DEVELOPMENT OF FORECASTING IN SUNGAI MUDA, KUALA MUDA, KEDAH BY UTILIZING ARTIFICIAL NEURAL NETWORK (ANN).

NURUL MURSHIDA BINTI MOHD SABRI

Bachelor of Engineering (Hons) in Civil Engineering

UNIVERSITI MALAYSIA PAHANG

DEVELOPMENT OF FORECASTING IN SUNGAI MUDA, KUALA MUDA, KEDAH BY UTILIZING ARTIFICIAL NEURAL NETWORK (ANN).

NURUL MURSHIDA BINTI MOHD SABRI

Report submitted in fulfillment of the requirements for the award of the degree of

Bachelor of Engineering (Hons) in Civil Engineering

Faculty of civil Engineering and Earth Resources

UNIVERSITI MALAYSIA PAHANG

JUNE 2015

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS / UNDERGRADUATE PROJECT PAPER AND COPYRIGHT

Author's full name : **NURUL MURSHIDA BINTI MOHD SABRI**

Date of birth : **25 NOVEMBER 1992**

Title : **DEVELOPMENT OF FORECASTING IN SUNGAI MUDA, KUALA MUDA, KEDAH BY UTILIZING ARTIFICIAL NEURAL NETWORK (ANN).**

Academic Session : 2014/2015

I declare that this thesis is classified as:

☐	**CONFIDENTIAL**	(Contains confidential information under the Official Secret Act 1972)*
☐	**RESTRICTED**	(Contains restricted information as specified by the organization where research was done)*
√	**OPEN ACCESS**	I agree that my thesis to be published as online open access (full text)

I acknowledged that University Malaysia Pahang reserves the right as follows:

1. The thesis is the property of University Malaysia Pahang.
2. The Library of University Malaysia Pahang has the right to make copies for the purpose of research only.
3. The Library has the right to make copies of the thesis for academic exchange.

_______________	_______________
SIGNATURE	SIGNATURE OF SUPERVISOR
921125126632	DR. MUHAMMAD @ S.A. KHUSHREN BIN SULAIMAN
(NEW IC NO./PASSPORT NO.)	NAME OF SUPERVISOR
Date: 29 June 2015	Date: 29 June 2015

SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of degree of Bachelor of Civil Engineering and Earth Resources.

Signature :

Name of Supervisor : DR. MUHAMMAD @ S.A. KHUSHREN BIN SULAIMAN

Position : SENIOR LECTURER

Date : 29 JUNE 2015

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged. The thesis has not been accepted for any degree and is not concurrently submitted for award of other degree.

Signature :

Name : NURUL MURSHIDA BINTI MOHD SABRI

ID Number : AA11197

Date : 29 JUNE 2015

ACKNOWLEDGMENTS

In the name of Allah, the Beneficent, the Merciful. Alhamdulillah, thanks to Allah for the strength and endurance during the preparation of this thesis. First of all, I want to thank my supervisor, Dr. Muhammad @ S.A Khushren bin Sulaiman, who always patient to my lot of hassle and provide support and improvements in the production of this thesis. In addition, thanks to all my friends those who took under same supervisor (Amir Aliff Amri, Muhamad Afiq Mustafa, Wan Nurulhafizah Abd Razak and Nurul Qamar Roszaidi) also to my close friends (Fatin Nabilah Mender, Noridayu Zulkifli, Nor Shafiqah Omar and Nor Faezah Affri), who always accompanying me throughout the completion of this thesis. Not to forget, to my family, especially my mother, Puan Kartini Mangka and my father, Encik Mohd Sabri Zainol with their continuing of encouragement to me.

Finally, I want to apologize to those who recognize me, if any wrong words or behaviour intentionally or unintentionally during the preparation of this thesis. Everything I faced during the duration of completion of this thesis is a pleasure and invaluable experiences to me.

ABSTRACT

This report deals with flood problem which is usually happened in Malaysia when it coincides with monsoon and gave harm and damages to human life, as it had took many lives each time it happens. A case study of flood is going to be conduct to analyze the pattern of water level and to determine other causes that contributes to the flood. The main aim of the study is to minimize the effect of flood problems. It is also used to develop high accuracy model utilizing Artificial Neural Network (ANN) in predicting flood. Furthermore, it used to forecast flood occasion in the study area of station number of 5606410 of Sungai Muda (Jambatan Syed Omar) which is the main river that supplies water to Kedah and Penang. Besides, it used to investigate whether water level data alone can be used to produce modelling and to determine whether ANN is functioning in the forecasting. In this case study, a computational model will be used to stimulate the input data and generate the result, which is called Artificial Neural Network. ANN, which are modelled on the operating behaviour of the brain, are tolerant of some imprecision and are especially useful for classification and function approximation or mapping problems, to which hard and fast rules cannot be applied easily. The terminology of artificial neural networks has created form an organic biological model of neural system, which it comprises an asset of joined cells, the neurons. The neurons receive impulses or response from either input cells or any other neurons. It will perform some kind of transformation of the input and then, it will transfer the outcome to other neurons or also known as output cells. The neural networks are developed from many layers of connected neurons. The result showed that input 7+1 had the highest NSC value of 0.979 with RMSE value of 288.332 for 6 hour interval time, while input 6+1 had the highest NSC value of 0.977 with RMSE value of 134.801 for 3 hour interval time. In conclusion, this research contributes toward the development of forecasting using Artificial Neural Network for flood problems.

ABSTRAK

Laporan ini membincangkan masalah banjir yang kerap berlaku di Malaysia kebetulan dengan berlakunya monsun dan telah banyak memberikan kemudaratan dan kerosakan kepada kehidupan manusia, sehinggakan telah banyak mengambil nyawa manusia setiap kali ianya berlaku. Satu kajian kes banjir telah dijalankan untuk menganalisa corak paras air. Tujuan utama kajian ini adalah untuk menyelesaikan masalah banjir. Ianya juga digunakan untuk membangunkan model ketetapan yang tinggi dengan menggunakan ANN dalam meramalkan berlakunya banjir. Tambahan pula, ANN digunakan untuk meramal peristiwa banjir di kawasan kajian iaitu di stesen bernombor 5606410 Sungai Muda (Jambatan Syed Omar) yan merupakan sungai utama yang membekalkan air ke negeri Kedah dan Pulau Pinang. Selain itu, untuk menyiasat sama ada paras air sahaja boleh digunakan unruk menghasilkan model dan untuk menentukan sama ada ANN berfungsi dalam ramalan tersebut. Dalam kajian kes ini, satu model pengiraan akan digunakan untuk merangsang data input dan menjana hasil yang dipanggil *Artificial Neural Network* (ANN), yang mencontohi tingkah laku operasi otak. Ia adalah toleran terhadap beberapa ketakpersisan dan amat berguna untuk pengkelasan dan fungsi anggaran atau pemetaan masalah bahawa peraturan yang cepat dan susah tidak boleh diaplikasikan dengan mudah. Teminologi ANN ini telah mencipta model biologi organik sistem saraf yang terdiri daripada asset sel yang bergabung iaitu neuron. Neuron akan menerima impuls atau maklum balas daripada sama ada sel-sel input atau mana-mana neuron lain. Neuron tadi akan melaksanakan beberapa jenis transformasi input dan kemudian , ia akan memindahkan hasil untuk neuron lain atau juga dikenali sebagai sel-sel output. Rangkaian neural yang dibangunkan adalah hasil dari pelbagai lapisan neuron yang disambungkan. Keputusan menunjukan bahawa input 7+1 mempunyai nilai NSC tertinggi iaitu 0.979 dengan niai RMSE iaitu 288.332 untuk masa selang 6 jam, manakala input 6+1 mempunyai nilai NSC tertinggi iaitu 0.977 dengan nilai RMSE iaitu 134.801 untuk masa selang 3 jam. Kesimpulannya, kajian ini telah menyumbang ke arah pembangunan ramalan dengan menggunakan ANN untuk masalah banjir.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ANN	Artificial Neural Network.
FOR	Fixed Operating Regulations.
MRA	Multiple Regression Analysis.
NEM	North East Monsoon.
NSC	Nash-Sutch Coefficient.
RMSE	Root Mean Square Error.
SWM	South West Malaysia.

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Rain is a natural phenomenon which it can give benefit and advantages to people, animals and plants. Besides, it also can cool down the hot temperature, especially during hot day. Rain could be occur cause by frontal activity, convection and orographic effects and also human influence. Rain is formed through the process of condensation of water vapour which it rises up to cloud and when the cloud become saturated and heavy enough, then it will fall under gravity in form of droplets of water. Rainwater is important in water cycle and it is also the main contributors to the fresh water in Earth. Rainfall usually considered as mercy from God as it helps in agriculture and it is also an essential for human, animals and plants to survive in their life. But once it raining heavily, it can cause a big disaster to human life known as flood.

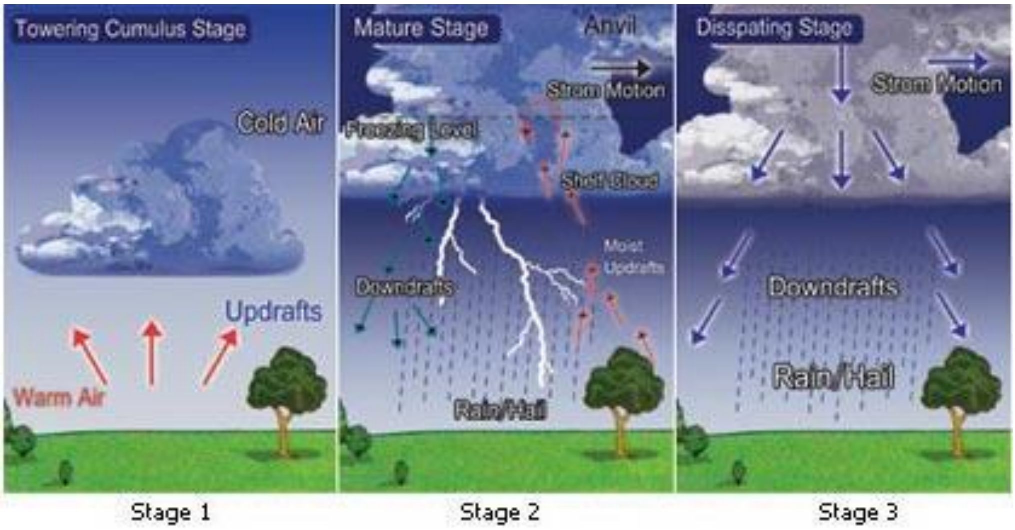


Figure 1.1: Formation of rain.

Flood is an event of overflow big quantity of water which it the soil cannot infiltrates more of the water. Flood also can occur when there is tide period. Tides are the effects of combination gravitational forces exerted by the Moon and the Sun and the rotation of the Earth in orbit which cause to rise and fall of sea levels. When the sea level start to raise, the flow rate of water will exceeding the capacity of water flow channel and this could cause spillway of rainfall. Therefore, when there is too much of rainfall, soil cannot absorb the rainfall because it is already fully saturated of water. In mean time, water level will increase which there is possibility of water level increase to a maximum point, exceed the normal water level. Flood disaster often cause harm and damages to living life, especially to human. For those living and doing business nearer river or other bodies of water will getting much the effect. This is because river is also another way in providing an easy travel and access to commerce industry.

