objective and expected outcomes of each phase as shown in Figure 3.1.



UUM
Universiti Utara Malaysia

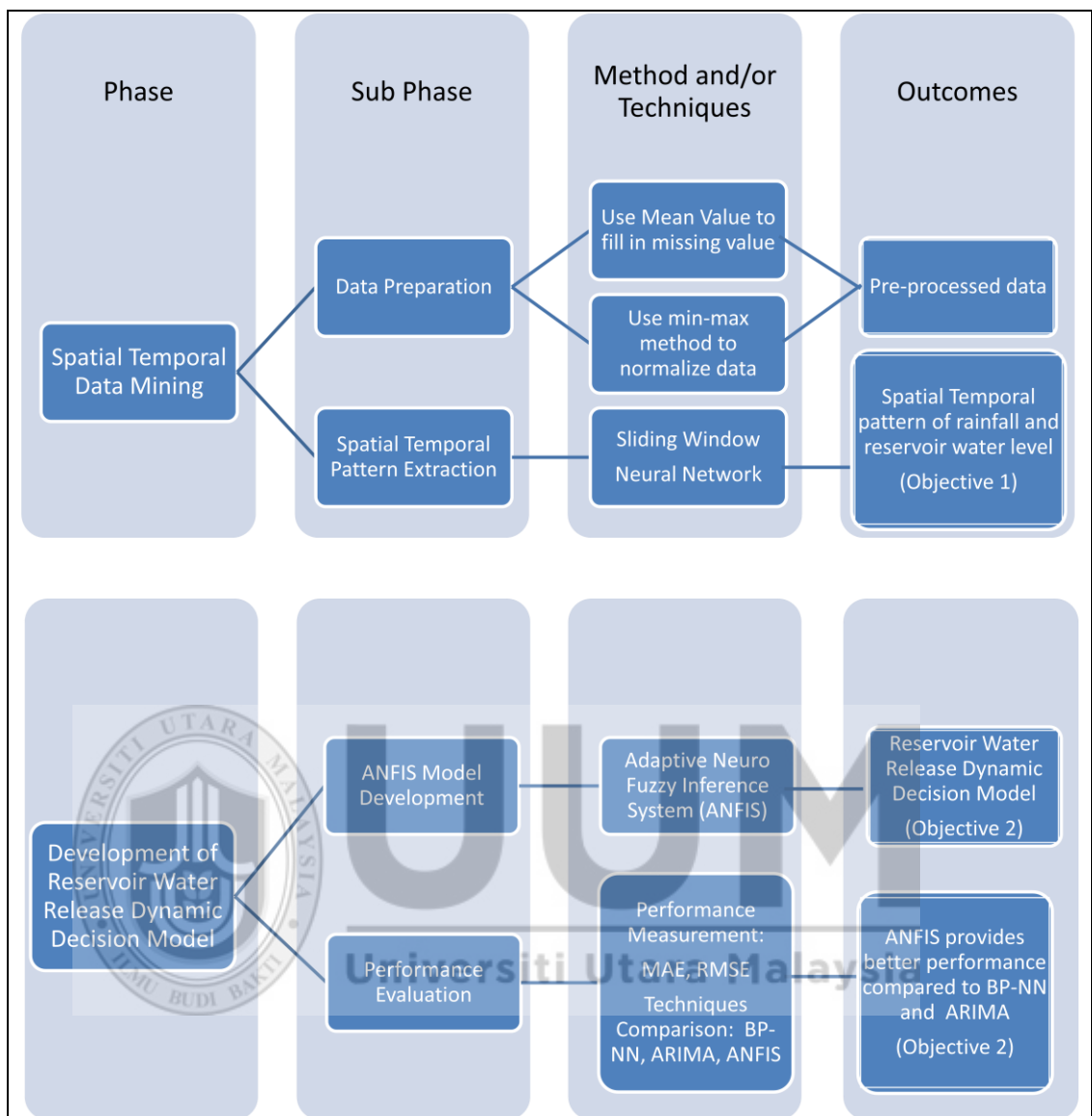


Figure 3.1. Research Framework

3.2 Research Methodology

In this study, the research methodology adopted in this study will be discussed in detail correspondingly in the subsequent sections.

3.2.1 Spatial Temporal Data Mining

This phase involves two major activities, which are data preparation and spatial temporal pattern extraction. These activities will be elaborated in detail in the subsequent sections:

i) Data Preparation

In this study, the Timah Tasoh reservoir was chosen and used as a case study that is located in the state of Perlis, Malaysia. The Timah Tasoh reservoir is one of the largest multi-purpose reservoirs in Northern Peninsular Malaysia. The Timah Tasoh reservoir serves as flood mitigation and is the only reservoir that has a gate structure that involves decisions. The data used has been retrieved from Timah Tasoh operational data from 1999 until 2012.

The data have been obtained from two sources which are from reservoir operational logbook and Department of Irrigation and Drainage (DID). Reservoir operational logbook or also known as operational data has information of daily reservoir operation such as reservoir water level and gate operation. Meanwhile, data which has been retrieved from DID or also known as hydrological data consist of hydrological information's such as rainfall volume, river water level and evaporation rate. The attributes selected as the input and

output variables which are reservoir water level, rainfall from 5 gauging station, and gate opening event. Table 3.1 shows the five categories of rainfall measurement. These five categories of rainfall measurement are applied to measure the volume of rainfall from five different gauging stations. Table 3.2 shows the four categories of flood stage which are used to measure the water level at the Timah Tasoh reservoir. Both categories of rainfall measurement and flood stage follow the standard measurement from DID.

Table 3.1

Five Categories of Rainfall Measurement

Min	Max	Code Number	Code	Stage
0.00	0.00	0	N	None
1.00	10.00	1	L	Light
11.00	30.00	2	M	Moderate
31.00	60.00	3	H	Heavy
61.00	200.00	4	VH	Very Heavy

Table 3.2

Four Categories of Flood Stage

Min	Max	Code Number	Code	Stage
0.00	28.99	0	N	Normal
29.00	29.39	1	A	Alert
29.40	29.59	2	W	Warning
29.60	30.00	3	D	Danger

In this study, the selected data such as reservoir water level, rainfall from five gauging stations and gate opening event has been retrieved and computerized into a Microsoft Excel file. The computerized data has undergone data pre-processing that involves several major tasks such as data cleaning and data

transformation as shown in Figure 3.2. Table 3.3 shows an example of the reservoir operational data.

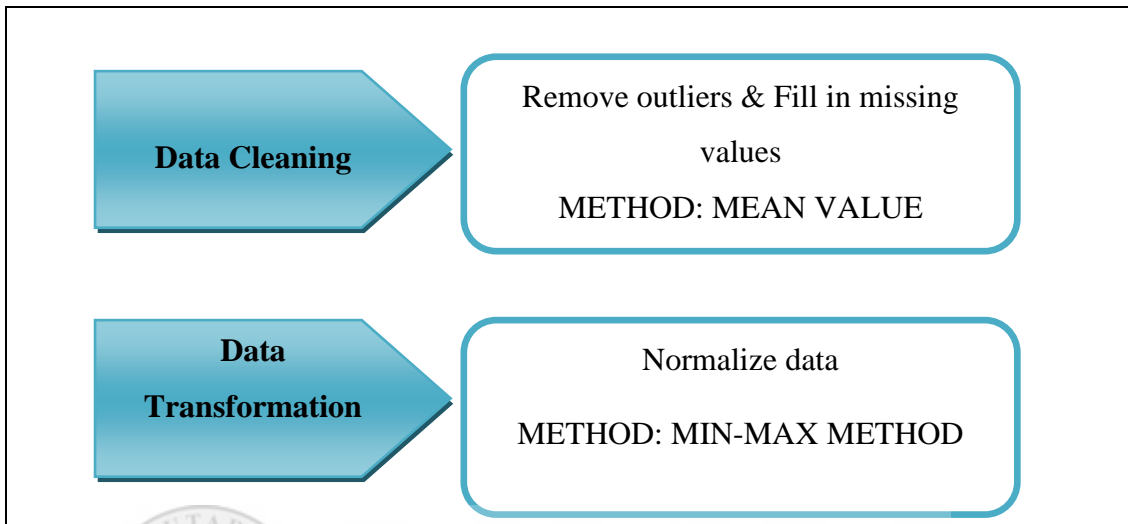


Figure 3.2. Data Pre-processing



Table 3.3

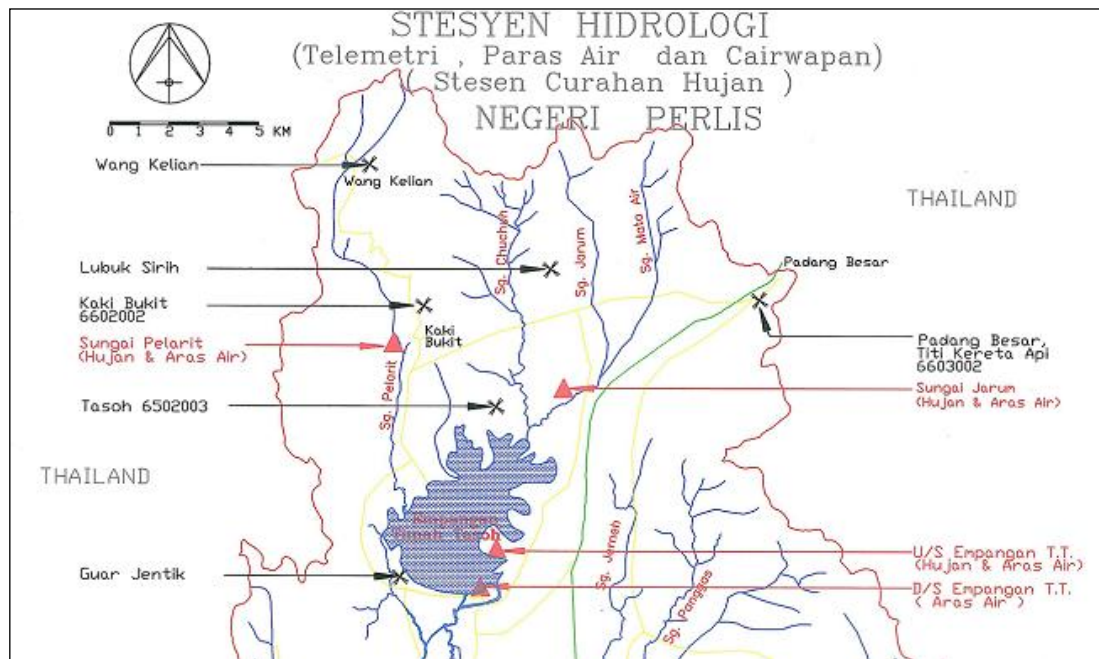
Example of Reservoir Operational Data

PARAS AIR (m)	CURAHAN HUJAN (mm)						LUAHAN MELALUI SPILLWAY				LUAHAN MELALUI OUTLET	
	Padang Besar	Tasoh	Lubuk * Sireh	Kaki Bukit	Wang * Kelian	Guar Jentik	Jumlah gate di buka	Level gate di buka (mm)	Discharge In meter 3/s	Jumlah (MCM)	Discharge In meter 3/s	Jumlah (MCM)
29.250	0.00	3.00	6.00	0.00	0.00	0.00	0	0	0	0	3.796	0.328
29.250	0.00	0.00	0.00	0.00	15.00	14.00	0	0	0	0	3.796	0.328
29.255	0.00	0.00	0.00	10.00	4.50	1.00	0	0	0	0	3.796	0.328
29.240	7.50	0.00	0.00	9.00	0.00	0.00	0	0	0	0	3.796	0.328
29.225	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	3.796	0.328
29.205	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	3.796	0.328
29.190	0.00	0.00	0.00	20.00	0.00	0.00	0	0	0	0	3.796	0.328
29.170	0.00	0.00	0.00	0.00	0.00	0.50	0	0	0	0	3.796	0.328
29.140	0.00	0.00	0.00	0.00	65.50	0.00	0	0	0	0	3.796	0.328
29.115	0.00	0.00	0.00	0.00	0.00	8.00	0	0	0	0	3.796	0.328
29.100	10.00	4.50	38.00	6.00	58.50	41.00	0	0	0	0	3.796	0.328
29.155	0.00	0.00	0.00	0.00	1.50	0.50	0	0	0	0	3.796	0.328
29.165	0.00	0.00	0.00	7.00	0.00	0.00	0	0	0	0	3.796	0.328
29.145	0.00	0.00	3.00	0.00	9.50	8.50	0	0	0	0	3.796	0.328
29.150	0.00	5.00	0.00	10.00	0.00	7.50	0	0	0	0	3.796	0.328
29.135	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	3.796	0.328
29.125	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	3.796	0.328
29.105	10.00	0.00	14.00	6.00	4.50	35.00	0	0	0	0	3.796	0.328

ii) Spatial Temporal Pattern Extraction

The second activity of this phase is to extract spatial temporal pattern of reservoir water release decision by using the sliding window technique and neural network (NN). A sliding window technique is selected because of its ability to capture the time delay within the dataset (Ku-Mahamud et al., 2009; Wan Ishak et al., 2011a; Wan Ishak et al., 2011b). In this phase, a few datasets have been formed based on different sliding window sizes. The window size is defined by a fixed period of time. Each of the sliding window size represents time of the delays. The best size of sliding window has been determined in order to obtain the best pattern of reservoir's upstream rainfall.

As mentioned earlier, the Timah Tasoh reservoir is influenced by upstream rainfall which is automatically recorded through five upstream gauging stations. Figure 3.3 shows the locations of the TIMAH TASOH reservoir and five gauging stations: PADANG BESAR, TASOH, LUBUK SIREH, KAKI BUKIT and WANG KELIAN. The temporal pattern of rainfall from these five gauging station has been trained using Neural Network for discovering spatial information exists between the reservoir and gauging stations. This spatial information represents different time delays for water to flow into the reservoir from the five different gauging stations.



(Source: Department of Irrigation and Drainage, 2012)

Figure 3.3. Location of Timah Tasoh Reservoir and Five Gauging Stations

For the NN models, input and output has been formed based on sliding window technique begins with size 2 that represent 2 days of delay. The maximum window size was set to 6 which represent 6 days of delay. Rainfall at the time t , $t-i$, $t-1$, ..., $t-w$ were used as the input data, while the current reservoir water level (WL_t) is used as the target. The constant t and w symbolize time and days of delay. Table 3.4 shows the example of input and output of NN model which formed using window size 3 that represent 3 days of delay. Input of all model are normalized using min-max method and the output was represented based on Binary-Coded-Decimal (BCD).

Table 3.4

Example Input and Output of NN Model

Input		Output	
rf t-2	rf t-1	rf t	WL t
-0.5	-0.5	-1.0	1.00
-1.0	0.0	0.0	1.00
1.0	0.0	-1.0	1.00
0.0	-1.0	-1.0	1.00
-0.5	0.0	0.5	1.00
-0.5	-0.5	0.5	1.00
0.0	-0.5	-1.0	1.00
-0.5	-1.0	-1.0	1.00
0.5	1.0	1.0	1.00
1.0	1.0	-1.0	1.00
1.0	-1.0	-1.0	1.00
-1.0	-1.0	-0.5	1.00
1.0	0.5	-1.0	1.00
0.5	-1.0	-1.0	1.00
0.5	0.5	-1.0	1.00
-1.0	-1.0	-1.0	1.00

For the ANFIS model, the input and output of the model has been formed based on the findings from spatial temporal pattern extraction. The spatial temporal patterns of rainfall from five gauging stations are used as the input variables and the status of gate opening at time $t+i$ ($Go t+i$) is used as the output variables. Figure 3.4 shows the example of windows formed with the real operational data for ANFIS model. The decision point is when there is a change in value of decision variables. In this study, the gate opening events either open or close gate (1 or 0) is used as a decision point. Zero (0) indicates gate is closed and one (1) represent the gate is opened.

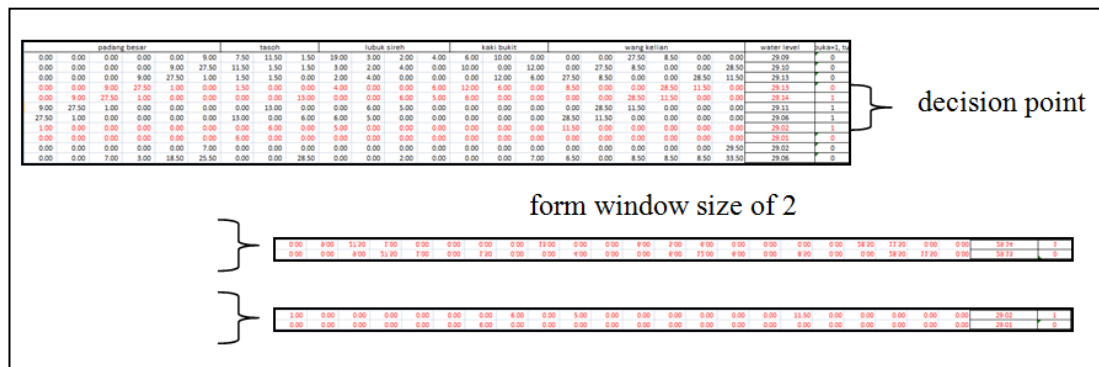


Figure 3.4. Example of Window Slice of 2

3.2.2 Development of Reservoir Water Release Dynamic Decision Model

In this study, Adaptive Neuro-Fuzzy Inference System (ANFIS) (Jang, 1993; Chang & Chang, 2006; Valizadeh & El-Shafie, 2013) is used to develop a reservoir water release dynamic decision model. ANFIS is preferred because it combines the learning ability of both techniques: neural network and fuzzy logic. The data in ANFIS are usually divided into three sets: training set, checking set, and testing set. The training data are used for the training of ANFIS, while the checking data are used for verifying the identified ANFIS which prevents over-fitting networks. The testing data are used to evaluate the model performance.

The ANFIS models were developed to create a fuzzy inference system in order to determine reservoir water release decision based on the given input-output patterns. Since the input variables used in this study are quite large: 23 and 22, which can generate different possible conditions, the basic structure of ANFIS model becomes expandable. In order to solve this problem, the subtractive clustering has been applied to establish the rule base relationship between the input and output variables.