Moreover, it is often to get news about flood problem in Malaysia, especially Peninsular Malaysia. In Peninsular Malaysia, they are having a humid tropical climate which the weather is warm and humid with a low temperature ranging from 20°C to 32°C and normal air temperature 37°C to 38°C is rarely recorded. There is a monthly definite variation that coincides with the monsoons (Jasim et al., 2013). The climate of Peninsular Malaysia meaningfully influences by the monsoons and experiences two rainy seasons throughout the year, associated with the South West Monsoon (SWM) from May to August and the North East Monsoon (NEM) from November to February (Wong et al., 2009).

As the case study that going to be conducted in Sungai Muda located in Kuala Muda, Kedah, with catchment basin area of 4,192 km^2. The population of citizen who is living in the area of basin is approximately to 220, 740 people. This event had caused a lot of damages and cost to citizen's life. A case study of flood by using Artificial Neural Network (ANN) will be conducted which it require a large number of water level data in study area to determine the water level that will produce flood and other causes of flood to occur. Daily water level data for 14 years is going to be selected based on the completeness of the water level data and the length of records. The input data that will be used in this case study are obtained from Drainage and Irrigation Department for the period from year 2000 to 2014.

Figure 1.2: Map of state of Kedah.

1.2 BACKGROUND

Artificial Neural Network also known as ANN was found in 1943, by a neurophysiologist, Warren McCulloh and a mathematician, Walter Pitts. This founding was happened during both of the neurophysiologist and mathematician wrote a paper on how neurons might work. Incidentally, in order to describe how neurons in brain are might work they have modelled a simple neural network using electrical circuit. From a biological neuron (Figure 1.3) they have developed an artificial neuron (perceptron) (Figure 1.4). It then, developed in 1949, by Donald Hebb, in 1959, by Bernard Widrow and Marcian Hoff of Stanford and in 1972, by Kohonen and Anderson were developed a similar network independently of one another.

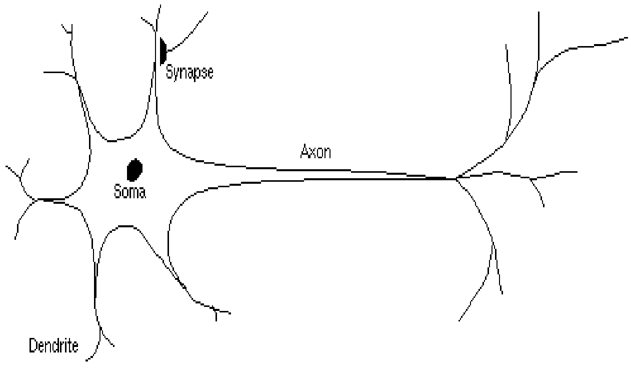


Figure 1.3: A biological neuron.

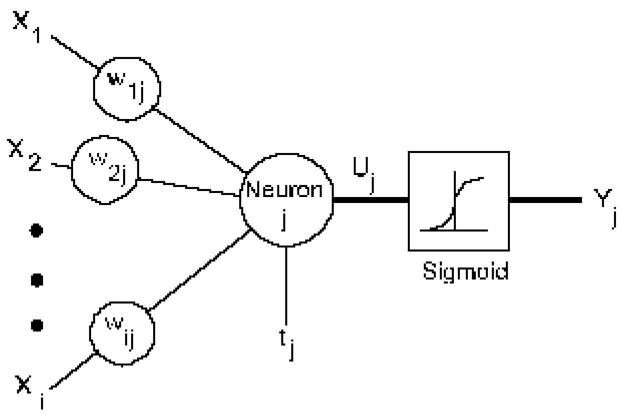


Figure 1.4: An artificial neuron (perceptron).

For this case study, data will be gathered and stimulate it by using Artificial Neural Network (ANN). This computational model was inspired based on animal’s central system which they prefer it as brain. It functions as estimation result based on a large number of input data and generally unknown. This model usually presented as system of interconnected “neurons”, which it means it can compute or generate values from the input data as well as the pattern of data that lead to the final result. In this case of study, rainfall data will be gathered in a specific study area, which flood often to occur. ANN is going to be used to stimulate data and generate the values to produce a result. Therefore, it can help in analyzing the hydrological data in determine the water level that could cause of flood and also in determine the relationship between water level and flood.

1.3 PROBLEM STATEMENT

When the temperature is getting too hot and the evaporation process of sea water could occur. As the cloud getting saturated and heavy enough to produce heavy rainfall, it will also cause metropolitan like Sungai Petani to be flooded. Besides, people will be in danger if there is none of the flood estimation equipment to control flooding in a certain place, especially places nearer to river or any body of water.

Furthermore, it is important to provide a prediction of flood as the flood estimation can give warning and alarm people to take precaution steps. Therefore, flood could occur anytime without any warning and this will lead transferring people to another safe place. Everything needs, full of cooperation and human energy to allow the process going smoothly.

In addition, there were previous studies about using mathematical equation such as, Multiple Regression Analysis (MRA) that also can be used in prediction of flood but it does not produce an accurate result. This is because of the independent variables are correlated with each other and this cause difficulties in its approach.

Currently, there was an alternative approach, which is based on Artificial Neural Network (ANN). Through this report, ANN will be the main platform of flood forecasting, in order to determine whether it is functioning well in prediction of flood occurrence.

1.4 RESEARCH OBJECTIVE

The main objectives of this research to minimize the effect of flood problems, while the sub-objectives of this research are:

1. To develop high accuracy model utilizing ANN in predicting flood.

2. To forecast flood occasion in Sungai Muda, Kuala Muda, Kedah.

3. To investigate whether water level data alone can be used to produce modelling.

4. To determine whether ANN is functioning in the flood forecasting.

1.5 SCOPE OF STUDY

1. The case study will be conducted in the area of Sungai Muda, Kuala Muda, Kedah Darul Aman and data will be gathered from a related department and reliable sources.

2. Water level data will be predicting in hourly.

3. Input data for 14 years will be stimulated by utilizing ANN.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

Rainfall is a variable component combination of atmosphere and land surface processes. The complex transient heterogeneity of storm precipitation coupled with hard mountainous physiographic setting makes the improvement of precise estimating repository inflow a few hours early an extraordinary test. The extent that precipitation estimating is concerned, the precision of precipitation items and their nowcasting is persistently enhanced and gets to be more solid for handy applications in recent years (F.J. Chang et al., 2013). Water resources as precipitation of rainfall is a profitable item and needed for the essential of human life. Substantial scale varieties of the precipitation may lead to floods and droughts, two extremely hydrological nature phenomenons regarding their effects on the environment and impacts on the ecosystem (Y.M. Chiang et al, 2006).

Flash flood is known as the effect of an extreme downpour occasion delivering a few many droplets of rain (mm), which is it can produced a large capacity of water in few hours. Flash flood normally happens in brisk reaction watersheds for two principle reasons:

(i) A short fixation time because of the size by and large under few hundred km^2,

(ii) Flood streams that are basically made out of surface overflow water or at any rate quick reacting spill over forms.

That makes extremely troublesome for crisis administrations to be involved and reveal the flash flood warnings progressively in real time (P.A. Versini, 2012).

In Peninsular Malaysia, as is attributes for a moist normal atmosphere, the climate is warm and damp during the time with temperatures of 38°C is exceptionally uncommon. There are yearly vacillations of the mean temperature of around 1.5-2°C. The most reduced normal month to month temperature happens from April-May and July-August in many places (J.M. Rajab et al., 2013). There is an exact month which is it coincides with the monsoon. The atmosphere of peninsular Malaysia definitively impacts by the storms and encounters two stormy seasons as the year progressed, connected with the South West Monsoon (SWM) from May to August and the North East Monsoon (NEM) from November to February (J.M. Rajab et al., 2013). Natural disasters are the main cause of irrecoverable damages that happened in our world and could be happen in any places and anytime without our expectations (M.S. Tehrany et al., 2014). Malaysia continually faced flood events annually mostly during the bad rainstorm seasons (M.S. Tehrany et al., 2014). These flooding have brought on impressive harm to thruways, settlement, agribusiness and occupation (M.S. Tehrany et al., 2014)

Figure 2.1: Map of Peninsular Malaysia.

The lightning and variable wind rainstorm create, amid the between storm months (for the most part in the middle of April and October), bringing on significant precipitation in each of the two move periods, particularly in the west drift states. The principle considers that influence the precipitation dispersions are the rainstorm changes and impacts of geography (J.M. Rajab et al., 2013). Other than the commitment of the numerous territorial toxin sources, these storms have distinctive impacts on the air parameters, as far as the consequences for atmosphere or the measures of contaminations, they bring to Malaysia.

Malaysia, in the same way as most tropical nations endures it impact of floods. Despite the fact that the storms were once to be rebuked for the yearly event of surge waters, today, Malaysia bears the worst thing about improvement and its uncharted outcomes when unreasonable water incapacitates groups and reason boundless decimation. The primary instrument giving rules to flood hazard management in Malaysia is the National Security Council Directive No.20 and Fixed Operating Regulations (FOR). It is just an arrangement that is regulated from government to enact flood management measures as it outlining the objectives of Policy and Mechanism on Disaster and Relief Management on Land. This management also describes the objectives and responsibilities of various agencies involved in disaster management. Flood hazard management requires more than simply reacting to the crisis that has emerged. Much the same as other flood management; it obliges a cycle that guarantees measures of counteractive action, readiness and alleviation in accordance with the Hyogo Framework for Action 2005-2015 which sets out the International Strategy for Disaster Reduction Framework. A more viable overseeing instrument could likely speed up and support in determining issues emerging on incorporation of strategies and instruments in flood management. In the journey to defend the presence of a functional lawful instrument in a solitary flood danger based legitimate skeleton in Malaysia, the English encounter that has prompted the institution of the Flood and Water Management Act 2010 will be analyzed (Zaharah Elias et al., 2013).

As the heavy storm brings floods, it will also bring big impacts to the community and ecosystem of Earth. A fast estimation of flood harm may help the first ones in allotting assets for recuperation and reproduction after a flooding occasion or in arranging sufficient flood control measures in long haul and in completing solid expense advantage investigation of these measures (V. Notaro et al., 2014). In the meantime, the learning about the normal outcomes of a flooding may encourage the conception of a flood strong society, which is the readiness of included individuals about flood dangers and harms and how to act in the occasion of a flood (V. Notaro et al. 2014). These days, the worldwide writing incorporates a few strategies for flood harm estimation in urban zones which regularly contrast about systems received, points sought after and accessibility of source information needed. A harsh order is possible between ex-post or ex-stake examination. In the first case, a harm examination at nearby scale is acquired by bookkeeping in detail the article particular harms after a flooding occasion. This sort of investigation is typically centred to assign assets for recuperation and remaking after the catastrophe occasion. Results are particular of the examined territory and generally influenced by a few lapses because of a repetitive overestimation of promptly uncovered harms, (for example, to family unit outfitting) and to an underestimation of the flood impacts on structures in long haul, (for example, the deterioration of advantages). In the recent case, ex-bet investigation gives the normal harm to a potential flooding occasion in the examined zone (V. Notaro et al. 2014).

Besides, it could influence the health of community. Flooding can have a real effect on the strength of a populace, for example, wounds, faecal–oral ailment, vector-borne infection, rodent borne infection, intense asthma, skin rashes, outbreaks of gastroenteritis also respiratory disease and mental wellbeing issues. A decent understanding of the effects of surges on populace wellbeing is important to breaking point potential wellbeing effects. Flood water can enter sub-carpets, divider cavities, and can be consumed into building materials, transporting both human and crafty pathogens into the building envelope. The wet condition could prompt to the development of germs from the flood water, and the existing living beings in structures, with the level of development and perseverance reliant on the volume and compound and organic substance of the flood water, the properties of the tainting organisms, and the encompassing natural conditions. The ecological conditions that effect microbial survival incorporate the hotness and dampness transport properties of the envelope outline, the rebuilding strategies, which will focus the pace of drying also level to which the building dries, and the accessible supplements or development inhibitors in the development materials and tainting water and silt (J. Taylor et al., 2011) These harms could lead danger in human life.

Sulafa Hag Elsafi, 2014, has stated in his journal that there are two ways approaches involved in prediction of flood. The first method is based on mathematical modelling. It is a method in models physical dynamics between principal interacting components of the hydrological system. Generally, a rainfall-runoff model is used in transforming the data values of rainfall, evaporation and flow data into a hydrograph prediction by considering the spatial variation in storage capacity. Then, a hydraulic channel flow routing model will be used in calculating the flow. The second method is based on modelling the relationship between hydrologic input and output data, without terminate the relationships that exist among the involved physical process. Therefore, stimulating data of flood forecasting by using modelling is one of main goals of this report been prepared.

CHAPTER 3

METHODOLOGY

3.1 STUDY AREA

The study is going to be conducted in Sungai Muda, Kuala Muda, Kedah Darul Aman which is situated in the North western part of Peninsular Malaysia. The study area is the longest river in Kedah, located in Kuala Muda which now having their centre administrative in Sungai Petani. The river provides water supply to Kedah as well as Penang. The study is going to cover the area of Sungai Muda which is located at latitudes 5°12’57 N and longitudes 100°2716 E, (Figure 3.1). Sungai Muda is having basin area of 4,302 km^2 with length about 203km is the main river for district of Sungai Petani. District of Sungai Petani is located at the southest of northern Peninsular Malaysia and also known as the largest town in the state of Kedah. Kuala Muda has a population of about 456, 605 people (Department of Statistics Malaysia, 2010)

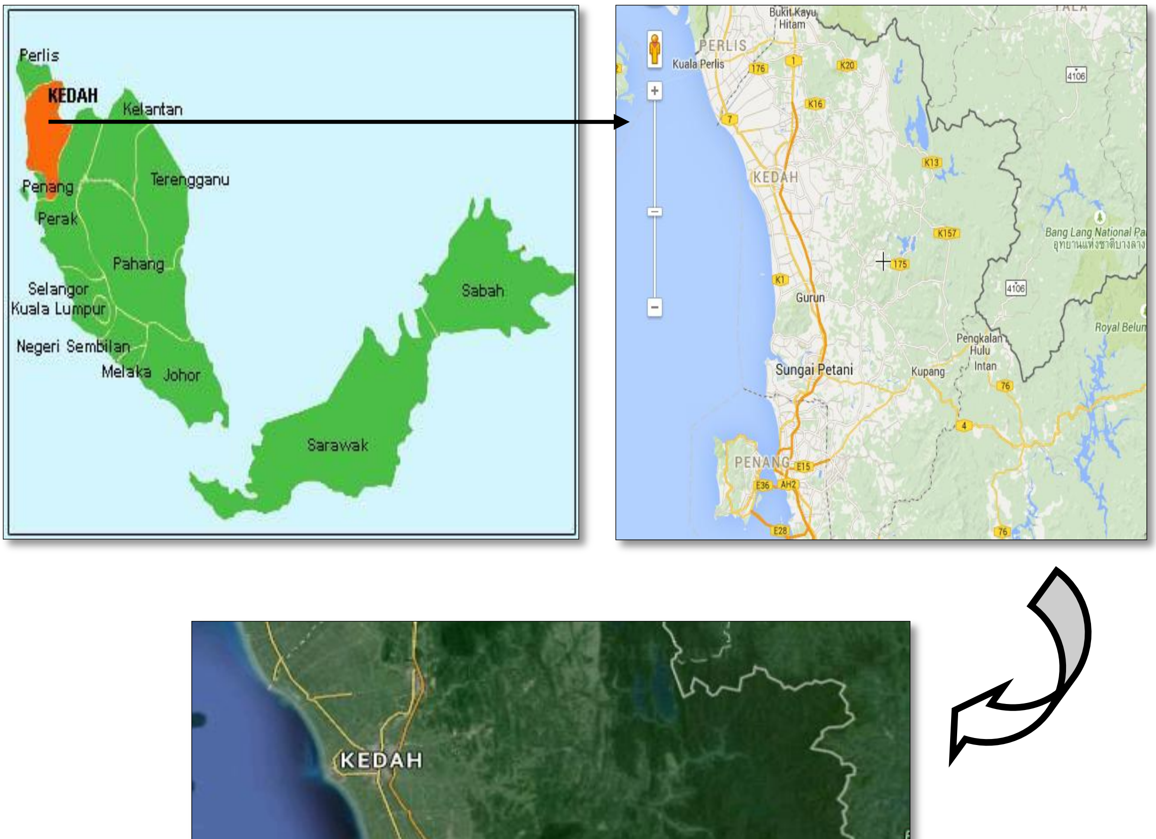


Figure 3.1: Study area in Kuala Muda, Kedah.

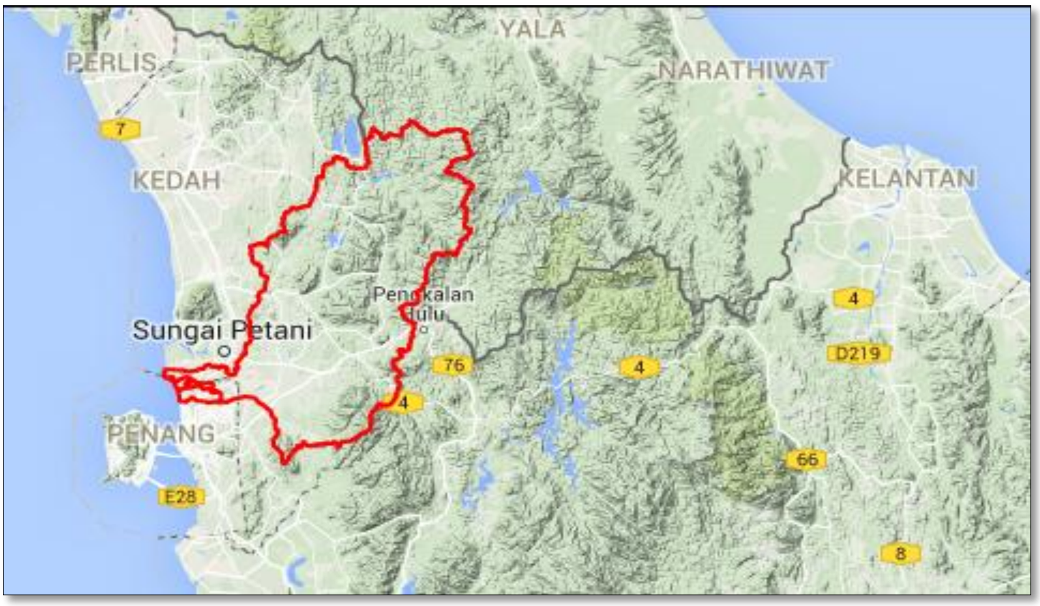


Figure 3.2: Catchment basin area of Sungai Muda in Kuala Muda, Kedah.

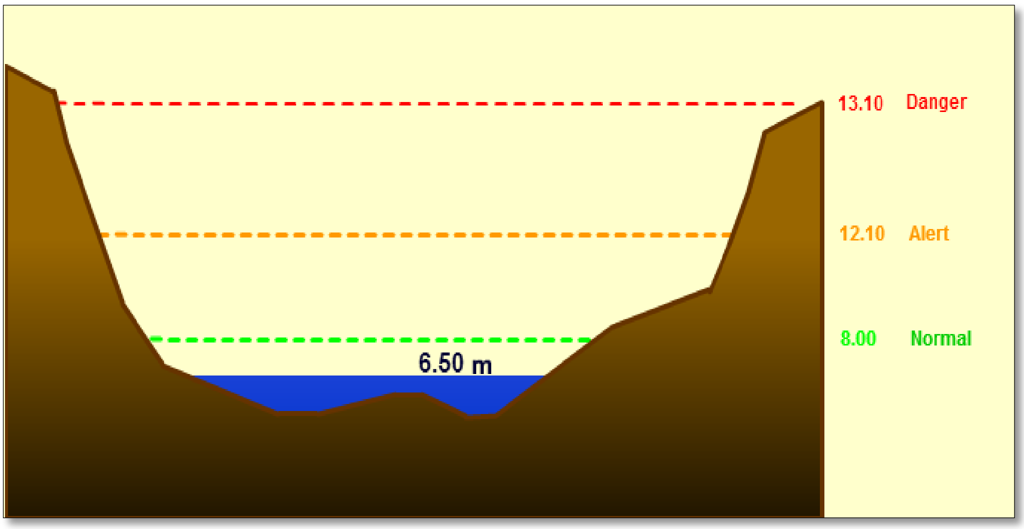


Figure 3.3: Cross-section of Sungai Muda at Jambatan Syed Omar.

3.2 METHOD OF ANALYSIS

3.2.1 Introduction of Artificial Neural Network (ANN)

In order to make a prediction or assume the future flood occurrence in certain of area, analyzing the pattern of past record of its occasion is necessary. Therefore, reading of rainfall data is considered as the most important factor for prediction of future disaster occurrence and it can represent single or multiple events in a specific area. There is an alternative, which is based on Artificial Neural Network (ANN) had approach in flow flood forecasting and it has been developed in the recent years. In order utilizing an ANN to solve a problem, the first step is to train the ANN to “study” the relationship between the input data and the output data.

ANN, which are modelled on the operating behaviour of the brain, are tolerant of some imprecision and are especially useful for classification and function approximation/mapping problems, to which hard and fast rules cannot be applied easily (David, 2000, John, 2000). The mind fundamentally gains as a matter of fact. It is common confirmation that a few issues that are past the extent of current machines are in fact resolvable by little vitality proficient bundles. This mind displaying likewise guarantees a less specialized approach to create machine arrangements. This new approach to figuring additionally gives a more effortless debasement amid framework over-burden than its more conventional partners.

These systems are "neural' as in they may have been propelled by neuroscience yet not so much in light of the fact that they are reliable models of organic neural or cognitive phenomena. Actually dominant part of the system is all the more nearly identified with conventional numerical and/or measurable models, for example, non-parametric example classifiers, grouping calculations, nonlinear channels, and factual relapse models than they are to neurobiology models. Neural networks have been utilized for a wide mixture of use where factual routines are customarily utilized. They have been utilized as a part of use issue, for example, recognizing submerged sonar flaws, perceiving discourse, and anticipating the optional structure of globular proteins. In time-arrangement applications, neural network have been utilized as a part of anticipating stock exchange execution.