The subtractive clustering usually can automatically verify the number of clustering for the data sets. It considers each data point is a potential cluster centre and calculates a measure of the probability that each data point would define the cluster centre, based on the density of surrounding data points. The density index D_i for each data point x_i is defined by:

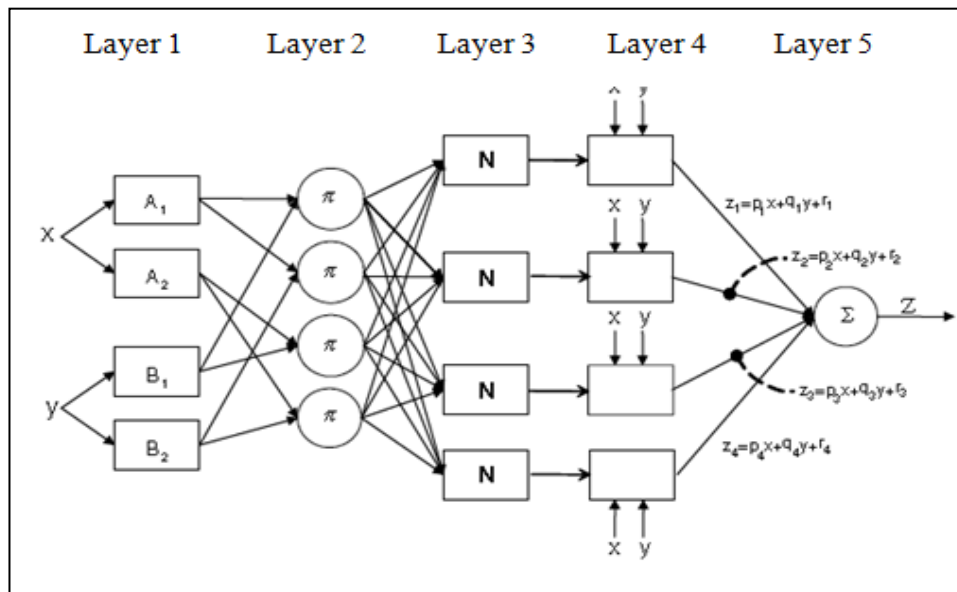
$$D_i = \sum_{j=1}^n \exp \left\{ - \frac{||x_i - x_j||^2}{(r_a/2)^2} \right\} \quad (3.1)$$

Where the value of r_a radius represents a positive constant which defines a neighbourhood of cluster centre. In order to avoid neighbourhood points of the first cluster centre being selected as the second centre, the density measure of each data point x_i is revised as:

$$D_i = D_i - D_{C1} \exp \left\{ - \frac{||x_i - x_j||^2}{(r_b/2)^2} \right\} \quad (3.2)$$

In general, r_b is a positive constant where the default value of setting r_b is set equal to $1.5 r_a$. The density measure will be modified in the remaining point set which is applied in each step. The process of determining the cluster centre will be terminated until the stop conditions have been reached. After several numbers of clustering are evaluated, each model adopts the appropriate number of rules. It appears that the subtractive fuzzy clustering does significantly reduce the number of rules.

Figure 3.5 shows the ANFIS Architecture for two inputs of Sugeno fuzzy model with four rules. This architecture consists of five layers that represent layer 1- input nodes, layer 2- rules nodes, layer 3- average nodes, layer 4- consequent nodes and layer 5- output nodes (Wang et al., 2009).



(Source: Chang & Chang, 2006)

Figure 3.5. ANFIS Architecture

The Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are recommended as performance evaluation criteria which have been used in this study. Figure 3.6 shows the process development of reservoir water release dynamic decision model.

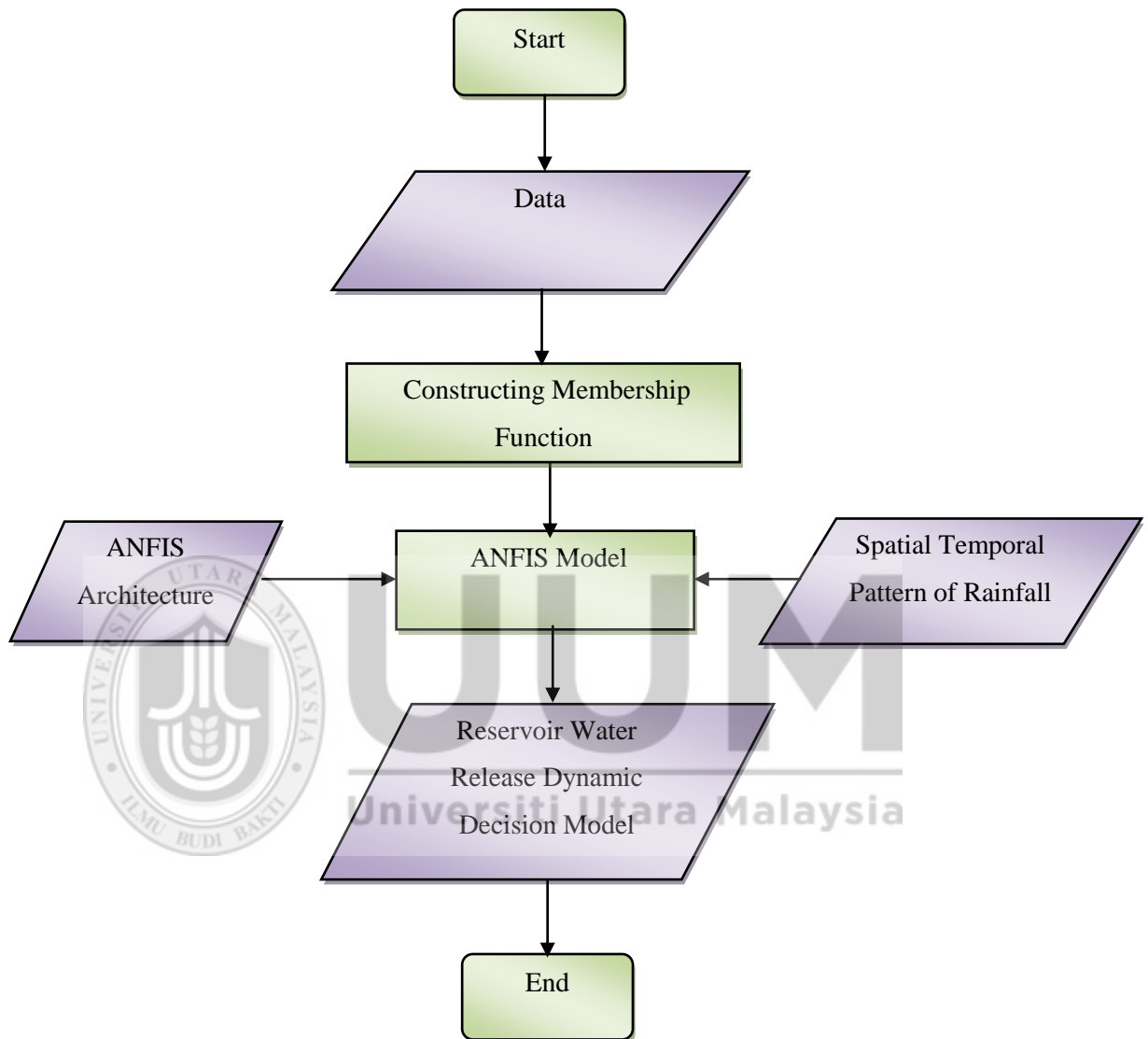


Figure 3.6. Process Flow for Development of Reservoir Water Release Dynamic Decision Model

i) Performance Evaluation

To assess the model's performances, the criteria used are follows:

1. Root Mean Square Error (RMSE) is used to measure the differences between values predicted by a model and the observed values that are being modelled.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obsi} - X_{model,i})^2}{n}} \quad (3.3)$$

2. Mean Absolute Error (MAE) is used to measure the average of the absolute error and if the lower value is better.

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (3.4)$$

3.3 Summary

In this chapter, the research methodology was designed and implemented for experimentation purposes. Two main phases were applied in order to develop a reservoir water release dynamic decision model. The performance of the decision model was measured by using two performance criteria: RMSE and MAE. The next chapter will discuss the steps for extracting spatial temporal patterns using spatial temporal data mining technique.

CHAPTER FOUR

SPATIAL TEMPORAL DATA MINING

This chapter discusses in detail on how to extract spatial temporal patterns from reservoir operational data. This chapter also provides an explanation on how spatial and temporal relationships are related in modelling reservoir water release decision.

4.1 Spatial Temporal Data Mining

Spatial temporal data mining is a process for finding spatial and temporal relationships or other patterns from spatio-temporal data (Kalyani & Chaturvedi, 2012). In this study, data mining techniques specifically sliding window and neural network have been used to identify spatial temporal patterns from reservoir operational data.

4.1.1 Data Pre-processing

As mentioned in Chapter 3, data pre-processing involves several major tasks such as data cleaning and data transformation. First, the data that has been collected from Timah Tasoh Dam and DID must undergo the cleaning process. Cleaning process is vital for detecting missing data, noisy data and outliers. In this study, missing data is replaced with the Mean value. Table 4.1 shows the mean value for each category of the flood stage: Normal, Alert, Warning and Danger. The missing data have been filled with mean value based on the flood stage. The following formula is used to find the mean value for each stage.

$$N = AVERAGEIFS(B:B, B:B, "≥0.00", B:B, "≤28.99") \quad (4.1)$$

$$N = AVERAGEIFS(B:B, B:B, "≥29.00", B:B, "≤29.39") \quad (4.2)$$

$$N = AVERAGEIFS(B:B, B:B, "≥29.40", B:B, "≤29.59") \quad (4.3)$$

$$N = AVERAGEIFS(B:B, B:B, "≥29.00", B:B, "≤30.00") \quad (4.4)$$

Where *B* refer to the name of the column

Table 4.1

Mean Value for Flood Stage

Min	Max	Stage	Mean
0.00	28.99	Normal	28.43
29.00	29.39	Alert	29.13
29.40	29.59	Warning	29.46
29.60	30.00	Danger	29.72

Table 4.2 shows the example on how to fill in the missing value with the mean value. For example, if there are missing values which value for the previous or the next data are classified into normal stage then the mean value of 28.43 is used to fill in the missing value.

Table 4.2

Fill in Missing Value

TARIKH	PARAS AI	Tasoh	Padang Be	Lubuk Sir	Kaki Buk	Wang Kel	Gate Open
01-Jan-07	28.550	0.00	0.00	0.00	0.00	0.00	0
02-Jan-07	28.530	0.00	0.00	0.00	0.00	0.00	0
03-Jan-07	28.510	0.00	0.00	0.00	0.00	0.00	0
04-Jan-07	28.505	0.00	0.00	0.00	0.00	0.00	0
05-Jan-07	28.500	0.00	4.00	0.00	0.00	2.00	0
06-Jan-07		2.00	1.00	14.00	1.00	0.00	0
07-Jan-07	28.490	59.00	91.00	89.00	77.00	83.00	0
08-Jan-07	28.590	10.00	13.00	5.00	12.00	30.00	0
09-Jan-07	28.735	0.00	8.00	0.00	2.00	4.00	0
10-Jan-07		0.00	0.00	0.00	0.00	0.00	0
11-Jan-07	28.745	0.00	0.00	0.00	0.00	0.00	0
12-Jan-07	28.750	1.00	0.00	0.00	0.00	0.00	0

Fill in Missing Value
with Mean Value

The values of data sets were normalized by using min-max method (Equation 4.5). Min-max method transforms the value of x into a new range [$new_min(A)$, $new_max(A)$]. It transforms the value x into a range between $[-1, 1]$ as shown in Table 4.3 and Table 4.4. Min- max normalization is recommended as the relationship of the original data values can be maintained (Han et al., 2006).

$$x = \frac{x - \min(A)}{\max(A) - \min(A)} * (new_max(A) - new_min(A)) + new_min(A) \quad (4.5)$$

Table 4.3

Normalize Value for Rainfall Measurement

Min	Max	Code	Stage	Normalize Value (x)
0.00	0.00	N	None	-1
1.00	10.00	L	Light	-0.5
11.00	30.00	M	Moderate	0
31.00	60.00	H	Heavy	0.5
61.00	200.00	VH	Very Heavy	1

Table 4.4

Normalize Value for Flood Stage

Min	Max	Code	Stage	Normalize Value (x)
0.00	28.99	N	Normal	-1
29.00	29.39	A	Alert	-0.33
29.40	29.59	W	Warning	0.33
29.60	30.99	D	Danger	1

Table 4.5 shows the example of data normalization using min-max method. The real value has been transform into a new value based on the rainfall measurement and flood stage.

Table 4.5

Data Normalization Using Min-Max Method

tarikh	water level	normalize	Kaki bukit	normalize	Jumlah gate
36161.00	28.53	-1.00	0.00	-1.0	0
36162.00	28.53	-1.00	12.00	0.0	0
36163.00	28.53	-1.00	0.00	-1.0	0
36164.00	28.70	-1.00	0.00	-1.0	0
36165.00	28.70	-1.00	0.00	-1.0	0
36166.00	28.71	-1.00	6.00	-0.5	0
36167.00	28.71	-1.00	0.00	-1.0	0
36168.00	28.70	-1.00	0.00	-1.0	0
36169.00	28.70	-1.00	7.00	-0.5	0
36170.00	28.69	-1.00	0.00	-1.0	0
36171.00	28.68	-1.00	0.00	-1.0	0
36172.00	28.67	-1.00	0.00	-1.0	0
36173.00	28.67	-1.00	5.00	-0.5	0
36174.00	28.67	-1.00	0.00	-1.0	0
36175.00	28.66	-1.00	0.00	-1.0	0
36176.00	28.66	-1.00	0.00	-1.0	0
36177.00	28.66	-1.00	0.00	-1.0	0

The output of each class is represented by using the Binary Coded-Decimal (BCD) scheme where the value “0” has been represented by “-1” so that all output is in the range [-1,1] as shown in Table 4.6. BCD is used as the total number of output nodes can be reduced to the integer $\log_2 M$, where M is the number of classes (Chong & Jia, 1994).

Table 4.6

Output Class Using BCD

Output Class	BCD Representation	
0	-1	-1
1	-1	1
2	1	-1
3	1	1

4.1.2 Spatial Temporal Pattern Formation using Sliding Window

Spatial and temporal patterns of reservoir water release decision need to be extracted from the past reservoir operational data which contains heuristic knowledge of reservoir operation. Spatial information is identified by determining the time of delays for water to flow into the reservoir from the five different gauging stations which are Padang Besar, Tasoh, Lubuk Sireh, Kaki Bukit, and Wang Kelian.

The gauging stations are located at different locations and the distance between the gauging stations and the reservoir varies. Thus, there are spatial relationships between the reservoir and the gauging stations which need to be investigated in making accurate decisions regarding water release. In the process to identify the existence of spatial information in the temporal pattern of rainfall, the temporal data mining classification process is applied in this study as shown in Figure 4.1. This process shows the accuracy can be improved by considering the value until to a number of n time units back in time.

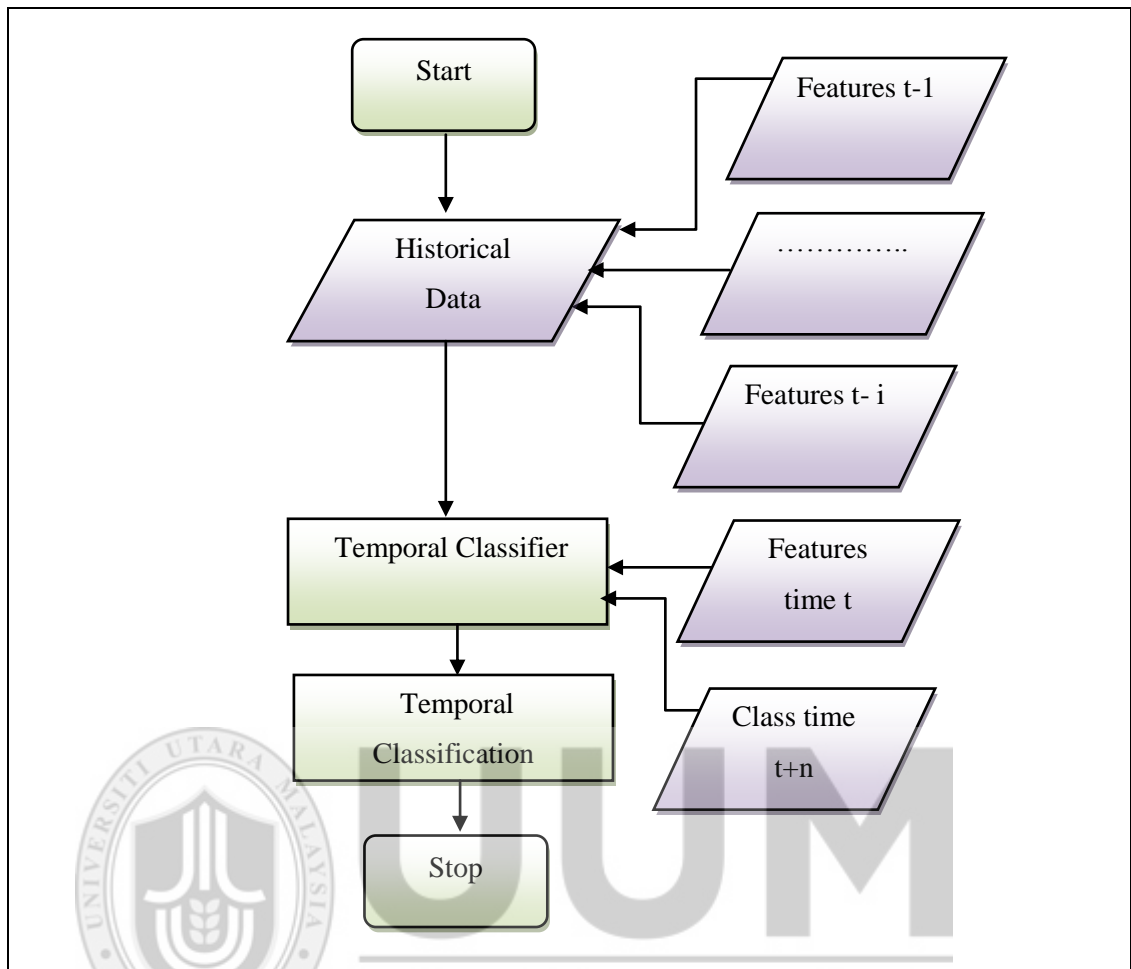


Figure 4.1. Temporal Data Mining Classification

The temporal information of the reservoir water level and rainfall are preserved by using sliding window. The sliding window technique was applied because of its ability to capture the time delay between the data sets (Wan Ishak et al., 2011c) and able to discover pattern from temporal data (Ku-Mahamud et al., 2009). Table 4.7 shows the temporal data representation of reservoir operational data using sliding window. It shows an example of data which has been formed using window size of six. An experiment has been carried out to identify the best window size for the sliding window technique.


```
Read data gauging stations
If delays=6 days Then
Form window slices wrft-5 to t
Else
If delays= 5 days Then
Form window slices wrft-4 to t
Else
If delays= 4 days Then
Form window slices wrft-3 to t
Else
If delays= 3 days Then
Form window slices wrft-2 to t
Else
If delays= 2 days Then
Form window slices wrft-1 to t
End if
Until no more data
Classify window slices
```



Figure 4.2. Pseudo-code for the Sliding Window

In this process, five data sets for five different gauging stations have been formed and each data set represents the time of delay. Each data set symbolizes different sliding window sizes. The window size begins with 2 which represent 2 days of delays and the maximum window size is 6 that represent 6 days of delay. Table 4.8 summarizes the number of instances extracted for each data set based on the window size.