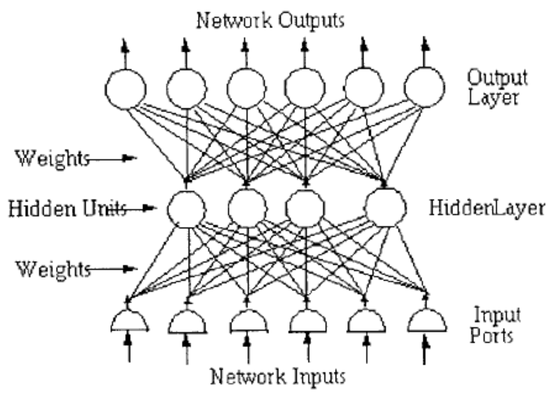


Figure 3.4: Diagram of Neural Network

3.2.2 Basics of Artificial Neural Network

The terminology of artificial neural networks has created form an organic biological model of neural system, which it comprises an asset of joined cells, the neurons. The neurons receive impulses or response from either input cells or any other neurons. It will perform some kind of transformation of the input and then, it will transfer the outcome to other neurons or also known as output cells. The neural networks are developed from many layers of connected neurons. Therefore, the one layer that receives impulse or response from the preceding layer of neurons and let it passes through the output on the next layer.

A neuron is a real function of the input vector (y_1,....., yk), while the output vector is obtained as $f(x_j) = f(\, a_j + \sum_{i=1}^{k}$ $w_{ij}\, y_j$), where f is a function usually known as the sigmoid (logistic or tangent hyperbolic) function.

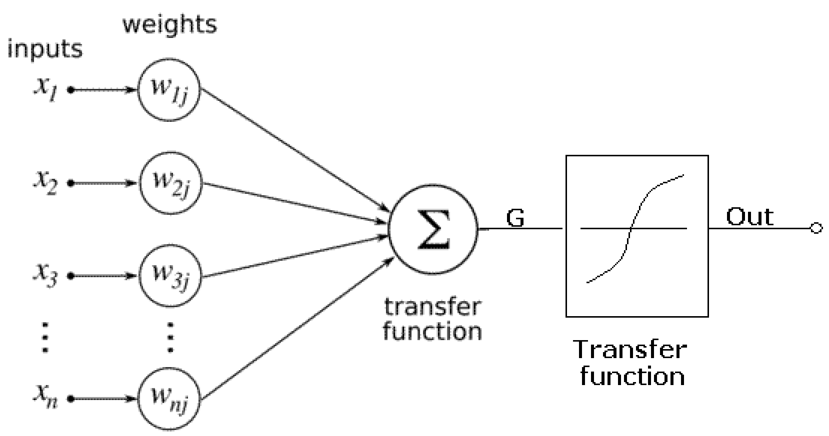


Figure 3.5: Basic of Artificial Neural Network.

3.2.3 Characteristics of Artificial Neural Networks:

- The neural networks exhibit mapping capabilities, that is, they can map input patterns to their associated output patterns.

- The neural networks learn by examples. So, neural network's architectures can be 'trained' with known examples of a problem before they are tested for their 'inference' capability on unknown instances of the problem. Therefore, they can identify new objects which previously untrained.

- The neural networks take over the capability to conclude. Thus, they can predict new results from previous results.

- The neural networks are robust systems and are fault tolerant. They can, therefore, repeat full patterns from incomplete, partial or complete patterns.

- The neural networks can process information in parallel, at high speed and in a distributed manner.

3.3 NEURAL NETWORKS ARCHITECTURES

3.3.1 Types of architecture of ANNs

3.3.1.1 Recurrent Networks

These type of networks are quite different from feed forward networks system, which it has at least one feedback loop. Therefore, in these networks there may be exist one layers with feedback connections a shown in figure 5. It also could have neurons with self feedback links.

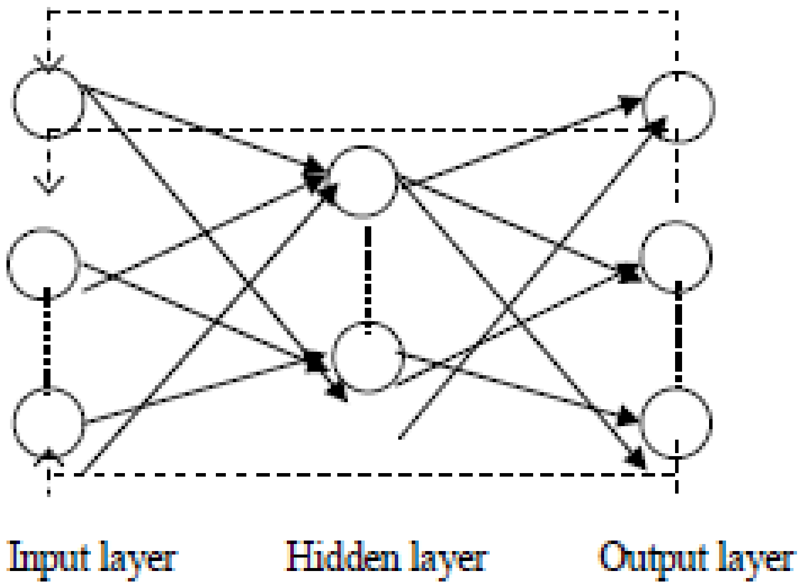


Figure 3.6: Recurrent Networks

3.3.1.2 Feed-Forward Networks

In a feed-forward network, information flows in one direction along the connecting pathways, from the input layer through the hidden layer to the final output layer. There are two types of feed-forward neural network:

a) Single-layer Perceptron is a single-layer perceptron network, which it consists of a single layer of output nodes. The inputs are fed directly to the outputs through a series of weights. Based on its behaviour of action, it can be considered as the simplest type of feed-forward network.

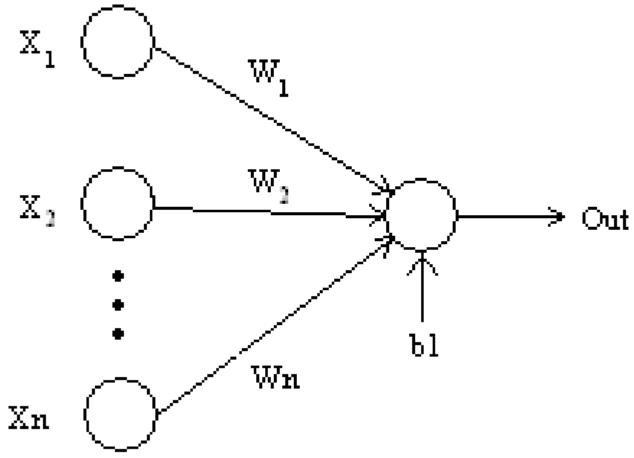


Figure 3.7: A Single-layer Feed-Forward Networks.

b) Multiple-layer Perceptron, is a multi-layers of computational units and it is usually interconnected in a feed-forward way. Each neuron in one layer has directed the connections to the other neurons at the next layer. This behaviour of unit network had applied a sigmoid function known as activation function.

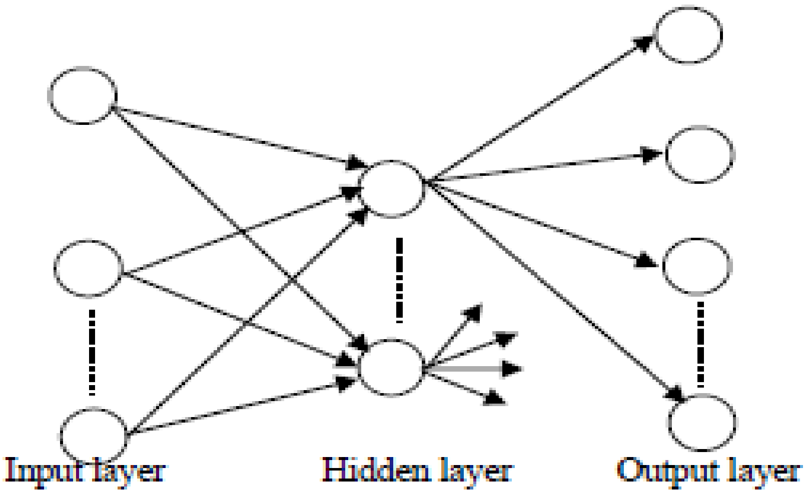


Figure 3.8: A Multi-layer Feed-Forward Networks.

Through this report, feed-forward artificial neural network has been in main setting to predict flood occurrence in the study area. Feed-forward neural network is the simplest type of artificial neural network been built. In this network system, the information from the input nodes just moves straight only in one direction, forward through the hidden nodes, if there is any, and moves to the output nodes. This network system also does not have cycles or loops.

3.4 ADVANTAGES OF UTILIZING ANNS.

The advantages of utilizing ANNs are the following (Silverman and Dracup, 1999):

- It does not required a priori knowledge of the underlying process.
- The existing complex relationships that occurred among the various aspects of the process do not need to be recognised.
- It does not give preset on any solution conditions such as those required by standard optimization or statistical models.
- It does not make assumption on constraints and priori solution structures.

3.5 DISADVANTAGES OF UTILIZING ANNS.

Masters, 1993, had stated before that there were some of disadvantages by utilizing ANNs, such as:

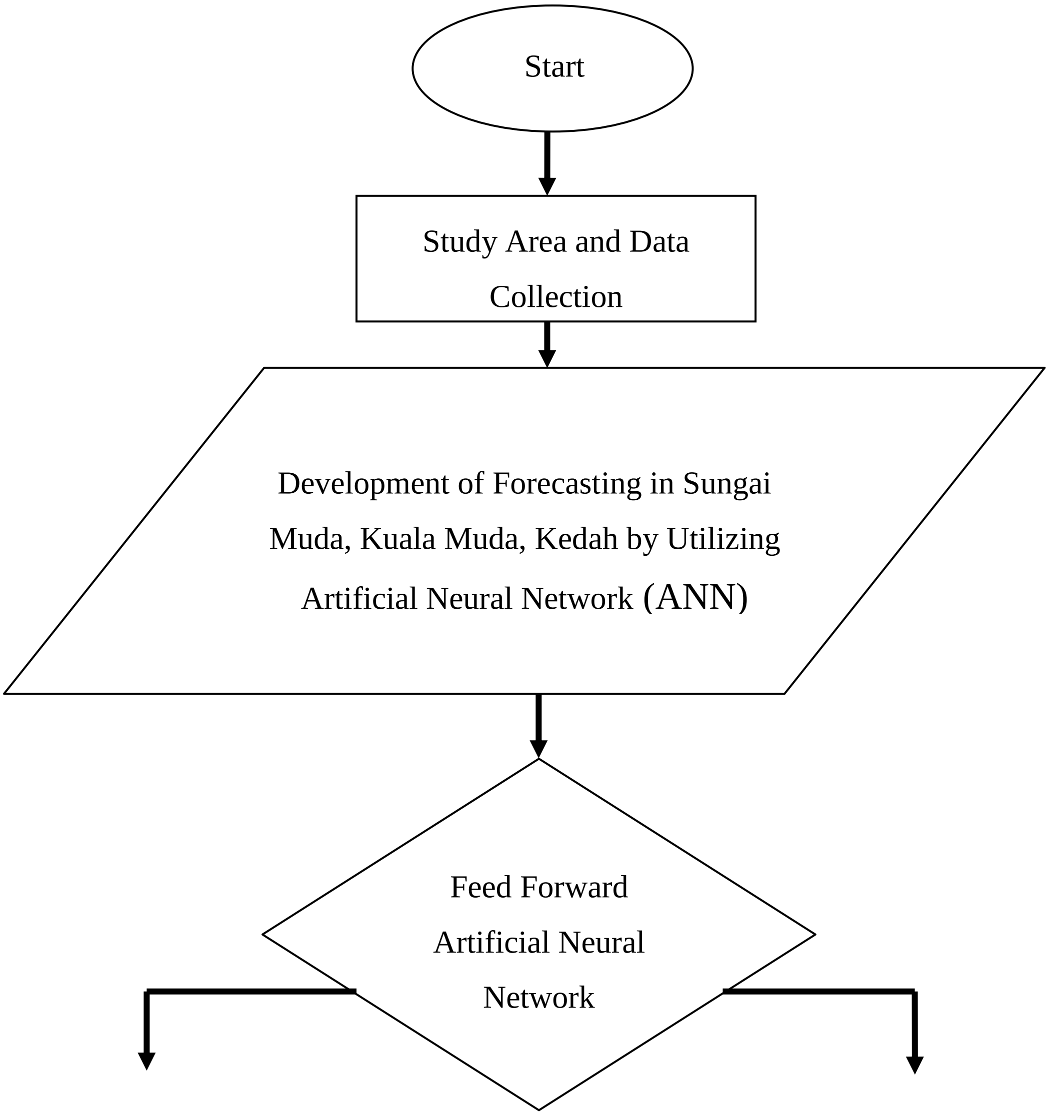
i. The exponential of training time would be increased with the increased of dataset size value.

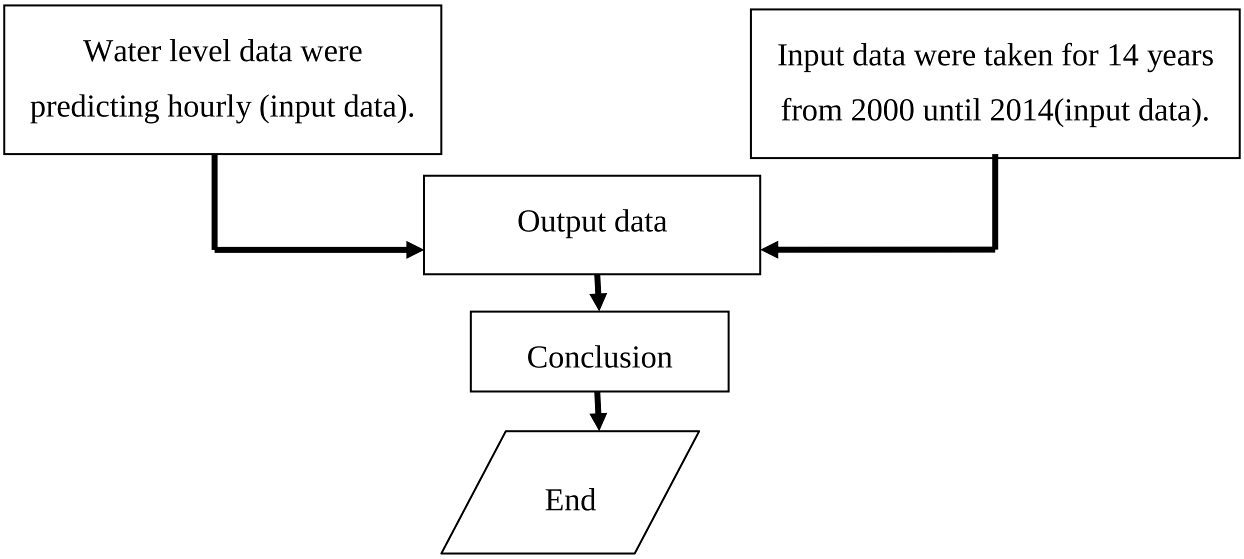
ii. Besides, ANNs do not have the certainty about the network and training set size requirements.

iii. It do not have the certainty about relationships that used by the network to produce the network's output.

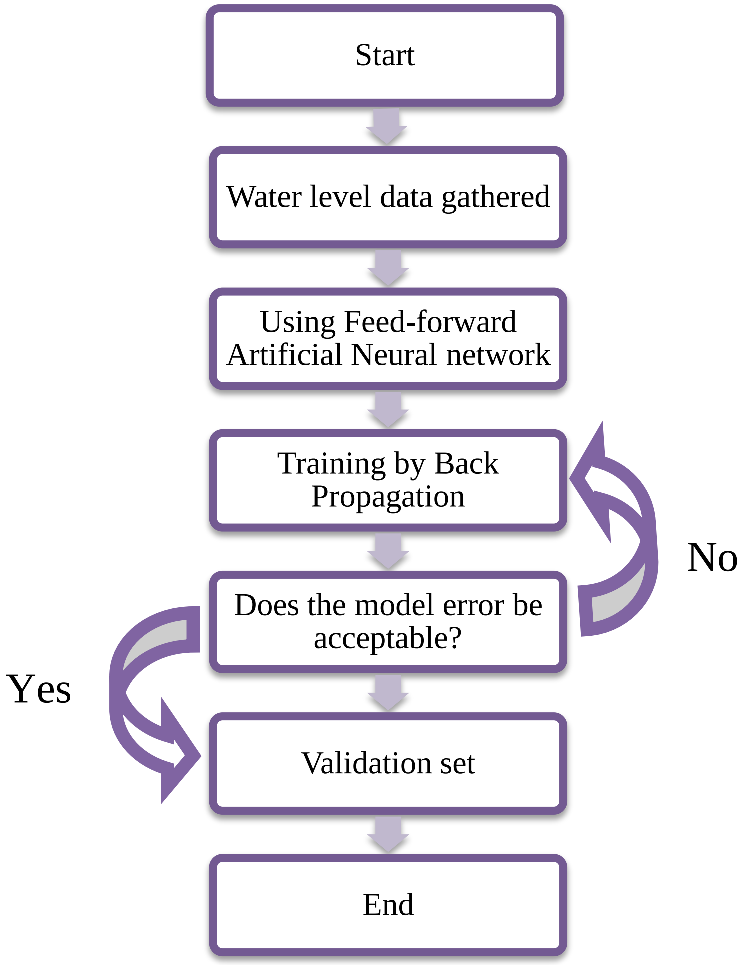
3.6 DATA COLLECTION

3.6.1 Flowchart





3.6.2 Process of Artificial Neural Network



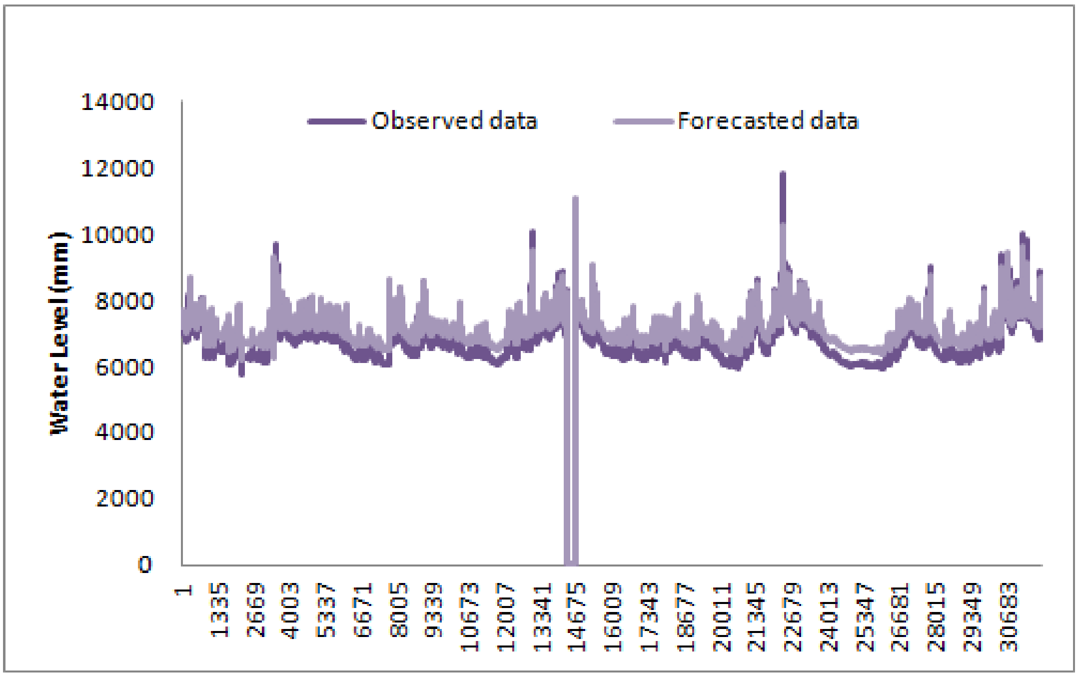
CHAPTER 4

RESULTS AND DISCUSSIONS

4.1 6 HOUR INTERVAL TIME.

4.1.1 1000 Iteration.

The data training and data validation results for the 6 hour interval time were having six inputs (2+1, 3+1, 4+1, 5+1, 6+1 and 7+1), which each input were using 2 types of iteration, 1000 and 5000 were presented herein. Furthermore, Figure 4.1 until Figure 4.6 shows data validation results for six inputs that using 1000 iteration trained by utilizing Feed-forward Back Propagation ANN and also shows the error for each input would have. Table 4.1 shows that the validation DT of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 is 0.959, 0.962, 0.972, 0.975, 0.974 and 0.978 respectively. Meanwhile, validation DV of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 are 0.821, 0.841, 0.881, 0.889, 0.888 and 0.905 respectively. Moreover, the NSC value for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 0.821, 0.841, 0.881, 0.889, 0.888 and 0.905 respectively. Besides, from the iteration it also had showed the RMSE for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 400.678, 377.43 327.047, 314.929, 316.431 and 292.357 respectively for the data validation The results indicate that more input used in the training, more accurate the trained pattern we could get. It indicates that for the 2+1 input had more scattered data than 7+1 input data.