Table 4.8

Data set and the Number of Instances based on the Window Size

Dataset	Window Size	Padang Besar	Tasoh	Lubuk Sireh	Kaki Bukit	Wang Kelian
1	2	25	25	25	25	25
2	3	110	106	107	100	114
3	4	313	296	336	274	386
4	5	665	620	755	587	910
5	6	1125	779	1355	1034	1576

4.2 Spatial Temporal Pattern Extraction Using Neural Network

The experiment on spatial temporal extraction was established by simulating the mapped rainfall and reservoir water level pattern by using neural network. The experiment was carried out to identify the best time of delays for water to flow into the reservoir from five different gauging stations. In this study, five data sets have been randomly divided into three: the training, validation/checking and testing data sets for neural network modelling. The training data are used to train or adjust the weight in NN model, while the validation data are used to find the best NN configuration and training parameters, and the testing data are used to evaluate the

performance of fully trained NN model. This study developed 25 neural network models (5 different gauging stations and each gauging station has 5 data sets). Each NN model is trained with different data set and with different combinations of hidden unit, learning rate and momentum. Figure 4.3 shows the pseudo-code for NN training. This procedure illustrates the value set for hidden unit, learning rate and momentum.

```

for each HIDDEN UNIT
where HU = {3,5,7,9,11,13,15,17,19,21,23,25}

    for each LEARNING RATE
    where LR = {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}

        for each MOMENTUM
        where MM = {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9}

            Training:
            Feedforward ()
            Backpropagation of error ()
            Weight update ()
            Validation ()
            End loop (MM)
        End loop (LR)
    End loop (HU)

```

Figure 4.3. Pseudo-code for NN Training

4.2.1 Experimental Results on Spatial Temporal Pattern Extraction

The experimental results on spatial temporal pattern extraction for five different gauging stations (PADANG BESAR, TASOH, KAKI BUKIT, LUBUK SIREH and WANG KELIAN) have been discussed in this section. The following findings have been obtained by analyzing the results that has best combination of highest percentage of correctness and minimum errors for training, validation and testing.

The percentage of correctness has been calculated by using the following equation (4.6).

$$correctness = \left(\frac{\text{number instances with correct output}}{\text{total number of instances}} \right) \times 100 \quad (4.6)$$

Table 4.9

Result for PADANG BESAR

Dataset	Training		Validation		Testing	
	MSE	%	MSE	%	MSE	%
1	0.24303	83.02	0.56973	57.14	0.33564	85.71
2	0.22770	86.52	0.19834	87.53	0.15026	90.15
3	0.10907	94.82	0.14455	93.55	0.08984	96.77
4	0.06904	96.69	0.05576	98.80	0.07408	97.59
5	0.08433	97.03	0.04112	100.00	0.07210	98.47

Table 4.9 shows the result for PADANG BESAR gauging station. Each column represents the MSE and percentage of correctness (%) for training, validation and testing for each dataset. Overall, the result shows that the minimum training, validation and testing errors for PADANG BESAR are 0.06904, 0.04112 and 0.07210 correspondingly, while the best percentage of correctness achieved for training, validation/checking and testing are 97.03, 100.00, and 98.47. However, the best combination of error and result achieved for PADANG BESAR are (0.08433, 97.03), (0.04112, 100.00) and (0.07210, 98.47). Based on the findings, data set 5 is selected as the best data set which was formed using sliding size 6 that represents six days of delay. This finding suggests that six days is the best time duration of delay for the observation of the upstream rainfall from PADANG BESAR which will significantly increase the water level at the reservoir.

Table 4.10

Neural Network Parameters for PADANG BESAR

Dataset	No of Input Unit	No of Hidden Unit	No of Output Unit	Learning rate	Momentum
1	2	7	1	0.9	0.7
2	3	13	1	0.7	0.6
3	4	3	1	0.1	0.7
4	5	15	1	0.3	0.4
5	6	17	1	0.6	0.7

The values of network parameters achieved after the training is completed are shown in Table 4.10. Overall, the results show that data set 5 provides better results with learning rate (LR) equalling to 0.6 and momentum (MM) equalling to 0.7. The best network architecture of neural network model achieved is 6-17-1.

Table 4.11

Results for TASOH

Dataset	Training		Validation		Testing	
	MSE	%	MSE	%	MSE	%
1	0.16743	94.74	0.78214	33.33	0.25441	66.67
2	0.15311	94.74	0.05235	100.00	0.03944	100.00
3	0.13394	94.49	0.29744	80.00	0.12373	93.33
4	0.09358	96.13	0.11112	96.15	0.10316	96.15
5	0.09378	96.63	0.09660	93.59	0.06265	97.44

Table 4.11 shows the result for TASOH gauging station. Each column represents the MSE and percentage of correctness (%) for training, validation and testing for each dataset. Overall, the result shows that the minimum training, validation and testing errors for TASOH are 0.13394, 0.05235 and 0.03944 respectively, while the best percentage correctness achieved for the training, validation/checking and testing are (94.74), (100.00) and (100.00). However, the best combination of error and result

achieved for TASOH are (0.15311, 94.74), (0.05235, 100.00) and (0.03944, 100.00). Based on the findings, data set 2 is selected as the best data set which was formed using sliding size 3 that represents three days of delay. This finding suggests that three days is the best time duration of delay for the observation of the upstream rainfall from TASOH which will significantly increase the water level at the reservoir.

Table 4.12

Neural Network Parameters for TASOH

Dataset	No of Input Unit	No of Hidden Unit	No of Output Unit	Learning rate	Momentum
1	2	9	1	0.4	0.5
2	3	9	1	0.2	0.7
3	4	3	1	0.5	0.7
4	5	7	1	0.9	0.4
5	6	25	1	0.5	0.1

The values of network parameters achieved after the completion of the training are shown in Table 4.12. Overall, the results show that data set 2 provides better results with learning rate equalling to 0.2 (LR) and momentum (MM) equalling to 0.7. The best network architecture of neural network model achieved is 3-9-1.

Table 4.13

Results for LUBUK SIREH

Dataset	Training		Validation		Testing	
	MSE	%	MSE	%	MSE	%
1	0.03539	100.00	0.11076	100.00	0.41828	66.67
2	0.16553	91.76	0.18354	90.91	0.12410	100.00
3	0.01263	100.00	0.02659	100.00	0.00662	100.00
4	0.06950	97.68	0.09649	96.05	0.09787	94.74
5	0.08254	98.06	0.09668	95.59	0.07760	97.79

Table 4.13 shows the result for LUBUK SIREH gauging station. Each column represents the MSE and percentage of correctness (%) for training, validation and testing for each dataset. Overall, the result shows that the minimum training, validation and testing errors for LUBUK SIREH are 0.01263, 0.02659 and 0.00662 respectively, while the best percentage of correctness achieved for the training, validation/ checking and testing are (100.00), (100.00), and (100.00). Based on the findings, data set 3 is selected as the best data set which was formed using sliding size 4 that represents four days of delay. This finding suggests that four days is the best time duration of delay for the observation of the upstream rainfall from LUBUK SIREH which will significantly increase the water level at the reservoir.

Table 4.14

Neural Network Parameters for LUBUK SIREH

Dataset	No of Input Unit	No of Hidden Unit	No of Output Unit	Learning rate	Momentum
1	2	3	1	0.8	0.2
2	3	3	1	0.2	0.9
3	4	3	1	0.3	0.2
4	5	13	1	0.5	0.2
5	6	7	1	0.8	0.7

The values of network parameters achieved after the training is completed are shown in Table 4.14. Overall, the results show that data set 3 provides the best results with learning rate equalling to 0.3 (LR) and momentum (Mom) equalling to 0.2. The best network architecture of neural network model achieved is 4-3-1.

Table 4.15

Results for KAKI BUKIT

Dataset	Training		Validation		Testing	
	MSE	%	MSE	%	MSE	%
1	0.02239	100.00	0.28815	66.67	0.97164	66.67
2	0.01347	100.00	0.00138	100.00	0.00013	100.00
3	0.11799	95.45	0.21663	85.19	0.08473	100.00
4	0.07642	97.44	0.09731	94.92	0.08712	96.61
5	0.05952	98.55	0.08294	97.09	0.04560	100.00

Table 4.15 shows the result for KAKI BUKIT gauging. Each column represents the MSE and percentage of correctness (%) for training, validation and testing for each dataset. Overall, the result shows that the minimum training, validation and testing errors for KAKI BUKIT are 0.01347, 0.00138 and 0.00013 respectively, while the best percentage of correctness achieved for the training, validation/ checking and testing are (100.00), (100.00), and (100.00). Based on the findings, data set 2 is selected as the best data set which was formed using sliding window size 3 that represents three days of delay. This finding suggests that three days is the best time duration of delay for the observation of the upstream rainfall from KAKI BUKIT which will significantly increase the water level at the reservoir.

Table 4.16

Neural Network Parameters for KAKI BUKIT

Dataset	No of Input Unit	No of Hidden Unit	No of Output Unit	Learning rate	Momentum
1	2	11	1	0.9	0.7
2	3	7	1	0.8	0.7
3	4	13	1	0.1	0.9
4	5	9	1	0.4	0.2
5	6	5	1	0.3	0.1

The values of network parameters achieved after the completion of the training are shown in Table 4.16. Overall, the results show that the data set 2 provides the best results with learning rate equalling to 0.8 (LR) and momentum (MM) equalling to 0.7. The best network architecture of neural network model achieved is 3-7-1.

Table 4.17

Results for WANG KELIAN

Dataset	Training		Validation		Testing	
	MSE	%	MSE	%	MSE	%
1	0.00841	100.00	0.39774	66.67	0.40395	66.67
2	0.10821	94.57	0.33501	81.82	0.23830	81.82
3	0.21539	92.21	0.19389	89.74	0.15009	94.87
4	0.08485	96.70	0.07559	96.70	0.09012	95.60
5	0.04518	99.22	0.04180	98.31	0.04760	99.44

Table 4.17 shows the result for WANG KELIAN gauging station. Each column represents the MSE and percentage of correctness (%) for training, validation and testing for each dataset. Overall, the result shows that the minimum errors of training, validation and testing for WANG KELIAN are 0.0084, 0.04180 and 0.04760 respectively, while the best percentage of correctness achieved for the training, validation/checking and testing are (100.00), (98.31), and (99.44). However, the best combinations of the error and results achieved for WANG KELIAN are (0.04518, 99.22), (0.04180, 98.31) and (0.04760, 99.44). Based on the findings, data set 5 is selected as the best data set which was formed using sliding window size 6 that represents six days of delay. This finding suggests that six days is the best time duration of delay for the observation of the upstream rainfall from WANG KELIAN which will significantly increase the water level at the reservoir.

Table 4.18

Neural Network Parameters for WANG KELIAN

Dataset	No of Input Unit	No of Hidden Unit	No of Output Unit	Learning rate	Momentum
1	2	11	1	0.3	0.7
2	3	3	1	0.2	0.5
3	4	5	1	0.8	0.8
4	5	11	1	0.9	0.3
5	6	23	1	0.3	0.1

The values of network parameters achieved after the training is completed are shown in Table 4.18. Overall, the results show that data set 5 provides the best results with learning rate equalling to 0.3 (LR) and momentum (MM) equalling to 0.1. The best network architecture of neural network model achieved is 6-23-1.

Table 4.19

Summary of the Findings

Gauging Station	Window Size	Training		Validation		Testing	
		MSE	%	MSE	%	MSE	%
Padang Besar	6	0.08433	97.03	0.04112	100.00	0.07210	98.47
Tasoh	3	0.15311	94.74	0.05235	100.00	0.03944	100.00
Lubuk Sireh	4	0.01263	100.00	0.02659	100.00	0.00662	100.00
Kaki Bukit	3	0.01347	100.00	0.00138	100.00	0.00013	100.00
Wang Kelian	6	0.04518	99.22	0.04180	98.31	0.04760	99.44

Table 4.19 shows a summary of the findings for five gauging stations: PADANG BESAR, TASOH, LUBUK SIREH, KAKI BUKIT and WANG KELIAN. In addition, a summary of network specification for each gauging station is shown in Table 4.20.

Table 4.20

Summary of Network Specification

Gauging Station	Dataset	#Input	#Output	Hidden units	Learning Rate	Momentum
Padang Besar	5	6	1	17	0.6	0.7
Tasoh	2	3	1	9	0.2	0.7
Lubuk Sireh	3	4	1	3	0.3	0.2
Kaki Bukit	2	3	1	7	0.8	0.7
Wang Kelian	5	6	1	23	0.3	0.1

4.2.2 Spatial Temporal Pattern of Rainfall

The finding of spatial temporal pattern extraction has been used as input variables for modelling reservoir water release decision using ANFIS in the next chapter. This model was derived based on the results of neural network modelling from spatial temporal pattern extraction. The input and output of the model was represented as:

$$\begin{aligned}
 WL(t) = RF & (pb(t-5), pb(t-4), pb(t-3), pb(t-2), pb(t-1), pb(t), ts(t-2), \\
 & ts(t-1), ts(t), ls(t-3), ls(t-2), ls(t-1), ls(t), kb(t-2), kb(t-1), kb(t), wk(t-5), \\
 & wk(t-4), wk(t-3), wk(t-2), wk(t-1), wk(t)) \quad (4.7)
 \end{aligned}$$

The output of the models is $WL(t)$ that represents the reservoir water level at time t ; RF represent rainfall at the each gauging station, where the subscripts pb , ts , ls , kb , and wk represent gauging stations of Padang Besar, Tasoh, Lubuk Sireh, Kaki Bukit, and Wang Kelian respectively.

4.3 Spatial Temporal Pattern Extraction Procedure

As mentioned earlier, the special pattern extraction procedure needs to be determined in order to maintain the existence of spatial temporal relationship in the reservoir water release decision pattern. Figure 4.4 shows the spatial temporal pattern extraction procedure which has been adopted in this study.

This procedure starts with data collection from two sources: reservoir operational data and hydrological data. The next step is data pre-processing that consists of two activities which are data cleaning and data normalization. The next step followed by spatial temporal information identification and spatial temporal pattern formation. These two steps are most important step to identify the spatial temporal pattern by using sliding window and neural network. The process to identify the best pattern is start from applying sliding window to find temporal pattern of rainfall. After the temporal pattern of rainfall has been extracted, neural network classifier has been applied to identify the best NN model based on the results. The best results have been compared based on the percentage of correctness (%) and error (MSE). The last step is to select the best model which forms the best pattern of rainfall. Finally, the best spatial temporal pattern has been identified and used as an input to develop ANFIS model.

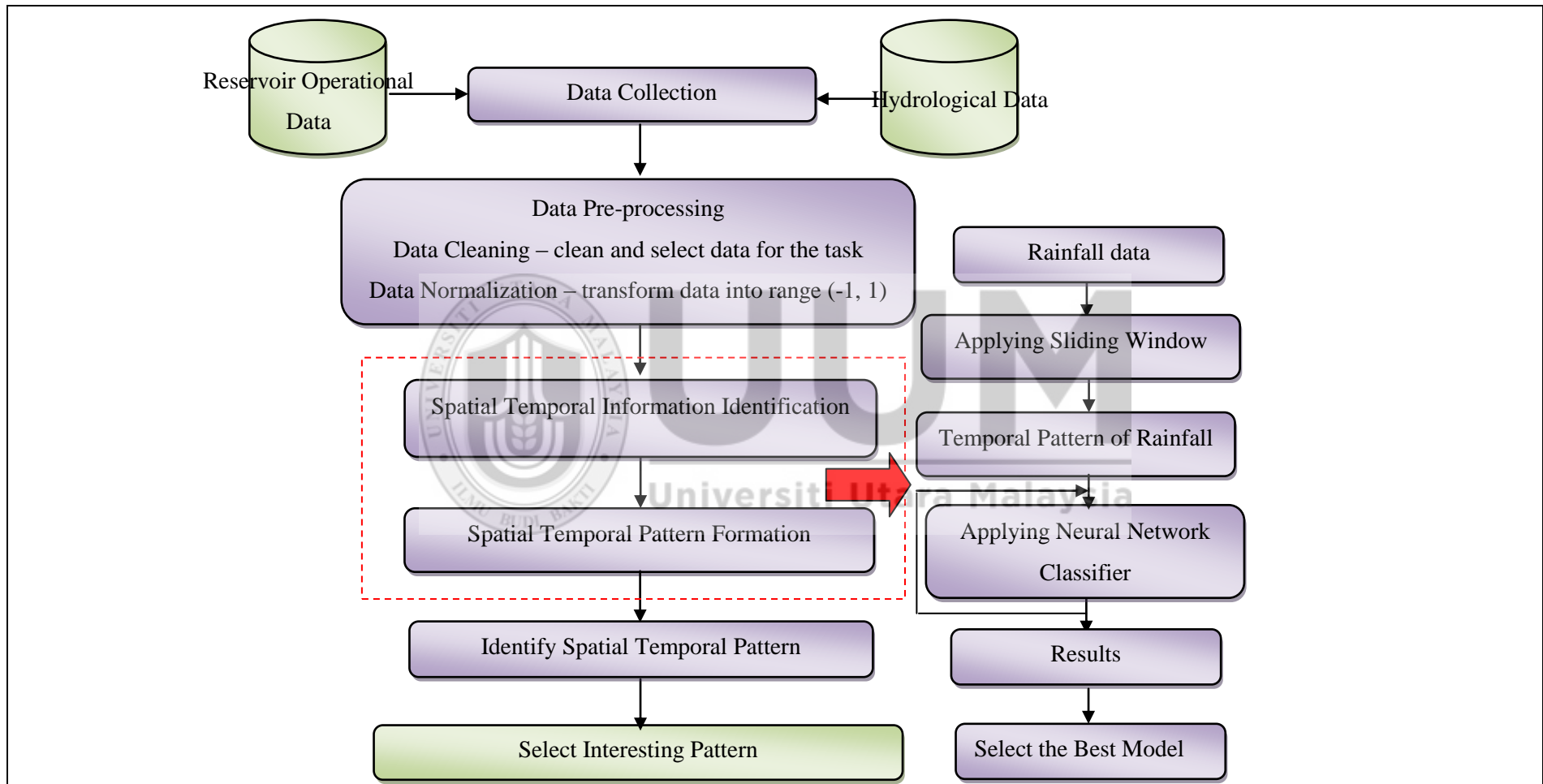


Figure 4.4. Steps in Spatial Temporal Pattern Extraction Procedure

4.4 Summary

In this chapter, the complete process of extracting spatial temporal pattern from reservoir operational data has been described. The spatial and temporal data mining technique was used to extract spatial temporal pattern from reservoir operational data that contains information of experts' knowledge and experiences. In this chapter, the experimental result on the spatial temporal pattern extraction has also been provided. In the next chapter, the explanation on the development reservoir water release dynamic decision model will be discussed.