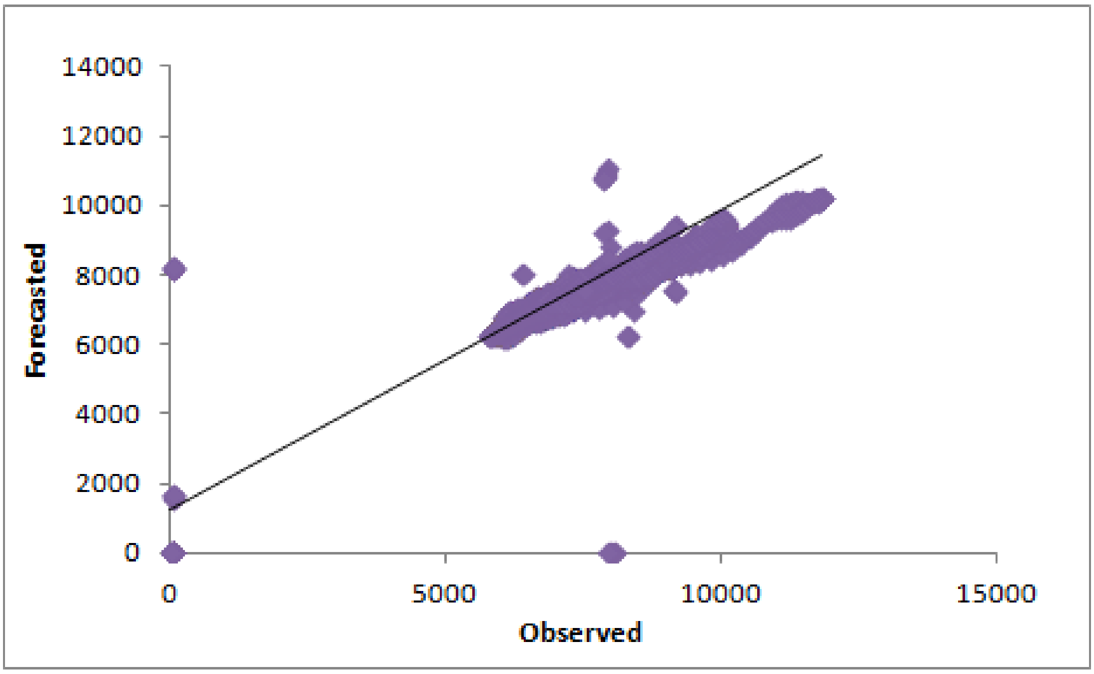
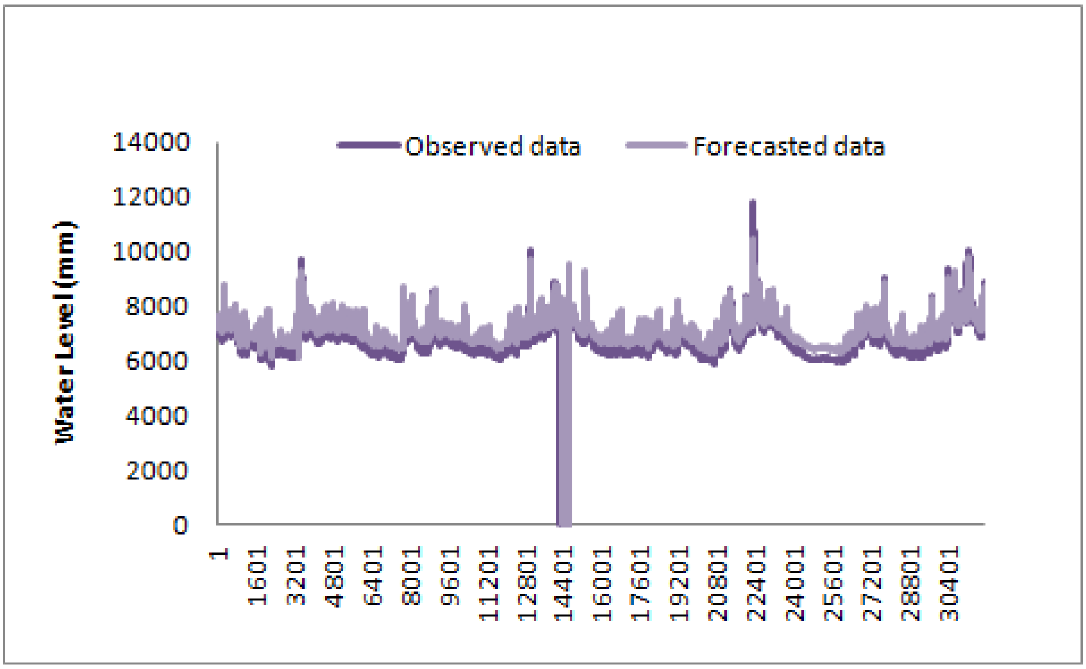


Figure 4.1: Data performance for 2 input with 1000 iteration.



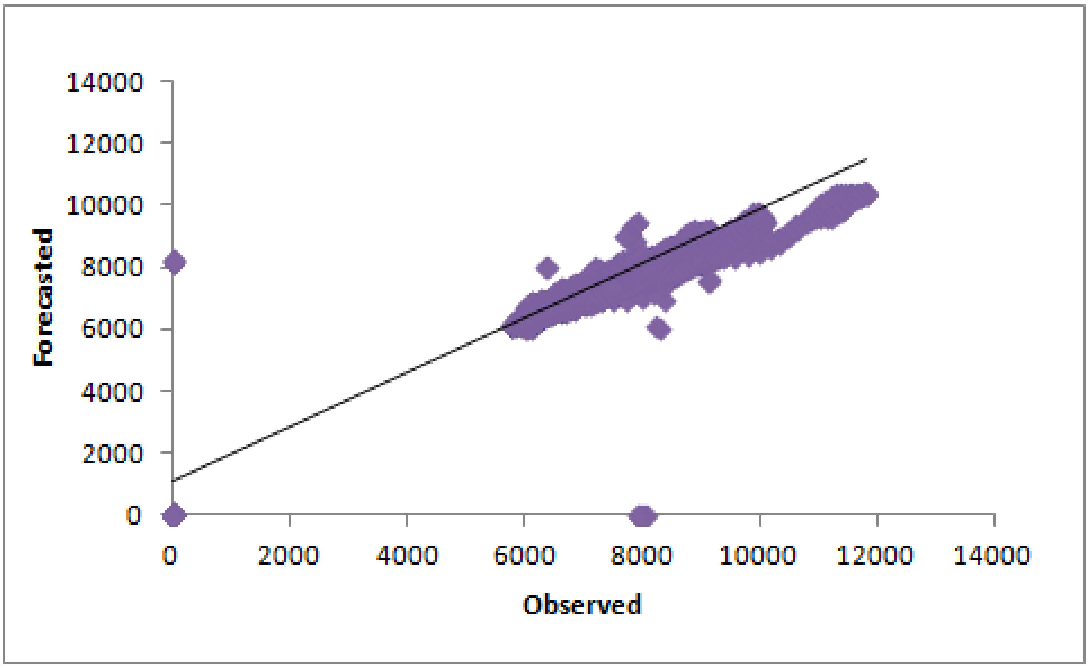
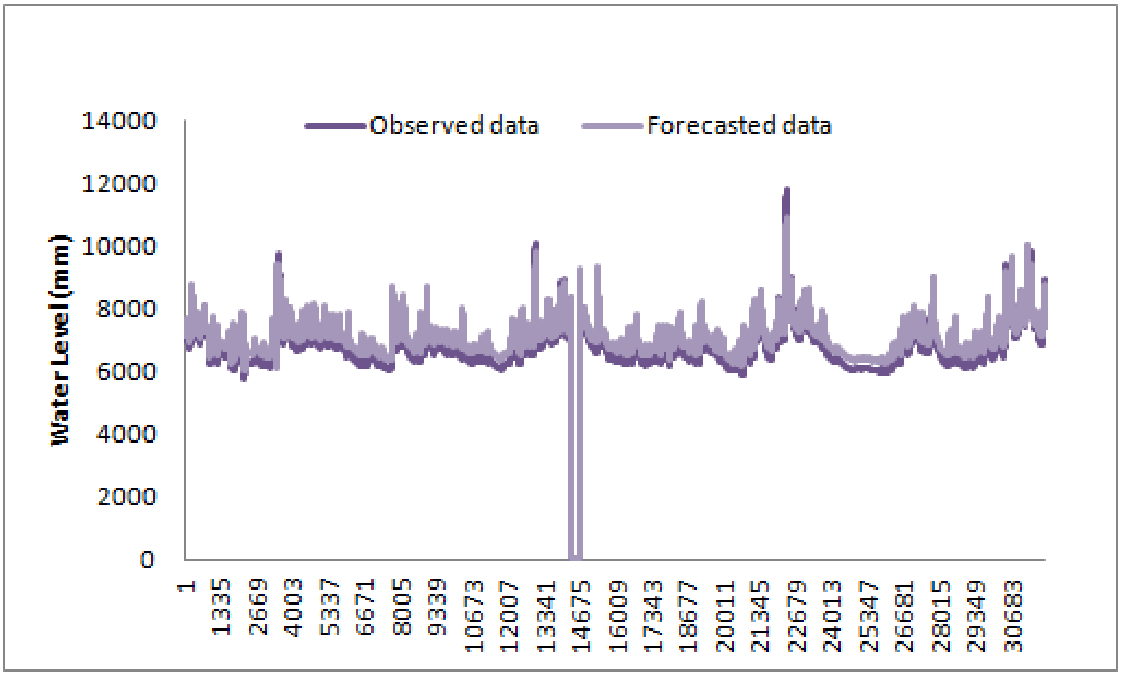


Figure 4.2: Data performance for 3 input with 1000 iteration.



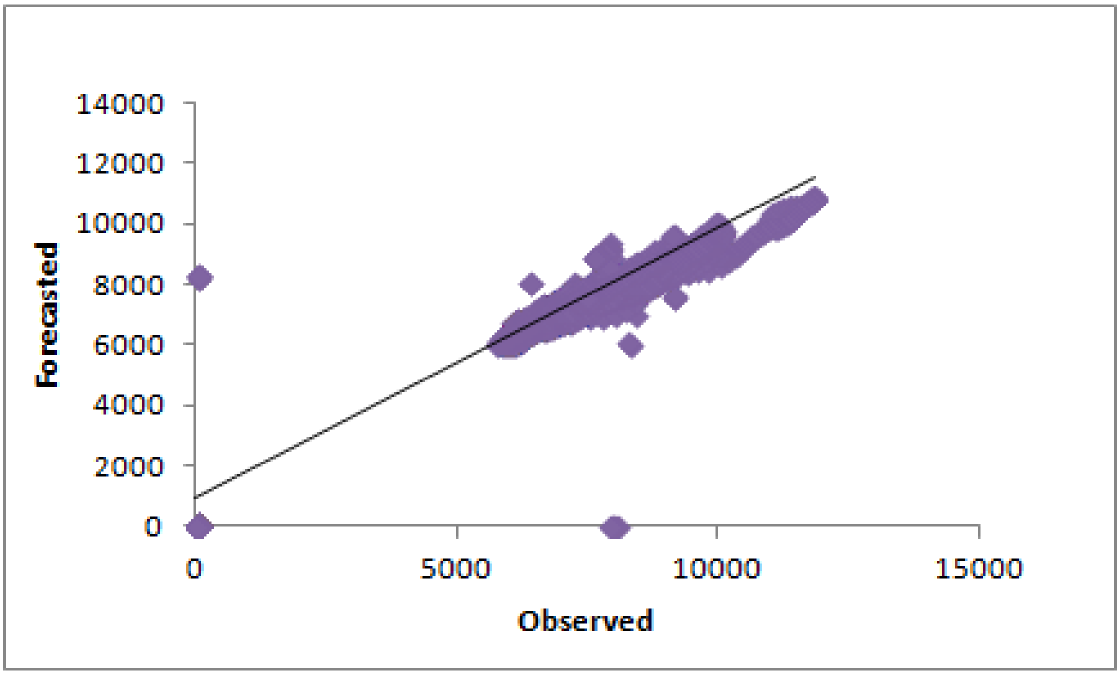
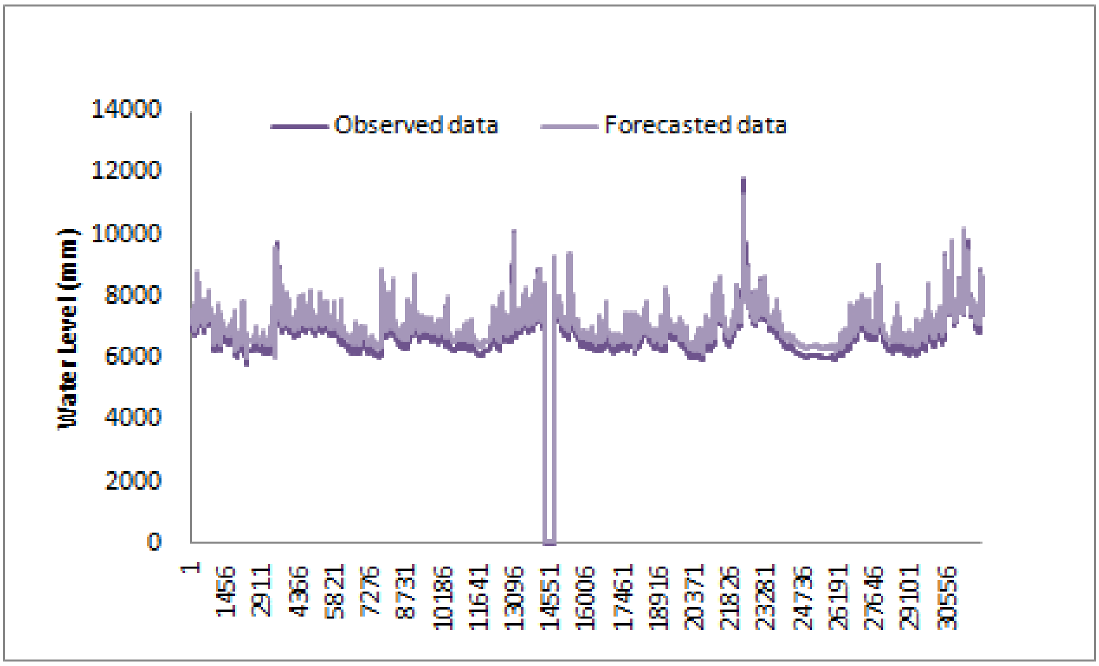


Figure 4.3: Data performance for 4 input with 1000 iteration.



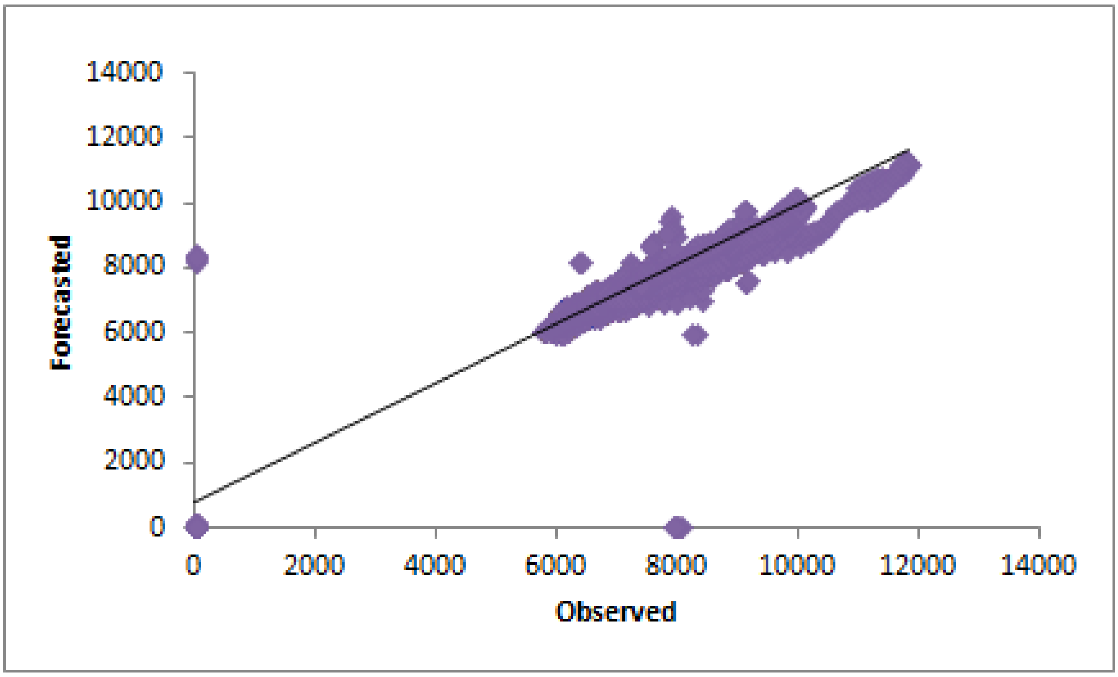
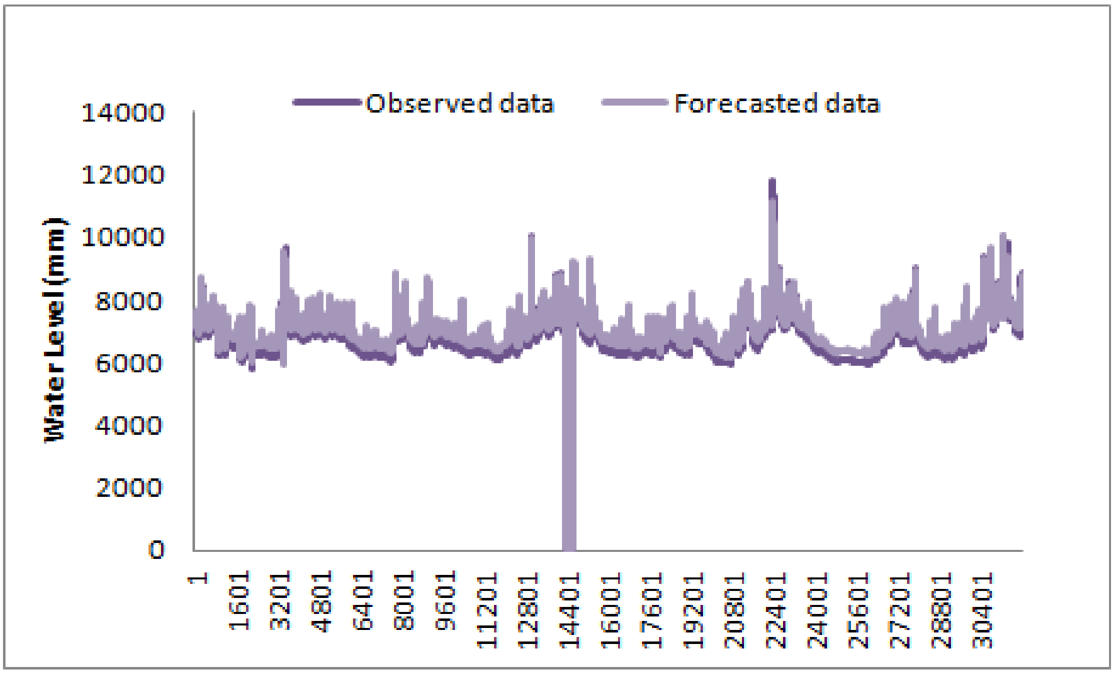


Figure 4.4: Data performance for 5 input with 1000 iteration.



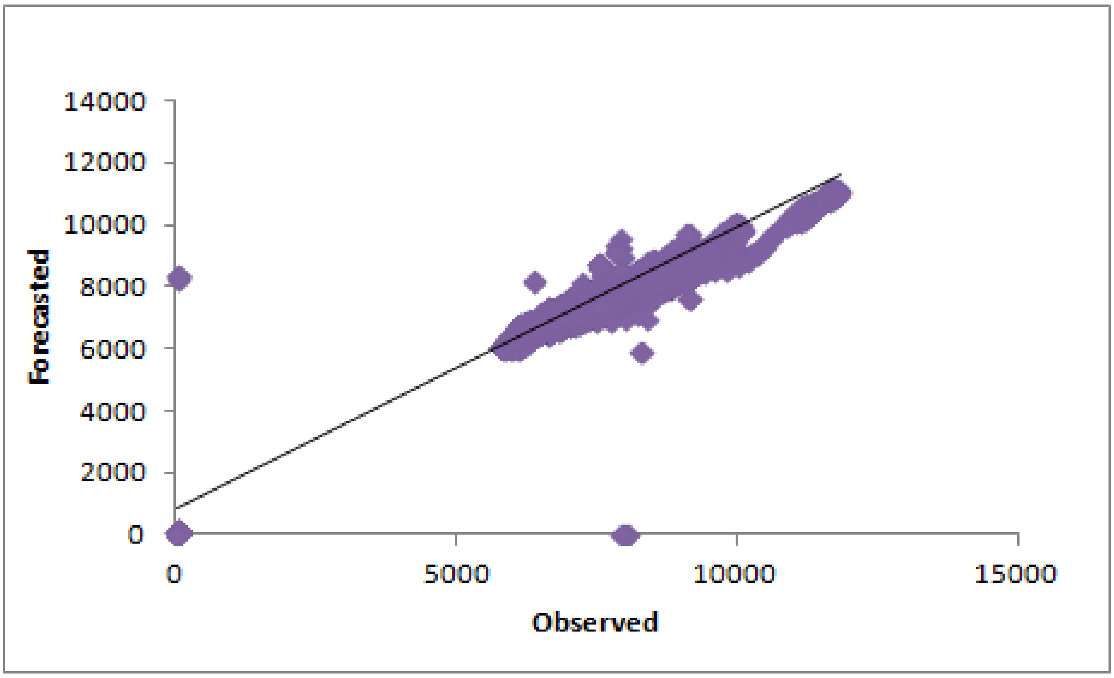


Figure 4.5: Data performance for 6 input with 1000 iteration.