CHAPTER FIVE

PROPOSED RESERVOIR WATER RELEASE DYNAMIC DECISION MODEL BASED ON SPATIAL TEMPORAL PATTERN

This chapter is presented to discuss on the development of reservoir water release dynamic decision model. The steps show the detailed process on how to develop a reservoir water release dynamic decision model using the Adaptive Neuro-Fuzzy Inference System (ANFIS).

5.1 Development of Reservoir Water Release Dynamic Decision Model

As mentioned in the previous chapter, the development of a reservoir water release dynamic decision model has been conducted by using ANFIS. The spatial temporal pattern of rainfall that has been extracted from the previous phase has been used as the input variables of the model. In this study, MATLAB tool is used to develop the ANFIS model for reservoir water release decision. The following sub section shows the steps for developing a reservoir water release dynamic decision model.

5.1.1 Constructing Fuzzy Membership Function

A *membership function* (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. In this study, a fuzzy membership function has been constructed for reservoir flood stage. The minimum entropy method has been used to identify the threshold line for classifying the samples while minimizing the entropy for an optimum partition. The threshold value has been calculated by using Equations (5.1-5.7).

$$S(x) = p(x)Sp(x) + q(x)Sq(x) \quad (5.1)$$

where

$$Sp(x) = -[p_1(x) \ln p_1(x) + p_2(x) \ln p_2(x)] \quad (5.2)$$

$$Sq(x) = -[q_1(x) \ln q_1(x) + q_2(x) \ln q_2(x)] \quad (5.3)$$

and where

$p_k(x)$ and $q_k(x)$ = conditional probabilities that the class k sample is in the region $[x_1, x_1 + x]$ and $[x_1 + x, x_2]$, respectively

$p(x)$ and $q(x)$ = probabilities that all samples are in the region $[x_1, x_1 + x]$ and $[x_1 + x, x_2]$, respectively 12

$$p(x) + q(x) = 1$$

$$pk(x) = \frac{n_{k(x)+1}}{n(x)+1} \quad (5.4)$$

$$qk(x) = \frac{N_{k(x)+1}}{N(x)+1} \quad (5.5)$$

$$p(x) = \frac{n_x}{n} \quad (5.6)$$

$$q(x) = 1 - p(x) \quad (5.7)$$

Where

$n_k(x)$ = number of class k samples located in $[x_1, x_1 + x]$

$n(x)$ = the total number of samples located in $[x_1, x_1 + x]$

$N_k(x)$ = number of class k samples located in $[x_1 + x, x_2]$

$N(x)$ = the total number of samples located in $[x_1 + x, x_2]$

n = total number of samples in $[x_1, x_2]$

Calculation of Membership Function using Minimum Entropy Method

This example shows the calculation of membership function using entropy method for reservoir flood stage. There are four categories of flood stage: Normal, Alert, Warning and Danger. Table 5.1 shows the segmentation of x value from the raw data.

Table 5.1

Segmentation of x into two arbitrary classes (from raw data)

N	1	2	3	4	5	6	7	8	9	10	11	12
$x=a/b$	27.80	28.10	28.30	28.50	28.65	28.85	28.90	29.00	29.25	29.55	29.75	29.94
class	1	1	1	1	1	2	1	1	2	2	2	2

The value has been selected randomly from raw data and the data are classified into two classes either class 1 or class 2. Figure 5.1 shows the illustration of threshold value idea which divides the space into two regions: p and q. Each region has a value from class 1 and class 2.

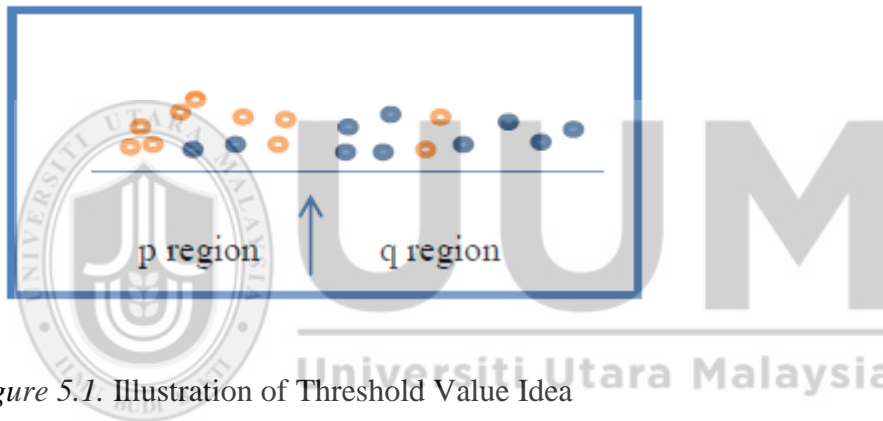


Figure 5.1. Illustration of Threshold Value Idea

The calculation of Primary Threshold Value (PRI) is compulsory for dividing the data into two sections: PO and NG sides. With this PRI value, the sections are divided in two, in which PO or known as positive refers to the right side of the primary threshold, and NG or known as negative refers to the left side of the primary threshold. Table 5.2 shows the calculation of Primary Threshold Value (PRI). The value of 29.10 has been selected as a Primary Threshold Value (PRI) because of having the minimum $S(x)$ value.

Table 5.2

Calculation of Primary Threshold Value (PRI)

<i>Value</i>	1	2	3	4	5	6	7	8	9	10	11
$x1$	27.90	28.20	28.40	28.60	28.80	28.87	28.95	29.10	29.40	29.60	29.80
$p1$	1	1	1	1	1	0.8571	0.875	0.8889	0.8	0.7273	0.6667
$p2$	0.5	0.3333	0.25	0.2	0.167	0.2857	0.25	0.2222	0.3	0.3636	0.4167
$q1$	0.5833	0.5455	0.5	0.444	0.375	0.4286	0.3333	0.2	0.25	0.3333	0.5
$q2$	0.5	0.5455	0.6	0.667	0.75	0.7143	0.8333	1	1	1	1
$p(x)$	0.0833	0.1667	0.25	0.333	0.417	0.5	0.5833	0.6667	0.75	0.8333	0.9167
$q(x)$	0.9167	0.8333	0.75	0.667	0.583	0.5	0.4167	0.3333	0.25	0.1667	0.0833
$Sp(x)$	0.347	0.366	0.347	0.322	0.299	0.490	0.463	0.439	0.540	0.599	0.635
$Sq(x)$	0.661	0.661	0.653	0.631	0.584	0.603	0.518	0.322	0.347	0.366	0.347
$S(x)$	0.635	0.612	0.576	0.528	0.465	0.547	0.486	0.400	0.491	0.561	0.611

The next calculation is to determine the secondary threshold value of NG side. The same procedure is applied to calculate this secondary threshold value. Table 5.3 shows the calculation of threshold value for NG side. The value of 28.80 has been selected as a secondary threshold value for NG side because of having the minimum $S(x)$ value. $S(x)$ values are selected based on the minimum entropy principle.

Table 5.3

Calculation Threshold Value of NG Side

<i>Value</i>	1	2	3	4	5	6	7
$x2$	27.9	28.2	28.40	28.60	28.80	28.87	28.95
$p1$	1	1	1	1	1	0.8571	0.875
$p2$	0.5	0.3333	0.25	0.2	0.167	0.2857	0.25
$q1$	0.875	0.8571	0.833	0.8	0.75	1	1
$q2$	0.25	0.2857	0.333	0.4	0.5	0.3333	0.5
$p(x)$	0.125	0.25	0.375	0.5	0.625	0.75	0.875
$q(x)$	0.875	0.75	0.625	0.5	0.375	0.25	0.125
$Sp(x)$	0.347	0.366	0.347	0.322	0.299	0.490	0.463
$Sq(x)$	0.463	0.490	0.518	0.545	0.562	0.366	0.347
$S(x)$	0.449	0.459	0.454	0.433	0.398	0.459	0.449

The next calculation is to determine the secondary threshold value of PO side. Table 5.4 shows the calculation of threshold value for PO side. The value of 29.40 has been selected as a threshold value for PO side because of having the minimum $S(x)$ value.

Table 5.4

Calculation Threshold Value of PO Side

<i>Value</i>	1	2	3
$x3$	29.40	29.60	29.80
$p1$	0.5	0.3333	0.25
$p2$	1	1	1
$q1$	0.25	0.3333	0.5
$q2$	1	1	1
$p(x)$	0.25	0.5	0.75
$q(x)$	0.75	0.5	0.25
$Sp(x)$	0.347	0.366	0.347
$Sq(x)$	0.347	0.366	0.347
$S(x)$	0.347	0.366	0.347

This last step is to calculate the next threshold value of PO side. Table 5.5 shows the calculation of threshold value for PO side. The value of 29.60 has been selected as last threshold value for PO side because of having the minimum $S(x)$ value.

Table 5.5

Calculation Threshold Value of PO Side

<i>Value</i>	1	2
$x4$	29.60	29.80
$p1$	0.5	0.3333
$p2$	1	1
$q1$	0.3333	0.5
$q2$	1	1
$p(x)$	0.3333	0.6667
$q(x)$	0.6667	0.3333
$Sp(x)$	0.347	0.366
$Sq(x)$	0.366	0.347
$S(x)$	0.360	0.360

Figure 5.2 illustrates the membership function graph for reservoir flood stages: Normal (N), Alert (A), Warning (W) and Danger (D). The threshold values for each partition are mapped into a graph after the calculation has been completed. This membership function which has been constructed will be used to develop an ANFIS model for modelling reservoir water release decision.

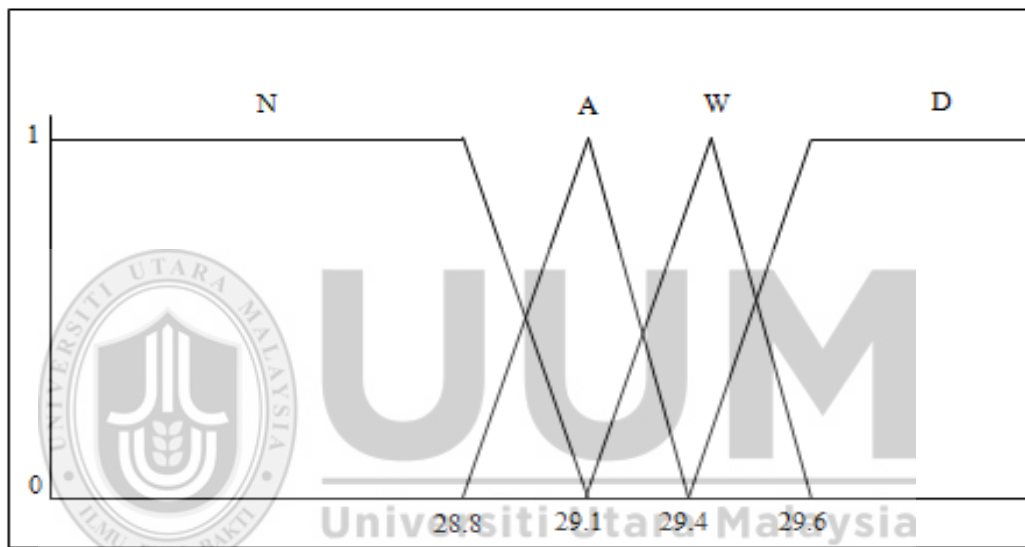


Figure 5.2. Membership Function Graph for Flood Stage

5.1.2 ANFIS Model Development

In this study, a hybrid computational intelligence approach, namely Adaptive Neuro-Fuzzy Inference System (ANFIS) (Jang, 1993; Chang & Chang, 2006; Valizadeh & El-Shafie, 2013) was used for modelling reservoir water release decision. The basic architecture of the ANFIS model was shown in Figure 3.4 in chapter 3. The input and output of the ANFIS model have been selected based on the gate opening events: open or close gates. After the selection, only 249 data have been selected and these data that represent the open or close gate events were divided into three datasets: the training (70%), checking (15%), and testing sets (15%).

Table 5.6 shows an example of training set for the ANFIS models. Two ANFIS models have been developed for gate opening decisions: Model 1 with rainfall patterns from five gauging stations and current reservoir water level as an input variable, Model 2 without reservoir water level. The rainfall data from five gauging stations and the water level of the TIMAH TASOH reservoir are used as the input variables. Because the distances between the gauging stations are different and the times of water to flow to the TIMAH TASOH reservoir vary, the travel times from five different gauging stations are being considered for constructing the models and the shifted inputs of the five gauging stations are adopted. In total, there are two models which have been developed with 23 input variables trained in Model 1 and 22 input variables trained in Model 2. The model inputs and outputs of the model are represented as:

Model 1 with 23 input variables (rainfall patterns from five gauging stations and current reservoir water level):

$$Go(t+i) = RF(pb(t-5), pb(t-4), pb(t-3), pb(t-2), pb(t-1), pb(t), ts(t-2), ts(t-1), ts(t), ls(t-3), ls(t-2), ls(t-1), ls(t), kb(t-2), kb(t-1), kb(t), wk(t-5), wk(t-4), wk(t-3), wk(t-2), wk(t-1), wk(t)), WL(t) \quad (5.8)$$

Model 2 with 22 input variables (rainfall patterns from five gauging stations without current reservoir water level):

$$Go(t+i) = RF(pb(t-5), pb(t-4), pb(t-3), pb(t-2), pb(t-1), pb(t), ts(t-2), ts(t-1), ts(t), ls(t-3), ls(t-2), ls(t-1), ls(t), kb(t-2), kb(t-1), kb(t), wk(t-5), wk(t-4), wk(t-3), wk(t-2), wk(t-1), wk(t)) \quad (5.9)$$

The output of the models is $Go(t+i)$ that represents the status of water release which is either to open or close the gate at the time $t+i$; where the subscripts pb , ts , ls , kb , and wk represent gauging stations of Padang Besar, Tasoh, Lubuk Sireh, Kaki Bukit, and Wang Kelian respectively. Besides that, $WL(t)$ represents the reservoir water level at time t .

Table 5.6

Example of Training Set for ANFIS Model 1

padang besar						tashoh			lubuk sireh				kaki bukit			wang kelian						water level	gate
0.00	2.00	0.00	13.50	0.00	12.00	15.50	0.00	0.00	0.00	5.00	5.00	6.00	0.00	9.00	7.00	0.00	0.00	0.00	7.50	0.00	18.50	28.91	1
2.00	0.00	13.50	0.00	12.00	2.50	0.00	0.00	0.00	5.00	5.00	6.00	0.00	9.00	7.00	6.00	0.00	0.00	7.50	0.00	18.50	5.50	28.91	0
58.00	0.00	0.00	0.00	7.00	8.00	39.00	21.00	0.00	26.00	2.00	0.00	0.00	0.00	0.00	10.00	0.00	0.00	0.00	16.50	0.00	11.50	29.14	0
0.00	0.00	0.00	7.00	8.00	32.00	21.00	0.00	0.00	2.00	0.00	0.00	67.00	0.00	10.00	11.00	0.00	0.00	16.50	0.00	11.50	35.50	29.21	1
7.00	8.00	32.00	1.00	26.50	1.00	28.00	38.00	28.00	67.00	15.00	39.00	0.00	20.50	0.00	0.00	0.00	11.50	35.50	0.00	18.50	1.50	29.00	1
8.00	32.00	1.00	26.50	1.00	0.50	38.00	28.00	2.00	15.00	39.00	0.00	31.00	0.00	0.00	0.00	11.50	35.50	0.00	18.50	1.50	0.00	29.07	0
32.00	1.00	26.50	1.00	0.50	0.00	28.00	2.00	21.50	39.00	0.00	31.00	0.00	0.00	0.00	9.00	35.50	0.00	18.50	1.50	0.00	0.00	29.15	0
1.00	26.50	1.00	0.50	0.00	45.50	2.00	21.50	0.00	0.00	31.00	0.00	0.00	0.00	9.00	10.50	0.00	18.50	1.50	0.00	0.00	13.50	29.24	1
0.00	0.00	10.00	5.00	7.50	0.00	0.00	0.00	0.00	0.00	0.00	5.00	3.00	0.00	6.00	0.00	0.00	0.00	2.50	0.00	56.50	0.00	28.91	1
0.00	10.00	5.00	7.50	0.00	0.00	0.00	0.00	0.00	0.00	5.00	3.00	10.00	6.00	0.00	0.00	0.00	2.50	0.00	56.50	0.00	2.00	28.91	0
0.00	0.00	0.00	0.00	0.00	0.00	5.00	7.00	0.00	0.00	3.00	0.00	8.00	20.50	0.00	0.00	2.50	2.50	3.50	22.50	0.00	0.00	29.03	0
0.00	0.00	0.00	0.00	0.00	0.00	7.00	0.00	0.00	3.00	0.00	8.00	0.00	0.00	0.00	0.00	2.50	3.50	22.50	0.00	0.00	0.00	29.21	1
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	29.00	1
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	28.99	0
8.50	29.00	19.50	15.50	1.00	0.00	29.00	26.00	38.00	3.00	50.00	35.00	35.00	0.00	0.00	11.00	3.50	88.50	35.50	18.00	0.00	2.50	29.19	0
29.00	19.50	15.50	1.00	0.00	0.00	26.00	38.00	0.00	50.00	35.00	35.00	0.00	0.00	11.00	0.00	88.50	35.50	18.00	0.00	2.50	40.00	29.21	1
0.00	7.00	1.50	10.00	1.00	13.00	0.00	7.50	3.00	5.00	10.00	1.00	16.00	0.00	8.00	6.00	0.00	8.50	1.50	17.50	36.00	2.50	28.97	1
7.00	1.50	10.00	1.00	13.00	40.00	7.50	3.00	0.00	10.00	1.00	16.00	0.00	8.00	6.00	0.00	8.50	1.50	17.50	36.00	2.50	0.00	28.91	0
4.00	45.50	1.50	13.00	8.00	65.00	11.00	15.00	0.00	32.00	60.00	2.00	20.00	14.50	10.00	11.00	24.50	144.00	10.50	34.50	6.50	17.50	29.04	0
45.50	1.50	13.00	8.00	65.00	39.50	15.00	0.00	19.00	60.00	2.00	20.00	5.00	10.00	11.00	14.00	144.00	10.50	34.50	6.50	17.50	51.50	29.09	1
0.00	0.00	4.00	10.00	25.00	9.50	0.00	0.00	0.00	4.00	0.00	9.00	45.00	1.00	10.00	1.00	0.00	0.00	7.50	2.50	51.50	39.50	28.90	1
0.00	4.00	10.00	25.00	9.50	18.00	0.00	0.00	30.00	0.00	9.00	45.00	0.00	10.00	1.00	0.00	0.00	7.50	2.50	51.50	39.50	42.50	28.90	0
25.00	9.50	18.00	1.00	0.00	2.00	30.00	60.00	8.00	0.00	22.00	37.00	0.00	20.00	31.50	34.50	51.50	39.50	42.50	2.50	0.00	14.50	28.96	0
9.50	18.00	1.00	0.00	2.00	0.00	60.00	8.00	30.50	22.00	37.00	0.00	0.00	31.50	34.50	0.00	39.50	42.50	2.50	0.00	14.50	20.50	28.98	1

5.2 Development of Reservoir Water Release Dynamic Decision Model using ANFIS

The experiment on the development of reservoir water release dynamic decision model has been conducted by using ANFIS in which implementation has been done in MATLAB tool. Table 5.7 shows the parameters used in these ANFIS model. The number of epochs refers to the frequency to train the model. Method of optimization refers to the optimization methods available for FIS training, namely hybrid (the default, a combination of mixed least squares estimation and back propagation) or back propa (back propagation) algorithm alone. A mixed least squares estimation is used to determine the consequent parameters in the forward pass learning algorithm and gradient descent is used to update the premise parameters in the backward pass algorithm. Meanwhile, error tolerance is used to generate a training stopping criterion, which is related to the error size (default=0). These numbers of parameters are chosen based on the experiment made by training the model using different number of parameters. The following parameters have been selected to train the model.

Table 5.7

ANFIS Parameters

Parameters	Value
Number of Epochs	20
Method of Optimization	Hybrid, Back propagation
Error Tolerance	0.05

As mentioned in Section 5.12, there are two ANFIS models which have been developed with different numbers of input variables, i.e. 23 for Model 1 and 22 for Model 2, respectively. The results are summarized in Table 5.8. The performances of the ANFIS models are compared using three different datasets: training sets, checking sets, and testing sets. Both Model 1 and Model 2 have been trained with different methods of optimization: hybrid and back propagation. The performances of the models were measured by using two performance criteria, RMSE and MAE as shown in Figure 5.3.

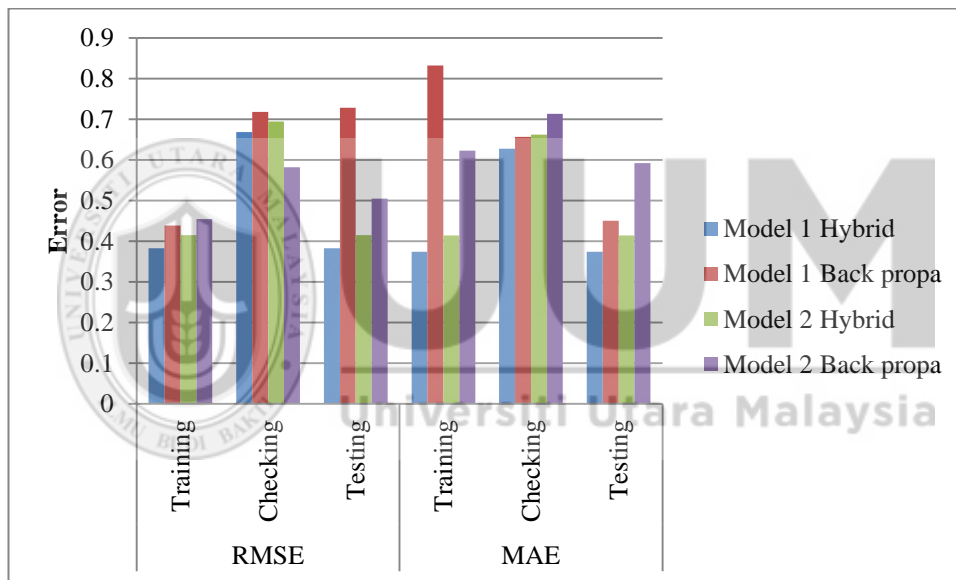


Figure 5.3. Performance Comparison between Hybrid and Backpropagation Model

Based on the results in Table 5.8, the hybrid method performed well compared to back propagation method. Both Model 1 and Model 2 have the lowest errors by using the hybrid method. The lowest errors achieved for training, checking and testing are (0.374083, 0.62738, 0.37408) and (0.38246, 0.66845, 0.38246) for Model 1 with 23 inputs. The development of Model 1 which includes the reservoir water level as an input variable has better performance compared to Model 2. This

situation demonstrates that the involvement of human decision through their experiences and instances are important for the future reservoir system because the previous decision is made based on their instances and past experiences is being considered for modelling reservoir water release decision. In addition, the results proved that the water level of the reservoir can be one of the variables that give significant effects on the water release decision.

Overall, based on the results, it shows the proposed two ANFIS model by using hybrid method perform well compared to backpropagation alone for modelling reservoir water release decision. The results illustrate that the application of ANFIS can be used to develop the decision models which might provide accurate water release prediction.

Table 5.8

Results for ANFIS Model

ANFIS Model	Method	Input	RMSE			MAE		
			Training	Checking	Testing	Training	Checking	Testing
Model 1	Hybrid	23	0.38246	0.66845	0.38246	0.37408	0.62738	0.37408
	Back propa	23	0.43847	0.71756	0.72843	0.83175	0.65645	0.45022
Model 2	Hybrid	22	0.41486	0.69457	0.41486	0.41391	0.6619	0.41391
	Back propa	22	0.45392	0.58141	0.50492	0.62303	0.71281	0.59164

The results of ANFIS model 1 is shown in Table 5.9 with the following information:

Table 5.9

ANFIS Model 1 Result

Information	Value
Number of nodes	122
Number of linear parameters	48
Number of nonlinear parameters	92
Total number of parameters	140
Number of training data pairs	248
Number of checking data pairs	42
Number of fuzzy rules	3

The ANFIS model structure is shown in Figure 5.4 and it illustrates the architecture of ANFIS model that consists of five layers: input, inputmf, rule, outputmf and output. This figure shows the number of rules is reduced significantly by using subtractive fuzzy clustering. Since the input variables used in this study are quite large: 23 and 22, which can generate different possible conditions, the basic structure of ANFIS model becomes expandable. In order to solve this problem, the subtractive clustering has been applied to establish the rule base relationship between the input and output variables.

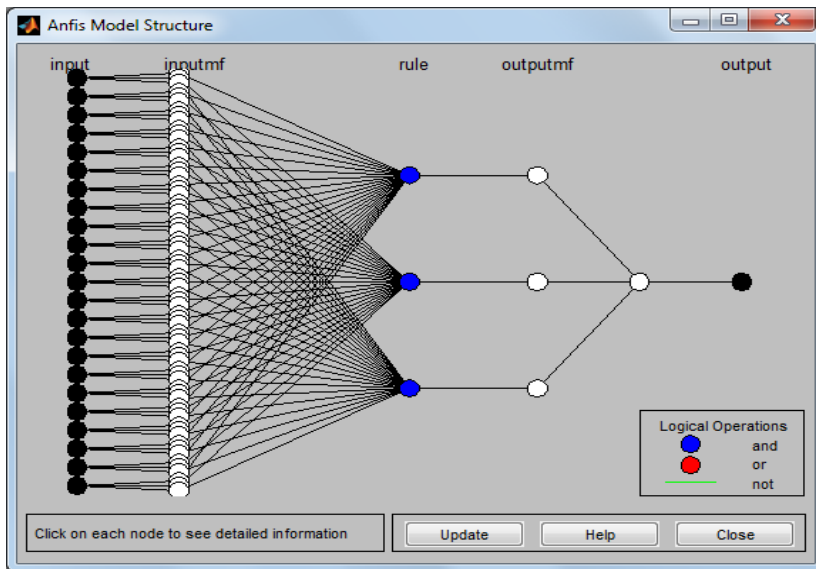


Figure 5.4. ANFIS Model Structure

Figure 5.5 shows the training error and checking error for Model 1. The training error was plotted with the checking error.

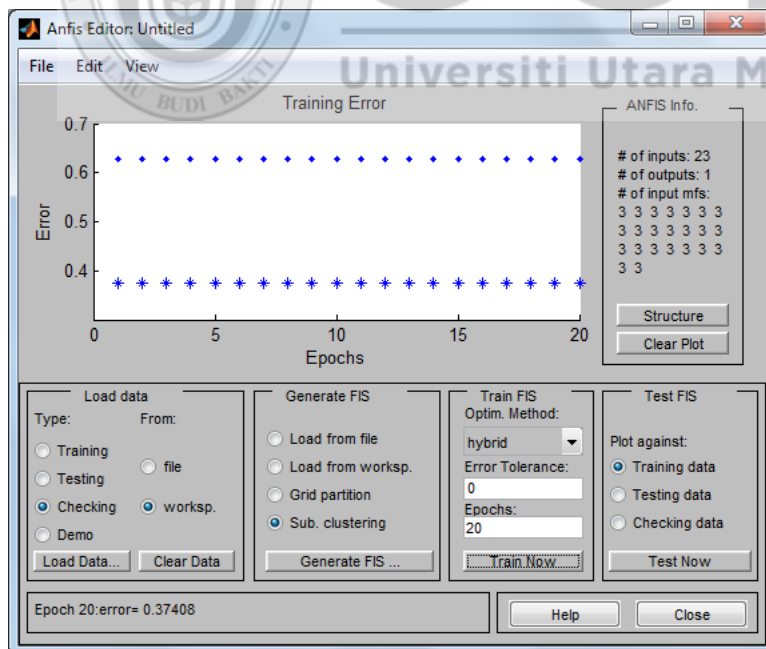


Figure 5.5. Training Error & Checking error for Model 1

Figure 5.6 shows the rule generated for Model 1. It shows three rules has been produce from training Model 1.

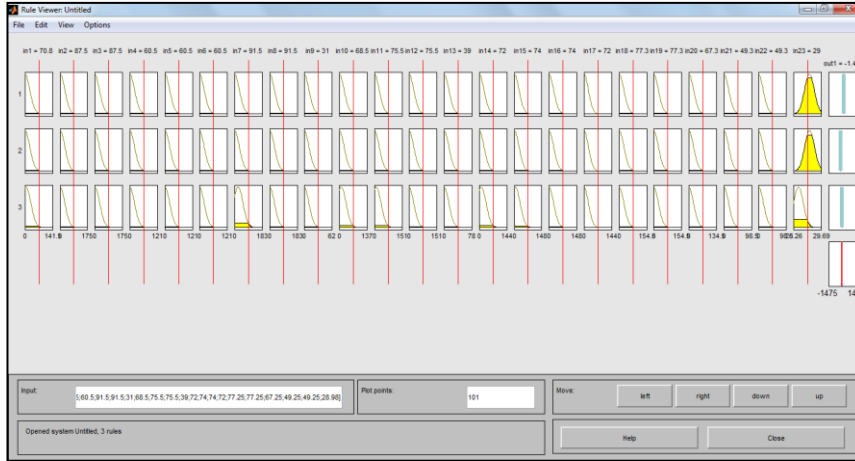


Figure 5.6. Rule Viewer for Model 1

The result of ANFIS Model 2 is shown in Table 5.10 with the following information:

Table 5.10

ANFIS Model 2 Result

Information	Value
Number of nodes	117
Number of linear parameters	46
Number of nonlinear parameters	88
Total number of parameters	134
Number of training data pairs	163
Number of checking data pairs	42
Number of fuzzy rules	2

Figure 5.7 demonstrates the training error and checking error for Model 2. The training error was plotted with the checking error.

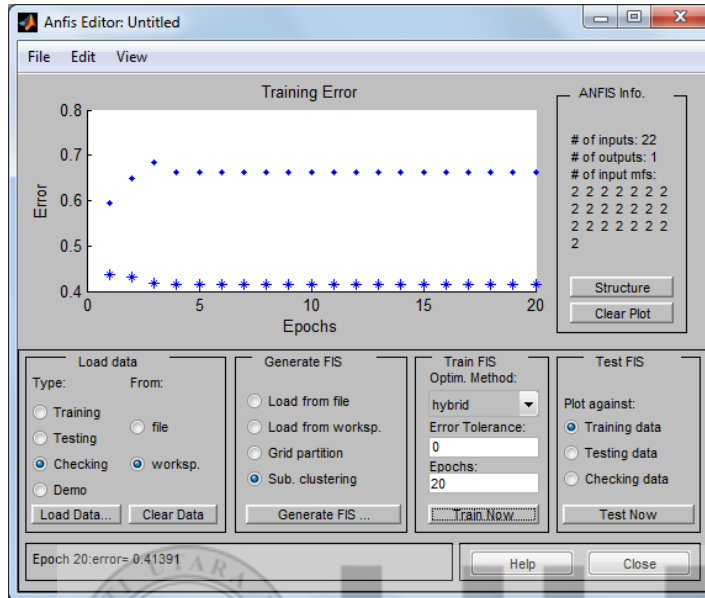


Figure 5.7. Training error & Checking error for Model 2

Figure 5.8 shows the rule generated for Model 2. It shows two rules has been produce from training Model 2.

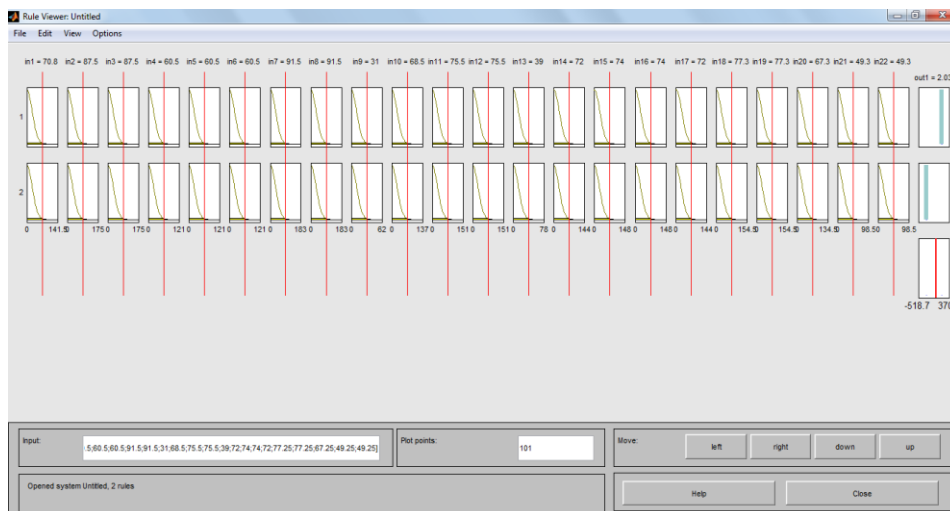


Figure 5.8. Rule Viewer for Model 1

5.3 Performance Evaluation

In this study, a comparison between BP NN, ANFIS, and ARIMA for modelling reservoir water release decision is presented. The experiment was carried out by using Matlab 2012b tool. The result has been obtained by run ANFIS, ARIMA and NN model by using the same input and output datasets. Figure 5.9 and Figure 5.10 shows the experiments conducted using Matlab tool for ARIMA and BPNN models. Two performance measurements, RMSE and MAE, are used for analysis and to evaluate the performances of the models. Overall, the result of this study shows that the performance of the ANFIS model is better compared to other models due to the fact that the values of MAE and RMSE are lower than other values.

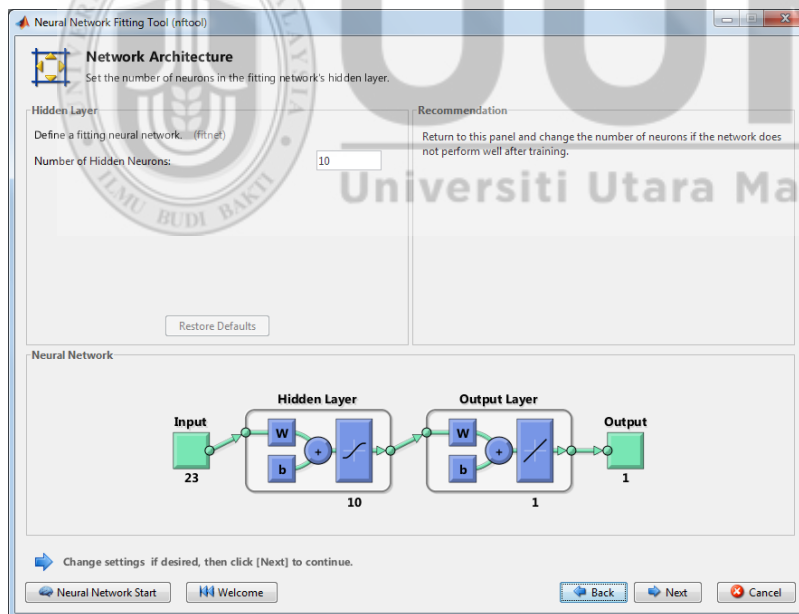


Figure 5.9. BPNN Model using Matlab tool

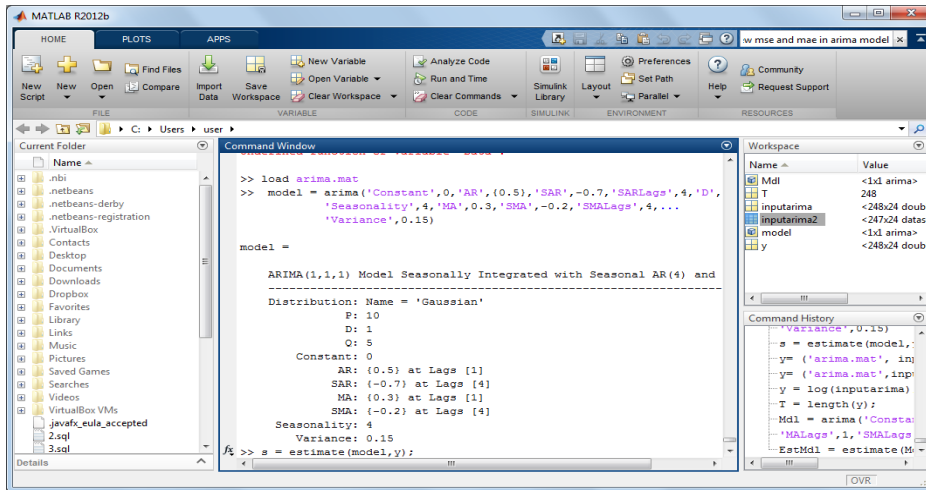


Figure 5.10. ARIMA Model using Matlab tool

The performance of the models was evaluated and compared based on their results in training, checking/verification and testing data sets. The results are summarized in Table 5.11.

Table 5.11

Performance Evaluation Result

Model	RMSE			MAE		
	Training	Checking	Testing	Training	Checking	Testing
BP NN	1.10532	1.26552	1.21812	0.84562	0.74562	0.96541
ANFIS	0.38246	0.66845	0.38246	0.37408	0.62738	0.37408
ARIMA	0.91234	0.89978	0.97688	0.76451	0.92167	0.89253

Figure 5.11 shows the RMSE and MAE of BP NN, ANFIS and ARIMA models. It shows the RMSE value of the ANFIS model is the lowest compared to BP NN and ARIMA models. Overall, the results of this study recommend that the ANFIS model is better compared to the BP NN and ARIMA models for modelling reservoir water release decision.

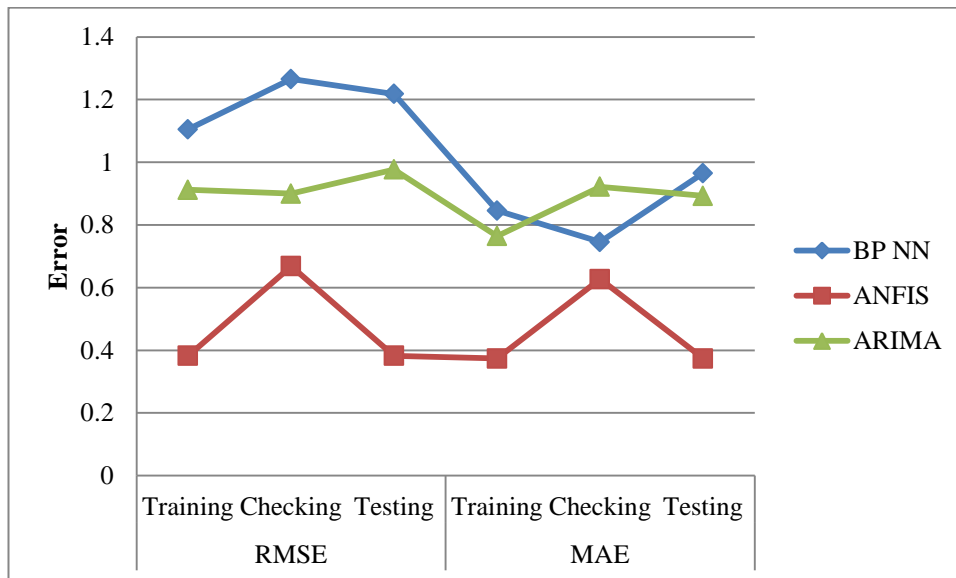


Figure 5.11. RMSE and MAE

On the other hand, the value of MAE of ANFIS model is also lowest compared to BP NN and ARIMA models. Overall, the results for training, testing and checking/validation of the models demonstrate that ANFIS model provides better performances than BP NN and ARIMA models for making accurate decision regarding water release according to the criteria.

5.4 Reservoir Water Release Dynamic Decision Model

The architecture of reservoir water release dynamic decision model adopted the ANFIS architecture which consists of five layers as shown in Figure 5.12. The first layer describes the 23 number of input variables (rainfall pattern from five gauging stations and reservoir water level). Layer 2 represents 23 memberships function constructed based on 23 input variables. Layer 3 produces rule generated from training the model which consist of three rules. Layer 4 represents the output for

membership function that has been constructed in Layer 2. The fifth layer generates the final output (gate opening at the time $t+i$).

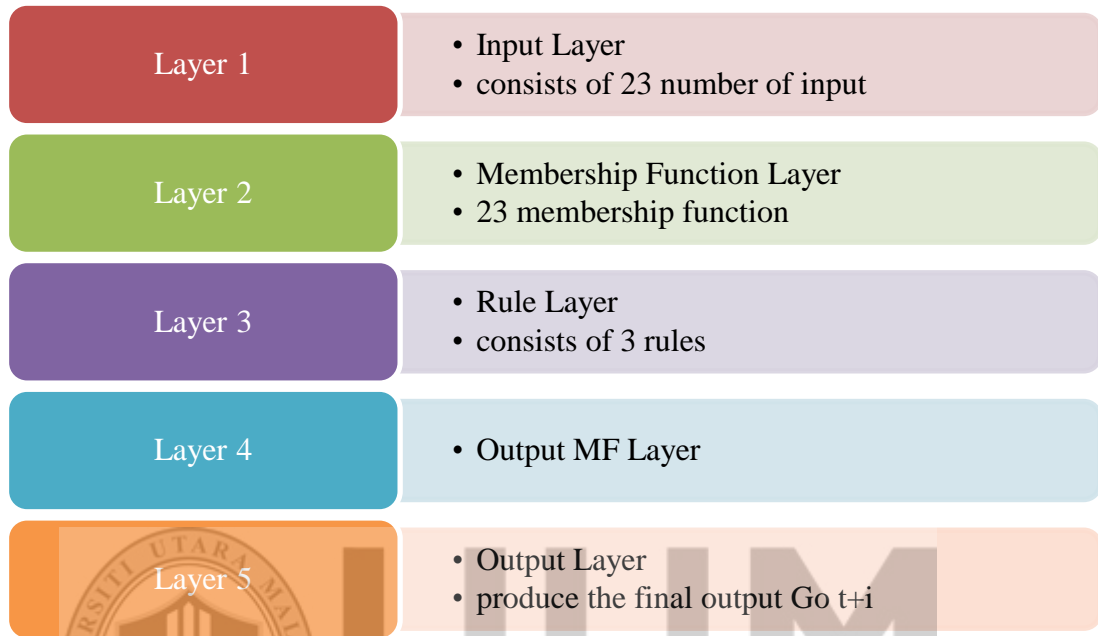


Figure 5.12. Five Layers of ANFIS Model

As mentioned in chapter one, the development of ANFIS model solves the issues of interpretability by providing the rules to explain the decision made from the knowledge embedded in the trained ANFIS. There are two ANFIS models have been developed for modelling reservoir water release decision which has been discussed in the previous section. Based on the finding, Model 1 with 23 input variables provides better performance compared to Model 2 with 22 input variables. Overall, it can be summarized here that Model 1 is selected as the best model in modelling reservoir water release decision and Model 1 produce rules as the following:

If (in1 is in1cluster1) and (in2 is in2cluster1) and (in3 is in3cluster1) and (in4 is in4cluster1) and (in5 is in5cluster1) and (in6 is in6cluster1) and (in7 is in7cluster1) and (in8 is in8cluster1) and (in9 is in9cluster1) and (in10 is in10cluster1) and (in11 is in11cluster1) and (in12 is in12cluster1) and (in13 is in13cluster1) and (in14 is in14cluster1) and (in15 is in15cluster1) and (in16 is in16cluster1) and (in17 is in17cluster1) and (in18 is in18cluster1) and (in19 is in19cluster1) and (in20 is in20cluster1) and (in21 is in21cluster1) and (in22 is in22cluster1) and (in23 is in23cluster1) then (out1 is out1cluster1) (1)

Else if

If (in1 is in1cluster2) and (in2 is in2cluster2) and (in3 is in3cluster2) and (in4 is in4cluster2) and (in5 is in5cluster2) and (in6 is in6cluster2) and (in7 is in7cluster2) and (in8 is in8cluster2) and (in9 is in9cluster2) and (in10 is in10cluster2) and (in11 is in11cluster2) and (in12 is in12cluster2) and (in13 is in13cluster2) and (in14 is in14cluster2) and (in15 is in15cluster2) and (in16 is in16cluster2) and (in17 is in17cluster2) and (in18 is in18cluster2) and (in19 is in19cluster2) and (in20 is in20cluster2) and (in21 is in21cluster2) and (in22 is in22cluster2) and (in23 is in23cluster2) then (out1 is out1cluster2) (1)

Else if

If (in1 is in1cluster3) and (in2 is in2cluster3) and (in3 is in3cluster3) and (in4 is in4cluster3) and (in5 is in5cluster3) and (in6 is in6cluster3) and (in7 is in7cluster3) and (in8 is in8cluster3) and (in9 is in9cluster3) and (in10 is in10cluster3) and (in11 is in11cluster3) and (in12 is in12cluster3) and (in13 is in13cluster3) and (in14 is in14cluster3) and (in15 is in15cluster3) and (in16 is in16cluster3) and (in17 is in17cluster3) and (in18 is in18cluster3) and (in19 is in19cluster3) and (in20 is in20cluster3) and (in21 is in21cluster3) and (in22 is in22cluster3) and (in23 is in23cluster3) then (out1 is out1cluster3) (1)

Reference:

Input cluster

in1 is in1cluster1 = input1 is
input1cluster1

in1 is in1cluster2= input1 is
input1cluster2

in1 is in1cluster3= input1 is
input1cluster3

Output cluster

out1 is out1cluster1= output1 is
output1cluster1

out1 is out1cluster2= output1 is
output1cluster2

out1 is out1cluster3= output1 is
output1cluster3

Figure 5.13 shows the input and output clustering using subtractive clustering method. Based on the rules produced by ANFIS, there are three main clusters which have been automatically identified in order to cluster the input and output (refer to the rules reference). The subtractive clustering usually can automatically verify the number of clustering for the data sets.

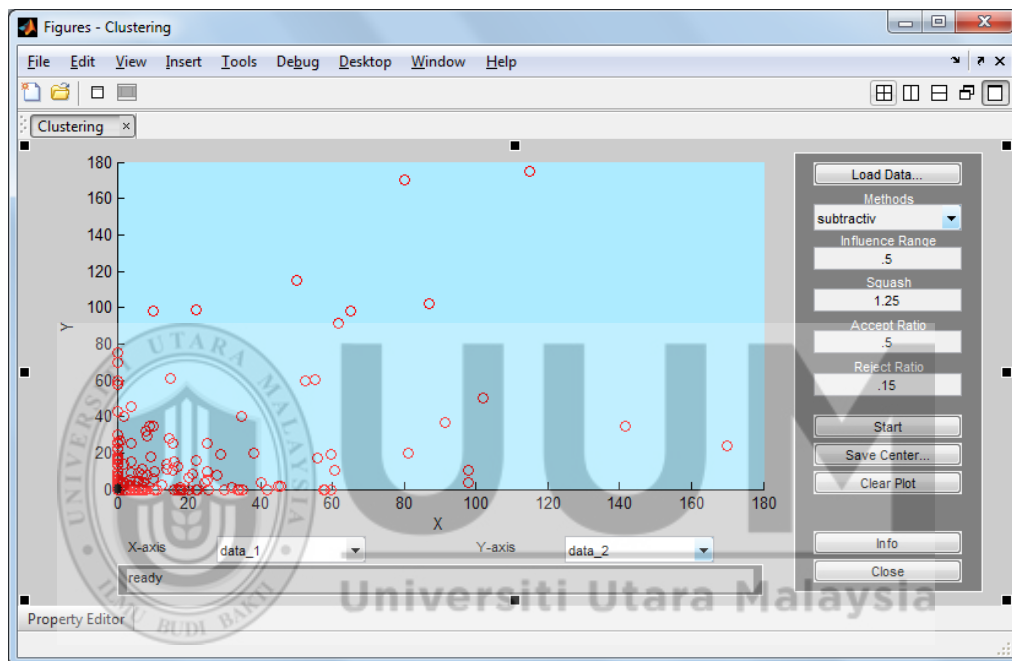


Figure 5.13. Input and Output Clustering

Figure 5.14 shows the architecture of the reservoir water release dynamic decision model which has been developed using the ANFIS.

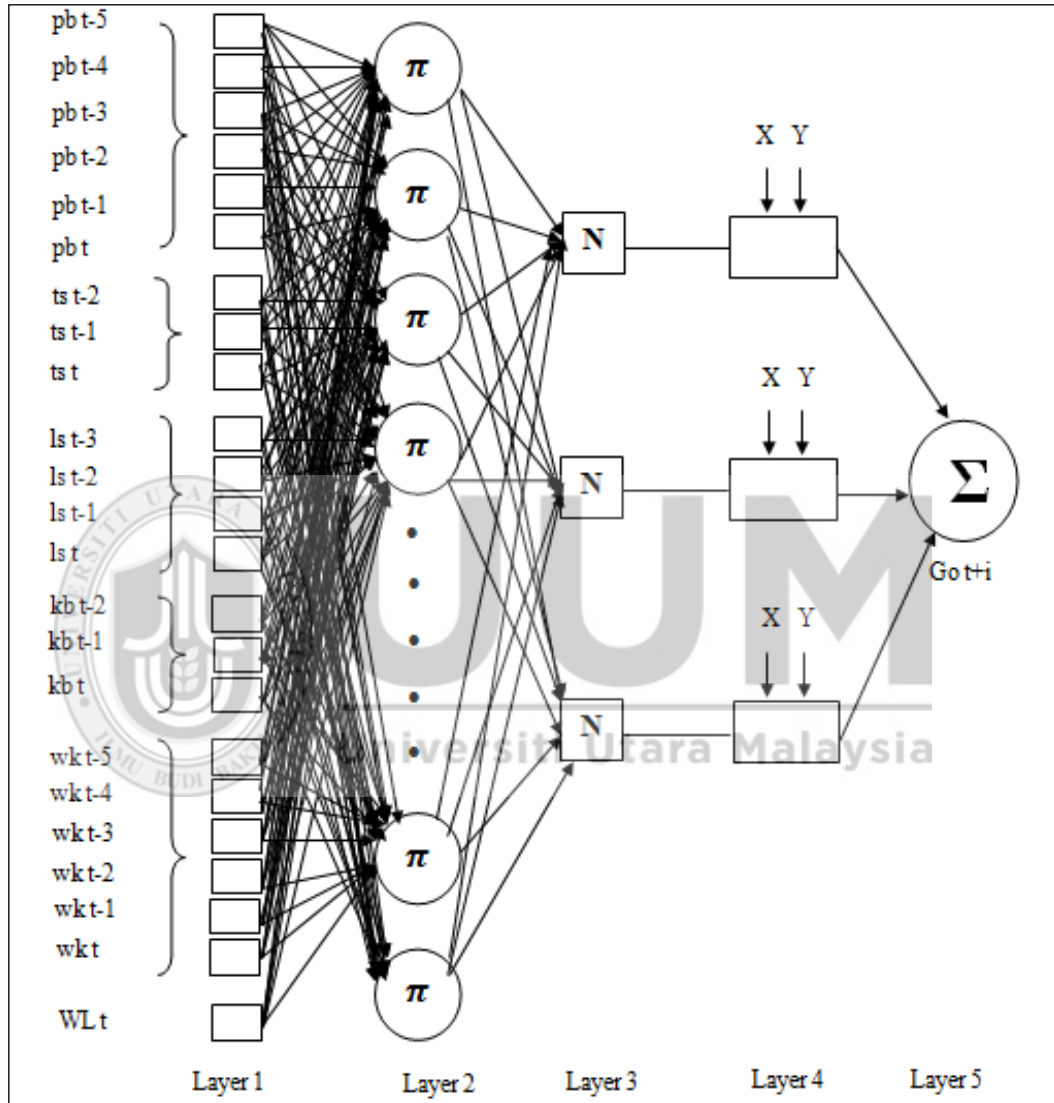


Figure 5.14. The Architecture of Reservoir Water Release Dynamic Decision Model

5.5 Summary

In this chapter, the experimental results and findings of the study have been discussed. The experiment was carried out to identify the size of sliding window that represents the time of delay. In addition, the findings show different time of delays for water to flow into the reservoir from five different gauging stations. The experiment on development reservoir water release dynamic decision model was carried out by using the ANFIS. Two ANFIS models have been developed for identifying the best model's performances. In the next chapter, the conclusion and future work will be presented as guidelines or direction to the researchers for making enhancement in the future.



CHAPTER SIX

CONCLUSION AND FUTURE WORK

This chapter will review all the findings found from this study by summarizing all the work done based on the research objectives. There are many factors that need to be considered for developing an efficient and effective reservoir water release dynamic decision model due to the complexity of the domain. This study highlights and recommends further work, especially in developing a decision model that consider other variables and factors for improving reservoir operations.

6.1 Contribution of the Study

The major contribution of this study can be summarized as follows:

- a. Spatial temporal pattern of reservoir water release decision has been extracted from reservoir operational data. This process involves data mining tasks such as data cleaning, data transformation and data normalization. This spatial temporal pattern was extracted by using data mining techniques. The modified sliding window algorithm was used to construct the rainfall temporal pattern, while spatial information was established by simulating the mapped rainfall and reservoir water level pattern by using neural network. In addition, spatial temporal pattern extraction procedure has been determined to identify both spatial temporal relationships that exist in modelling reservoir water release decision.

- b. A reservoir water release dynamic decision model was developed for improving reservoir operation during emergency situations. This model integrates spatial and temporal data mining approaches that utilize both spatial temporal information in the input pattern in order to determine when the reservoir water should be released. This model has been developed using a hybrid computational intelligence technique, specifically ANFIS (Adaptive Neuro- Fuzzy Inference System). It is found that the performance of the model has a small number of errors compared to the others techniques.

6.2 Future Work

There are a lot of areas for enhancement that can be made from this study. Basically, future works can be made from spatial temporal data mining and decision variables involved in the modelling techniques.

6.2.1 Spatial Temporal Data Mining

In this study, spatial temporal data mining technique has been applied to extract spatial temporal pattern from reservoir operational data. However, the spatial temporal relationship between the reservoir and gauging stations has been investigated by using temporal information, where other features of spatial data such as distance, slope, lines, polygons etc have not been considered. As for the techniques, suggestions for improvements are as follows:

- a. Consider spatial features such as distance between gauging stations and the reservoir to identify accurate time of delays for water to flow from the gauging station into the reservoir.
- b. The system of reservoir may be different from one to another. Thus, a comprehensive modelling needs to consider a system of reservoirs with their spatial relationship in order to improve the accuracy of the model.

6.2.2 Decision Variables Involved of the Modelling Techniques

In a reservoir operation, the existence of many decision variables associated with the operation models and complex structure of the reservoir system become one of the constraints for researchers to identify the appropriate methodology to solve problems that are related to reservoir operation and management. In real-time reservoir operations, there are many factors and variables that need to be considered in order to determine accurate reservoir water release decision. However, there is always a space for improvement in terms of accuracy, operational complexities, and decision variables involved in the modelling techniques. Future work can be summarized as follows:

- a. Other enhancement can be made either to include the other hydrology variables such as sediment, evaporation rate, river inflow etc. that could also improve the performance of the reservoir water release decision model. This effort needs to be done by working closely with the domain expert.

- b. Integrate both forecasting model and water release decision model to enhance predictions of future decision. This integration model will provide a better performance compared to a single model for improving reservoir operations.

6.3 Summary

Reservoir water release decision is one of the challenging task for reservoir operators since the decision is very critical to save human lives or to avoid water shortage problems in the future. Thus, reservoir water release dynamic decision model needs to be developed in order to support early and accurate water release decisions. This model integrates both spatial and temporal data mining as an approach to extract spatial temporal pattern from the reservoir operational data. This study demonstrates that spatial temporal relationship which exists in modelling reservoir water release decision has been identified.

This decision model was also developed based on the concept called dynamic decision making. This concept shows how humans naturally make decisions in emergency situations where the decisions made do not only depend on a set of alternatives, but also based on their experiences. The proposed model applied this concept in order to learn from experts on how to apply their past experiences to make accurate water release decisions. In addition, this model will hopefully improve the performances of reservoir operation and management in the future.

REFERENCES

- Ahn, H., & Kim, K. (2009). Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. *Applied Soft Computing*, 9(2), 599–607. doi:10.1016/j.asoc.2008.08.002
- Ahmed, J. A., & Sarma, A. K. (2005). Genetic Algorithm for Optimal Operating Policy of a Multipurpose Reservoir. *Water Resources Management*, 19 (2), 145–161. doi:10.1007/s11269-005-2704-7
- Azuma, R., Daily, M., & Furmanski, C. (2006). A review of time critical decision making models and human cognitive processes. In *Aerospace Conference, 2006 IEEE* (pp. 9-pp). IEEE
- Beach, L.R., & Mitchell, T.R. (1978). A Contingency Model for the Selection of Decision Strategies. *The Academy of Management Review*, Vol. 3, No. 3 (Jul., 1978), pp. 439-449, Retrieved from: <http://www.jstor.org/stable/257535>
- Becerra-Fernandez, I., Xia, W., Gudi, A., & Rocha, J. (2008, May). Task characteristics, knowledge sharing and integration, and emergency management performance: research agenda and challenges. In *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2008)* (pp. 88-92).
- Bohanec, M. (2001). "What is Decision Support?". *Proc. Information Society IS-2001: Data Mining and Decision Support in Action!*, (eds. Škrjanc, M., Mladenić, D.), Ljubljana, 86-89
- Bower, B. T., Hufschmidt, M. M., and Reedy, W. W. (1962), Operating procedures: Their role in the design of water-resources systems by simulation analyses, in *Design of Water -Resource Systems*, edited by A. Maass et al., pp. 443–458, Harvard Univ. Press, Cambridge, Mass
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81(3), 211–241. doi:10.1016/0001-6918(92)90019-A
- Bussemeyer, J. R. (1999). *Dynamic Decision Making*.
- Cancelliere, A., Giuliano, G., Ancarani, A., & Rossi, G. (2002). A Neural Networks Approach for Deriving Irrigation Reservoir Operating Rules. *Water Resources Management*, 16(1), 71–88. doi:10.1023/A:1015563820136
- Chandramouli, V., & Deka, P. (2005). Neural Network Based Decision Support Model for Optimal Reservoir Operation. *Water Resources Management*, 19(4), 447–464. doi:10.1007/s11269-005-3276-2
- Chang, F.-J., & Chen, L. (1998). Real-Coded Genetic Algorithm for Rule-Based Flood Control Reservoir Management. *Water Resources Management*, 12(3), 185–198. doi:10.1023/A:1007900110595
- Chang, F.-J., Chen, L., & Chang, L.-C. (2005). Optimizing the reservoir operating rule curves by genetic algorithms. *Hydrological Processes*, 19(11), 2277–2289. doi:10.1002/hyp.5674
- Chang, F.-J., & Chang, Y.-T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 29(1), 1–10. doi:10.1016/j.advwatres.2005.04.015
- Chaves, P., & Chang, F. J. (2008). Intelligent reservoir operation system based on evolving artificial neural networks. *Advances in Water Resources*, 31(6), 926-936.
- Chen, L., McPhee, J., & Yeh, W. W.-G. (2007). A diversified multiobjective GA for

- optimizing reservoir rule curves. *Advances in Water Resources*, 30(5), 1082–1093. doi:10.1016/j.advwatres.2006.10.001
- Chong, C. C., & Jia, J. C. (1994, October). Assessments of neural network output codings for classification of multispectral images using Hamming distance measure. In *Pattern Recognition, 1994. Vol. 2-Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on* (Vol. 2, pp. 526-528). IEEE.
- Club Managers Association of America (CMAA). (1999). *The Behavioral Theory of Decision Making*. Retrieved from <https://www.cmaa.org/bmiteam/decision/page4.asp>
- Cruz, A., & Mestrado, N. C. E. (2009). ANFIS: Adaptive Neuro-Fuzzy Inference Systems. *IM, UFRJ, Mestrado NCE*.
- Dawson, C. W., & Wilby, R. L. (2001). Hydrological modelling using artificial neural networks. *Progress in physical Geography*, 25(1), 80-108.
- Deka, P., & Chandramouli, V. (2005). Fuzzy neural network model for hydrologic flow routing. *Journal of Hydrologic Engineering*, 10(4), 302-314.
- Deka, P., & Chandramouli, V. (2009). Fuzzy Neural Network Modeling of Reservoir Operation. *Journal of Water Resources Planning and Management*, 135(1), 5–12. doi:10.1061/(ASCE)0733-9496(2009)135:1(5)
- Department of Irrigation and Drainage. (2012). Retrieved from <http://infobanjirperlis.dyndns.org/perlis/>
- Du, J., Qian, L., Rui, H., Zuo, T., Zheng, D., Xu, Y., & Xu, C. Y. (2012). Assessing the effects of urbanization on annual runoff and flood events using an integrated hydrological modeling system for Qinhuai River basin, China. *Journal of Hydrology*, 464, 127-139
- Dubrovin, T., Jolma, A., & Turunen, E. (2002). Fuzzy Model for Real-Time Reservoir Operation. *Journal of Water Resources Planning and Management*, 128(1), 66–73. doi:10.1061/(ASCE)0733-9496(2002)128:1(66)
- Draper, A., & Lund, J. (2004). Optimal Hedging and Carryover Storage Value. *Journal of Water Resources Planning and Management*, 130(1), 83–87. doi:10.1061/(ASCE)0733-9496(2004)130:1(83)
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.
- Feigh, K., & Pritchett, A. (2006). Design of Support Systems for Dynamic Decision Making in Airline Operations. In *2006 IEEE Systems and Information Engineering Design Symposium* (pp. 136–141). doi:10.1109/SIEDS.2006.278727
- Galavi, H., & Shui, L.T. (2012). Neuro-fuzzy modelling and forecasting in water resources. *Scientific Research and Essays*, 7(24). doi:10.5897/SRE11.2164
- Geetha, R., Sumathi, N., & Sathiabama, D. S. (2008). A Survey of Spatial, Temporal and Spatio-Temporal Data Mining. *Journal of Computer Applications*, 1(4), 31-33.
- Gilboa, I., & Schmeidler, D. (1995). Case-Based Decision Theory. *The Quarterly Journal of Economics*, 110(3), 605–639. doi:10.2307/2946694

- Gilboa, I., Schmeidler, D., & Wakker, P. P. (2002). Utility in Case-Based Decision Theory. *Journal of Economic Theory*, 105(2), 483–502. doi:10.1006/jeth.2001.2858
- Guo, X., Hu, T., Zeng, X., & Li, X. (2013). Extension of Parametric Rule with the Hedging Rule for Managing Multi-reservoir System during Droughts. *Journal of Water Resources Planning and Management*, 139(2), 139–148. doi:10.1061/(ASCE)WR.1943-5452.0000241
- Grosskopf, B., Sarin, R., & Watson, E. (2007). An Experiment on Case-Based Decision Making. Available at SSRN 1028442.
- Han, J., Kamber, M., & Pei, J. (2006). *Data Mining, Second Edition: Concepts and Techniques*. Morgan Kaufmann.
- Hashimoto, T., Stedinger, J. R., & Loucks, D. P. (1982). Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation. *Water Resources Research*, 18(1), 14–20. doi:10.1029/WR018i001p00014
- Hassin, M. H. B. M., Norwawi, N. B. M., & Aziz, A. (2006). Temporal Case-Based Reasoning for reservoir spillway gate operation. *International Conference on Computing Informatics, 2006. ICOCI '06* (pp. 1–4). Presented at the International Conference on Computing Informatics, 2006. ICOCI '06. doi:10.1109/ICOCI.2006.5276517
- Hejazi, M. I., Ruddell, B. L., & Cai, X. (2008). How Reservoirs Were Operated — Exploring the Role of Hydrologic Information. In *World Environmental and Water Resources Congress 2008* (pp. 1–9). American Society of Civil Engineers. Retrieved from <http://ascelibrary.org/doi/abs/10.1061/40976%28316%29593>
- Hersh, M. A. (1999). Sustainable decision making: the role of decision support systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 29(3), 395–408. doi:10.1109/5326.777075
- Hüllermeier, E. (2005). Experience-Based Decision Making: A Satisficing Decision Tree Approach. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 35(5), 641–653. doi:10.1109/TSMCA.2005.851145
- Hüllermeier, E. (2007). Case-Based Decision Making. In *Case-Based Approximate Reasoning* (pp. 253–305). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/1-4020-5695-8_7
- Husain, A. (2012). An Overview of Reservoir Systems Operation Techniques.
- Jain, S. K., & Singh, V. P. (2003). Chapter 11: Reservoir Operation. *Water Resources Systems Planning & Management*, 51, 615–679
- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on*, 23(3), 665–685.
- Jothiprakash, V., & Shanthi, G. (2006). Single Reservoir Operating Policies Using Genetic Algorithm. *Water Resources Management*, 20(6), 917–929. doi:10.1007/s11269-005-9014-y
- Jothiprakash, V., & Kote, A. S. (2011). Improving the performance of data-driven techniques through data pre-processing for modelling daily reservoir inflow. *Hydrological Sciences Journal*, 56(1), 168–186. doi:10.1080/02626667.2010.546358
- Jothiprakash, V., Shanthi, G., & Arunkumar, R. (2011). Development of Operational Policy for a Multi-reservoir System in India using Genetic Algorithm. *Water*

- Resources Management*, 25(10), 2405–2423. doi:10.1007/s11269-011-9815-0
- Kaiornrit, J., Wong, K. W., & Fung, C. C. (2013). A modular technique for monthly rainfall time series prediction. In *2013 IEEE Symposium on Computational Intelligence in Dynamic and Uncertain Environments (CIDUE)* (pp. 76–83). doi:10.1109/CIDUE.2013.6595775
- Kalyani, D., & Chaturvedi, S. K. (2012). A Survey on Spatio-Temporal Data Mining. *International Journal of Computer Science and Network (IJCSN). Volume 1, Issue 4, August 2012* www.ijcsn.org ISSN 2277-5420
- Karaboga, D., Bagis, A., & Haktanir, T. (2004). Fuzzy Logic Based Operation of Spillway Gates of Reservoirs during Floods. *Journal of Hydrologic Engineering*, 9(6), 544–549. doi:10.1061/(ASCE)1084-0699(2004)9:6(544)
- Karaboga, D., Bagis, A., & Haktanir, T. (2008). Controlling spillway gates of dams by using fuzzy logic controller with optimum rule number. *Applied Soft Computing*, 8(1), 232-238
- Kim, T., Heo, J.-H., Bae, D.-H., & Kim, J.-H. (2008). Single-reservoir operating rules for a year using Multiobjective Genetic Algorithm. *Journal of Hydroinformatics*, 10(2), 163. doi:10.2166/hydro.2008.019
- Koutroulis, A. G., Tsanis, I. K., Daliakopoulos, I. N., & Jacob, D. (2013). Impact of climate change on water resources status: A case study for Crete Island, Greece. *Journal of Hydrology*, 479, 146–158. doi:10.1016/j.jhydrol.2012.11.055
- Klein, G. A. (1993). A Recognition-Primed Decision (RPD) Model of Rapid Decision making. In *Decision Making in Action: Models and Methods*, G. A. Klein, J. Orasanu, R. Calderwood, and C. E. Zsombok (eds.), ISBN 0-89391-794-X, pp.138 - 147, 1993.
- Klein, G., & Klinger, D. (1991). Naturalistic decision making. *Human Systems IAC Gateway*, 11(3), 16-19.
- Klein, G. (2008). Naturalistic Decision Making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 456–460. doi:10.1518/001872008X288385
- Kudryavtsev, A., & Pavlodsky, J. (2012). Description-based and experience-based decisions: individual analysis. *Judgment and Decision Making*, 7(3), 316-331.
- Ku-Mahamud, K. R., Zakaria, N., Katuk, N., & Shbier, M. (2009). Flood Pattern Detection Using Sliding Window Technique. In *Third Asia International Conference on Modelling Simulation, 2009.AMS '09* (pp. 45–50). Presented at the Third Asia International Conference on Modelling Simulation, 2009.AMS '09. doi:10.1109/AMS.2009.15
- Labadie, J. (2004). Optimal Operation of Multireservoir Systems: State-of-the-Art Review. *Journal of Water Resources Planning and Management*, 130(2), 93–111. doi:10.1061/(ASCE)0733-9496(2004)130:2(93)
- Laxman, S., & Sastry, P. S. (2006). A survey of temporal data mining. *Sadhana*, 31(2), 173–198. doi:10.1007/BF02719780
- Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. *Department of Social and Decision Sciences*, 88.

- Li, B. (2008). The Classical Model of Decision Making Has Been Accepted as not providing an Accurate Account of How People Typically Make Decisions. *International Journal of Business and Management*.
- Lin, W., Orgun, M. A., and Williams, G. J. (2002). An Overview of Temporal Data Mining. In *Proceedings of the 1st Australian Data Mining Workshop (ADM'02)*, pages 83–90, Canberra, Australia
- Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, 14(5), 331-352.
- Liu, X., Guo, S., Liu, P., Chen, L., & Li, X. (2011). Deriving Optimal Refill Rules for Multi-Purpose Reservoir Operation. *Water Resources Management*, 25(2), 431–448. doi:10.1007/s11269-010-9707-8
- Luo, J., Qi, Y., Xie, J., & Zhang, X. (2015). A hybrid multi-objective PSO–EDA algorithm for reservoir flood control operation. *Applied Soft Computing*, 34, 526-538.
- Mehta, R., & Jain, S. K. (2009). Optimal Operation of a Multi-Purpose Reservoir Using Neuro-Fuzzy Technique. *Water Resources Management*, 23(3), 509–529. doi:10.1007/s11269-008-9286-0
- Mohan, A., & Revesz, P. (2012). Temporal data mining of uncertain water reservoir data. In *Proceedings of the Third ACM SIGSPATIAL International Workshop on Querying and Mining Uncertain Spatio-Temporal Data* (pp. 10–17). New York, NY, USA: ACM. doi:10.1145/2442985.2442987
- Momtahn, S., & Dariane, A. B. (2007). Direct Search Approaches Using Genetic Algorithms for Optimization of Water Reservoir Operating Policies. *Journal of Water Resources Planning and Management*, 133(3), 202–209. doi:10.1061/(ASCE)0733-9496(2007)133:3(202)
- Majumdar, V., & Mujumdar, S. V. P. P. (2005). *Water Resources Systems*. Tata McGraw-Hill. Retrieved from <http://books.google.com.my/books?id=oN87XeoHHM0C>
- Manikandan, G., & Srinivasan, S. (2013). An Efficient Algorithm For Mining Spatially Co-Located Moving Objects. *American Journal of Applied Sciences*, 10(3).
- MIKE Basin – An Integrated River Basin Planning and Management Tool. (n.d). Retrieved from <http://www.crwr.utexas.edu/gis/gishyd98/dhi/mikebas/Mbasbody.htm>
- Naik, P. K., & Jay, D. A. (2011). Distinguishing human and climate influences on the Columbia River: Changes in mean flow and sediment transport. *Journal of Hydrology*, 404(3–4), 259–277. doi:10.1016/j.jhydrol.2011.04.035
- Neelakantan, T. R., & Pundarikanthan, N. V. (1999). Hedging Rule Optimisation for Water Supply Reservoirs System. *Water Resources Management*, 13(6), 409–426. doi:10.1023/A:1008157316584
- Neelakantan, T., & Pundarikanthan, N. (2000). Neural Network-Based Simulation-Optimization Model for Reservoir Operation. *Journal of Water Resources Planning and Management*, 126(2), 57–64. doi:10.1061/(ASCE)0733-9496(2000)126:2(57)
- Ngo, L.L. (2006). Optimising reservoir operation: A case study of the HoaBinh reservoir, Vietnam. PhD Thesis, *Institute of Environment & Resources Technical University of Denmark*, 2006

- Ngoc, T. A., Hiramatsu, K., & Harada, M. (2013). Optimizing the rule curves of multi-use reservoir operation using a genetic algorithm with a penalty strategy. *Paddy and Water Environment*, 1–13. doi:10.1007/s10333-013-0366-2
- Norita, M. N. (2004). Computational Recognition-Primed Decision Model Based on Temporal Data Mining Approach in a Multiagent Environment for Reservoir Flood Control Decision (*Doctoral dissertation, Universiti Utara Malaysia*).
- Phillips-Wren, G. (2009). Adaptive Decision Support for Dynamic Environments. In K. Nakamatsu, G. Phillips-Wren, L. C. Jain, & R. J. Howlett (Eds.), *New Advances in Intelligent Decision Technologies* (pp. 235–243). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-00909-9_23
- Oliveira, R., & Loucks, D. P. (1997). Operating rules for multireservoir systems. *Water Resources Research*, 33(4), 839–852. doi:10.1029/96WR03745
- Ossadnik, W., Wilmsmann, D., & Niemann, B. (2013). Experimental evidence on case-based decision theory. *Theory and decision*, 75(2), 211–232.
- Othman, F., & Naseri, M. (2011). Reservoir inflow forecasting using artificial neural network. *International Journal of Physical Sciences*, 6(3), 434–440.
- Palit, A. K., & Popovic, D. (2006). *Computational intelligence in time series forecasting: theory and engineering applications*. Springer.
- Panigrahi, D. P., & Mujumdar, P. P. (2000). Reservoir Operation Modelling with Fuzzy Logic. *Water Resources Management*, 14(2), 89–109. doi:10.1023/A:1008170632582
- Pinthong, P., Gupta, A. D., Babel, M. S., & Weesakul, S. (2009). Improved Reservoir Operation Using Hybrid Genetic Algorithm and Neurofuzzy Computing. *Water Resources Management*, 23(4), 697–720. doi:10.1007/s11269-008-9295-z
- Prasad, M. A., & Mohan, S. (2005). Fuzzy Logic Model for Multi — Purpose Multi - Reservoir System. In D. Li & B. Wang (Eds.), *Artificial Intelligence Applications and Innovations* (pp. 335–347). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/0-387-29295-0_36
- Raje, D., & Mujumdar, P. P. (2010). Reservoir Performance under Uncertainty in Hydrologic Impacts of Climate Change. *Advances in Water Resources*, 33(3), 312–326. doi:10.1016/j.advwatres.2009.12.008
- Rakow, T., & B Rahim, S. (2010). Developmental insights into experience-based decision making. *Journal of Behavioral Decision Making*, 23(1), 69–82.
- Raman, H., & Chandramouli, V. (1996). Deriving a general operating policy for reservoirs using neural network. *Journal of water resources planning and management*, 122(5), 342–347
- Ramani Bai, V, Woo, C. S., Othman, F, & Ramadas, G. (2007). Data Mining In Reservoir Operation and Flood Control Using Artificial Neural Networks. *Proceeding of the 2nd International Conference on Informatics, 2007*. Retrieved February 15, 2013, from <http://myais.fsktm.um.edu.my/1467/>
- Rani, D., & Moreira, M. M. (2010). Simulation–Optimization Modeling: A Survey and Potential Application in Reservoir Systems Operation. *Water Resources Management*, 24(6), 1107–1138. doi:10.1007/s11269-009-9488-0

- Rashid, A. N. M. B., & Hossain, M. A. (2012). Challenging Issues of Spatio-Temporal Data Mining. *Computer Engineering and Intelligent Systems*, 3(4), 55–63.
- Reservoir. (2012). *Dictionary.com Unabridged*. Retrieved Dec 04, 2012, from Dictionary.com website: <http://dictionary.reference.com/browse/reservoir>
- Rittima, A. (2009). Hedging Policy for Reservoir System Operation: A Case Study of Mun Bon and Lam Chae Reservoirs. *Kasetsart Journal, Natural Sciences*, 43(4), 833–842.
- Roddick, J. F., & Spiliopoulou, M. (2002). A survey of temporal knowledge discovery paradigms and methods. *IEEE Transactions on Knowledge and Data Engineering*, 14(4), 750–767. doi:10.1109/TKDE.2002.1019212
- Russell, S. O., & Campbell, P. F. (1996). Reservoir Operating Rules with Fuzzy Programming. *Journal of Water Resources Planning and Management*, 122(3), 165–170. doi:10.1061/(ASCE)0733-9496(1996)122:3(165)
- Salehi, A., & Afshar, A. (2011). Gated Spillways Operation Rules Considering Water Surface Elevation and Flood Peak: Application to Karkheh Dam. In *World Environmental and Water Resources Congress 2011* (pp. 3007–3015). American Society of Civil Engineers. Retrieved from <http://ascelibrary.org/doi/abs/10.1061/41173%28414%29314>
- Shahnawaz, M., Ranjan, A., & Danish, M. (2011). Temporal Data Mining: An Overview. *International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-1, October 2011*
- Sharifi, F., Haddad, O. B., & Naderi, M. (2005). Reservoir optimal operation using DP-ANN. *Evolutionary Computation*, 496-281.
- Shekhar, S., Zhang, P., & Huang, Y. (2005). Spatial Data Mining. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 833–851). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/0-387-25465-X_39
- Shekhar, S., Zhang, P., & Huang, Y. (2010). Spatial Data Mining. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 837–854). Springer US. Retrieved from http://link.springer.com/chapter/10.1007/978-0-387-09823-4_43
- Shiau, J. (2009). Optimization of Reservoir Hedging Rules Using Multiobjective Genetic Algorithm. *Journal of Water Resources Planning and Management*, 135(5), 355–363. doi:10.1061/(ASCE)0733-9496(2009)135:5(355)
- Shiau, J. T., & Lee, H. C. (2005). Derivation of Optimal Hedging Rules for a Water-supply Reservoir through Compromise Programming. *Water Resources Management*, 19(2), 111–132. doi:10.1007/s11269-005-1502-6
- Shih, J., & ReVelle, C. (1994). Water Supply Operations during Drought: Continuous Hedging Rule. *Journal of Water Resources Planning and Management*, 120(5), 613–629. doi:10.1061/(ASCE)0733-9496(1994)120:5(613)
- Shih, J.-S., & ReVelle, C. (1995). Water supply operations during drought: A discrete hedging rule. *European Journal of Operational Research*, 82(1), 163–175. doi:10.1016/0377-2217(93)E0237-R
- Smith, K., & Ward, R. (1998). *Floods: Physical Processes and Human Impacts* (p. 394). Chichester: Wiley.

- Suriya, S., & Mudgal, B. V. (2012). Impact of urbanization on flooding: The Thirusoolam sub watershed – A case study. *Journal of Hydrology*, 412–413, 210–219. doi:10.1016/j.jhydrol.2011.05.008
- Srinivasan, K., & Philipose, M. C. (1996). Evaluation and selection of hedging policies using stochastic reservoir simulation. *Water Resources Management*, 10(3), 163–188. doi:10.1007/BF00424201
- Srinivasan, K., & Philipose, M. C. (1998). Effect of Hedging on Over-Year Reservoir Performance. *Water Resources Management*, 12(2), 95–120. doi:10.1023/A:1007936115706
- Talib, S. H. A., Yusoff, M. S., & Hasan, Z. A. (2012). Modeling of Sedimentation Pattern in Bukit Merah Reservoir, Perak, Malaysia. *Procedia Engineering*, 50, 201–210. doi:10.1016/j.proeng.2012.10.025
- Taghian, M., Rosbjerg, D., Haghghi, A., & Madsen, H. (2013). Optimization of Conventional Rule Curves Coupled with Hedging Rules for Reservoir Operation. *Journal of Water Resources Planning and Management*, doi:10.1061/(ASCE)WR.1943-5452.0000355
- Tu, M.-Y., Hsu, N.-S., & Yeh, W. W.-G. (2003). Optimization of Reservoir Management and Operation with Hedging Rules. *Journal of Water Resources Planning and Management*, 129(2), 86–97. doi:10.1061/(ASCE)0733-9496(2003)129:2(86)
- Tu, M.-Y., Hsu, N.-S., Tsai, F. T.-C., & Yeh, W. W.-G. (2008). Optimization of Hedging Rules for Reservoir Operations. *Journal of Water Resources Planning and Management*, 134(1), 3–13. doi:10.1061/(ASCE)0733-9496(2008)134:1(3)
- Tucci, C. E. (2002). Flood Flow Forecasting. Presented at 54th session of Executive Council of WMO World Meteorological Organization, Geneva
- U.S. Army Corps of Engineers (USACE). (1991). Optimization of Multiple-Purpose Reservoir System Operations: A Review of Modeling and Analysis Approaches, Hydrologic Engineering Center, R.A. Wurbs, RD-34.
- Valizadeh, N., & El-Shafie, A. (2013). Forecasting the Level of Reservoirs Using Multiple Input Fuzzification in ANFIS. *Water Resources Management*, 27(9), 3319–3331. doi:10.1007/s11269-013-0349-5
- Valizadeh, N., El-Shafie, A., Mirzaei, M., Galavi, H., Mukhlisin, M., & Jaafar, O. (2014). Accuracy Enhancement for Forecasting Water Levels of Reservoirs and River Streams Using a Multiple-Input-Pattern Fuzzification Approach. *The Scientific World Journal*, 2014.
- Wang, W.-C., Chau, K.-W., Cheng, C.-T., & Qiu, L. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374(3–4), 294–306. doi:10.1016/j.jhydrol.2009.06.019
- Wang, B. (2012). Reservoir Characterization and Horizontal Well Placement Guidance Acquisition by Using GIS and Data Mining Methods. *Master Thesis. Department of Geomatics Engineering, University of Calgary, Alberta, 2012.*
- Wan Ishak, Wan Hussain and Ku-Mahamud, Ku Ruhana and Md. Norwawi, Norita (2010). Reservoir Water Level Forecasting Model Using Neural Network. *International Journal of Computational Intelligence*, 6 (4). pp. 947-952.

- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2011a). Reservoir Water Release Decision Modeling. *Proceedings of the 3rd International Conference on Computing and Informatics (8-9 Jun 2011, Bandung, Indonesia)*, pp: 74-79
- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2011b). Mining Temporal Reservoir Data Using Sliding Window Techniques. *CiiT International Journal of Data Mining Knowledge Engineering* ISSN 0974 – 9683 (Print) 0974 – 9578 (Online), Vol. 3, No. 8, (Jun 2011), pp: 473-478 DOI: DMKE062011004
- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2011c). Intelligent Decision Support Model Based on Neural Network to Support Reservoir Water Release Decision. *ICSECS 2011, Part I, Communications in Computer and Information Science (CCIS) 179*, pp. 365–379 ISBN: 978-3-642-22169-9. Springer LNCS
- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2011d). Neural Network Application in Reservoir Water Level Forecasting and Release Decision. *Proceedings of the 3rd International Conference on Computing and Informatics (8-9 Jun 2011, Bandung, Indonesia)*, pp: 74-79
- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2011e). Conceptual Model of Intelligent Decision Support System Based on Naturalistic Decision Theory for Reservoir Operation during Emergency Situation. *IJCEE: International Journal of Civil & Environmental Engineering*, 11(2), 06-11, ISSN: 2077-1258
- Wan Ishak, W. H., Ku Mahamud, K. R., & Md Norwawi, N. (2012). Modelling Reservoir Water Release Decision Using Temporal Data Mining and Neural Network. *International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459 (Online), Vol. 2, No. 8, pp: 422-428
- Wei, C.-C., & Hsu, N.-S. (2008). Derived operating rules for a reservoir operation system: Comparison of decision trees, neural decision trees and fuzzy decision trees. *Water Resources Research*, 44(2), W02428. doi:10.1029/2006WR005792
- Wurbs, R. (1993). Reservoir-System Simulation and Optimization Models. *Journal of Water Resources Planning and Management*, 119(4), 455–472. doi:10.1061/(ASCE)0733-9496(1993)119:4(455)
- Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, 12(3), 387–402. doi:10.3758/BF03193783
- Yeh, W. W.-G. (1985). Reservoir Management and Operations Models: A State-of-the-Art Review. *Water Resources Research*, 21(12), 1797–1818. doi:10.1029/WR021i012p01797
- You, J.-Y., & Cai, X. (2007). Determining Forecast and Decision Horizons for Reservoir Operations under Hedging Policies (pp. 1–10). American Society of Civil Engineers. doi:10.1061/40927(243)553
- You, J.-Y., & Cai, X. (2008a). Hedging rule for reservoir operations: 1. A theoretical analysis. *Water Resources Research*, 44(1), n/a–n/a. doi:10.1029/2006WR005481

- You, J.-Y., & Cai, X. (2008b). Hedging rule for reservoir operations: 2. A numerical model. *Water Resources Research*, 44(1), n/a–n/a. doi:10.1029/2006WR005482
- You, J., & Cai, X. (2009). Reexamination of Critical Period for Reservoir Design and Operation. *Journal of Water Resources Planning and Management*, 135(5), 392–396. doi:10.1061/(ASCE)0733-9496(2009)135:5(392)
- You, J.-Y. (2008). Hedging Rule for Reservoir Operation: How Much, When and how Long to Hedge. *PhD Thesis, Graduate College of the University of Illinois at Urbana-Champaign, 2008*



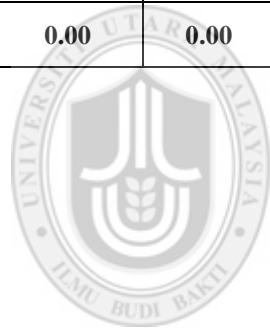
Appendix A

Sample of Reservoir Operational Data (Year 2007)

TARIKH	PARAS AIR (m)	Tail Water	CURAHAN HUJAN (mm)					LUAHAN MELALUI SPILLWAY			
			Padang Besar	Tasoh	Lubuk Sireh	Kaki Bukit	Wang Kelian	Jumlah gate di buka	Level gate di buka (mm)	Discharge In meter 3/s	Jumlah (MCM)
01-Jan-07	28.550	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0
02-Jan-07	28.530	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0
03-Jan-07	28.510	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0
04-Jan-07	28.505	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0
05-Jan-07	28.500	18.05	4.00	0.00	0.00	0.00	2.00	0	0	0	0
06-Jan-07	28.490	18.05	1.00	2.00	14.00	1.00	0.00	0	0	0	0
07-Jan-07	28.490	18.05	91.00	59.00	89.00	77.00	83.00	0	0	0	0
08-Jan-07	28.590	18.05	13.00	10.00	5.00	12.00	30.00	0	0	0	0
09-Jan-07	28.735	18.05	8.00	0.00	0.00	2.00	4.00	0	0	0	0
10-Jan-07	28.745	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0

11-Jan-07	28.745	18.05	0.00	0.00	0.00	0.00	0.00	0	0	0	0
12-Jan-07	28.750	17.50	0.00	1.00	0.00	0.00	0.00	0	0	0	0
13-Jan-07	28.760	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
14-Jan-07	28.760	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
15-Jan-07	28.745	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
16-Jan-07	28.740	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
17-Jan-07	28.725	17.50	1.00	0.00	1.00	0.00	0.00	0	0	0	0
18-Jan-07	28.725	17.50	0.00	1.00	0.00	0.00	0.00	0	0	0	0
19-Jan-07	28.720	17.50	0.00	1.00	0.00	0.00	0.00	0	0	0	0
20-Jan-07	28.710	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
21-Jan-07	28.700	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
22-Jan-07	28.700	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
23-Jan-07	28.700	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
24-Jan-07	28.690	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
25-Jan-07	28.680	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0

26-Jan-07	28.670	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
27-Jan-07	28.660	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
28-Jan-07	28.650	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
29-Jan-07	28.640	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
30-Jan-07	28.630	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0
31-Jan-07	28.620	17.50	0.00	0.00	0.00	0.00	0.00	0	0	0	0



UUM
Universiti Utara Malaysia

Appendix B

Sample of Hydrological Data (Year 2009)

RAINFALL FROM WANG KELIAN STATION

Day	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	2.5	0	10.5	0	16.5	0	0	21.5	3	0	0	4.5
2	16	0	25.5	5	15	5.5	0	0.5	0	1.5	0	2
3	1.5	0	20	6	17.5	0	0	0	0	8	0	0.5
4	0	0	15.5	24.5	19	0	57.5	5	0	0	30.5	0
5	0	0	1	0	9.5	0	0.5	19.5	22	44	73	0
6	0	0	46.5	1	17	0	0	5	2.5	2.5	58	0
7	0	0	0	7	14.5	0	3	8	0	8.5	0	0
8	4.5	0	8	1	5	0	0	0	0	15.5	0	0
9	0	0	1.5	2	1	0	0.5	0	0	0.5	0	0
10	0	0	0	89	0	0	1.5	0.5	15	21.5	0	0
11	0	0	0	1	0.5	0	104	1	2.5	0.9	8	0
12	0	0	0	0	0	0	25.5	0	0.5	7.6	81.5	0
13	0	16	0	0	17.5	0	0	0	0.5	0.5	19.5	0
14	0	0	0	4.5	3	9	0	2.5	2	0	0.5	0

15	0	0	0	35.5	5.5	10.5	6.5	0	41.5	13.5	0	0
16	0	0.5	0	0	7	2	3	2	3	2.5	1	7
17	0	0	39.7	0	6.5	0	0	4.5	7.5	0	5	1.5
18	0	0	3.8	0	1	0	0.5	53.5	12.5	19.5	4.5	35.5
19	0	0	2	2	0	0	0	2	15	13.5	5	0
20	0	0.5	41.5	1.5	45.5	0	5	0	0.5	3	26	0.5
21	0	3	1.5	17	0	12.5	34.5	4.5	0	5.5	37.5	0
22	0	0	0	0.5	0	0	10.5	16.5	0	0	48	0
23	0	0.5	6	12.5	0	0	5	59.5	0	0	17.5	0
24	0	0	3.5	5	0	0	0.5	36.5	10	0	0.5	0
25	0	1.5	9	4.5	0.5	0	0	0	5	9	0	0
26	0	0	0	0	0	0	0	0	28	22	0	0
27	0	3.5	2.5	17	0	3.5	0	28.5	0	23	0	0
28	0	1	32	14	2	0	5.5	35	0.5	25.5	0.5	0
29	0	8	1	39.5	0	0	20	44.5	0.5	0	0	
30	0	0	0	1.5	0	0	3	50	0	0	0	
31	0	0	0	1.5	20	0	18					