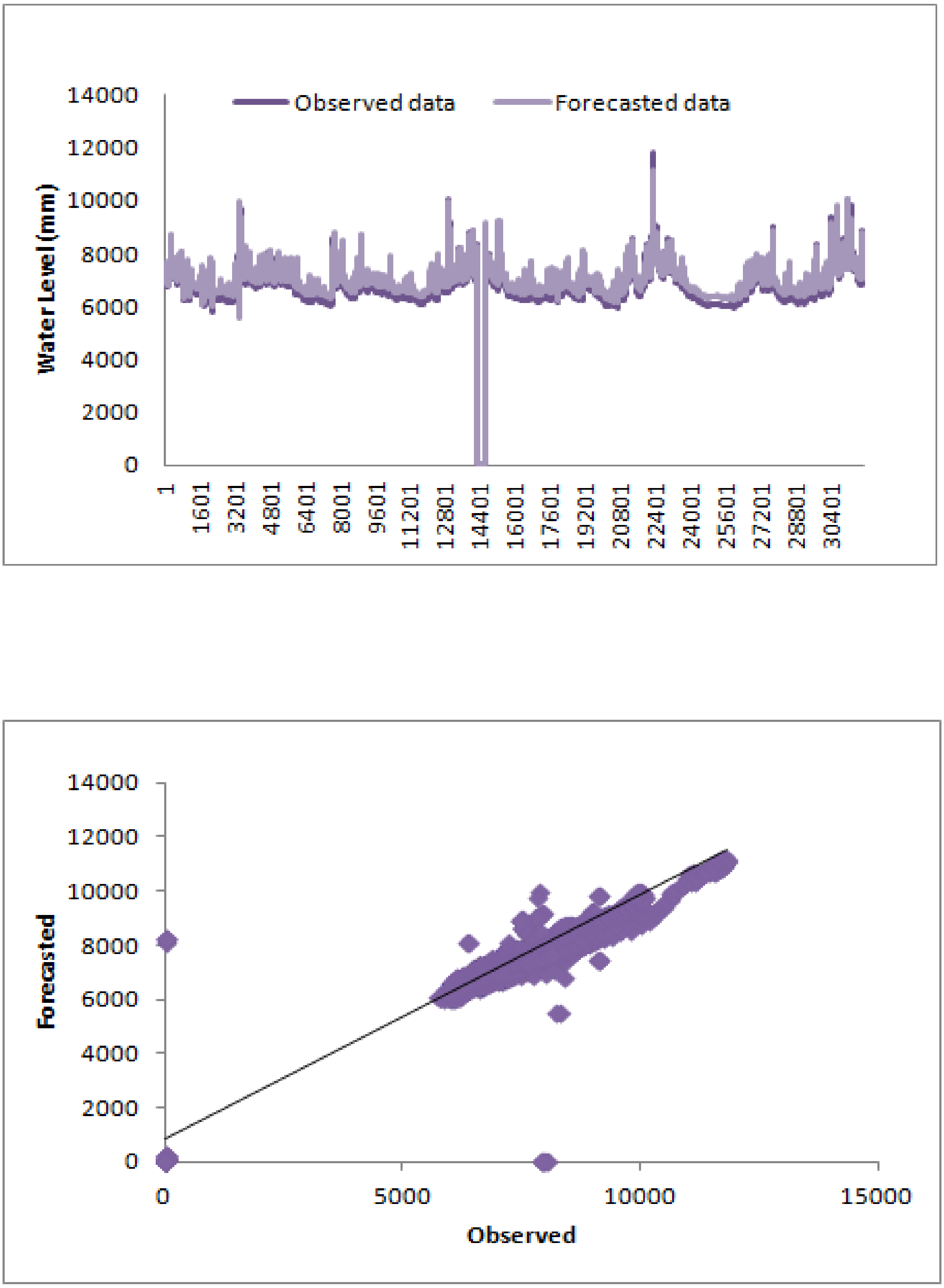


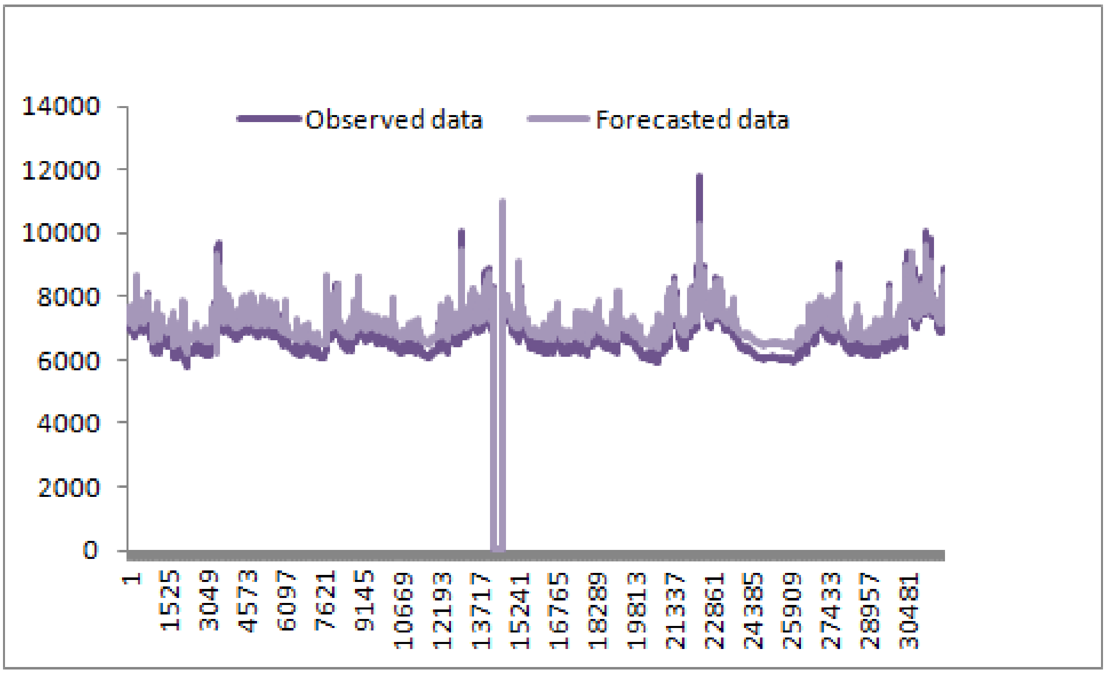
Figure 4.6: Data performance for 7 input with 1000 iteration.

Table 4.1: Data training and data validation results for all inputs trained for 1000 iteration 6 hour interaval time.

Interval 6 Hour							
Input + output	Iteration	Validation DT	Validation DV	NSC		RMSE	
				DT	DV	DT	DV
2+1	1000	0.959	0.821	0.999	0.821	60842.447	400.678
3+1	1000	0.962	0.841	0.962	0.841	375.836	377.43
4+1	1000	0.972	0.881	0.972	0.881	323.908	327.047
5+1	1000	0.975	0.889	0.975	0.889	307.280	314.929
6+1	1000	0.974	0.888	0.974	0.888	314.411	316.431
7+1	1000	0.978	0.905	0.978	0.905	290.628	292.357

4.1 2 5000 Iteration.

Besides than using 1000 iteration, 5000 iteration of 6 hour interval time also been used in this research. For each 2+1, 3+1, 4+1, 5+1, 6+1, and 7+1 input were also trained with 5000 iteration as the maximum iteration. Furthermore, Figure 4.7 until Figure 4.12 shows data validation results for six inputs that using 5000 iteration trained by utilizing Feed-forward Back Propagation ANN and also shows the error for each input would have. Table 4.2 shows that the validation DT of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 is 0.959, 0.962, 0.970, 0.972, 0.974 and 0.979 respectively. Meanwhile, validation DV of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 are 0.824, 0.837, 0.876, 0.885, 0.893 and 0.907 respectively. Moreover, the NSC value for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 0.824, 0.837, 0.876, 0.885, 0.893 and 0.907 respectively. Besides, from the iteration it also had showed the RMSE for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 397.472, 382.107, 333.365, 320.422, 309.566 and 288.332 respectively for the data validation. The results indicate that more input used in the training, more accurate the trained pattern we could get. It indicates that for the 2+1 input had more scattered data than 7+1 input data.



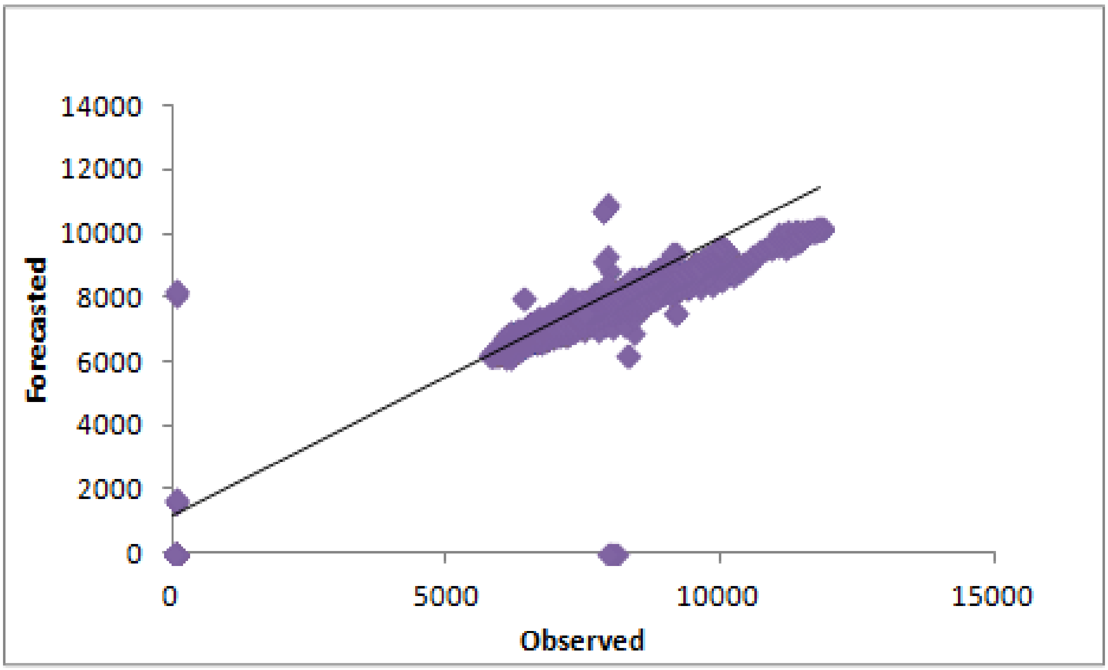
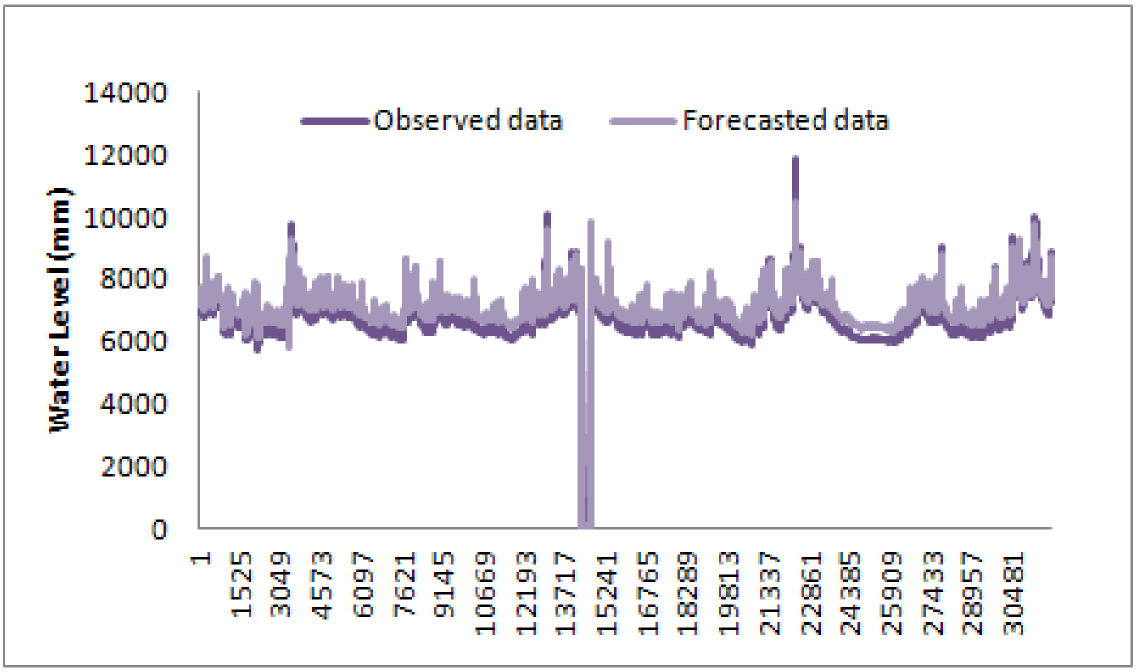


Figure 4.7: Data performance for 2 input with 5000 iteration.



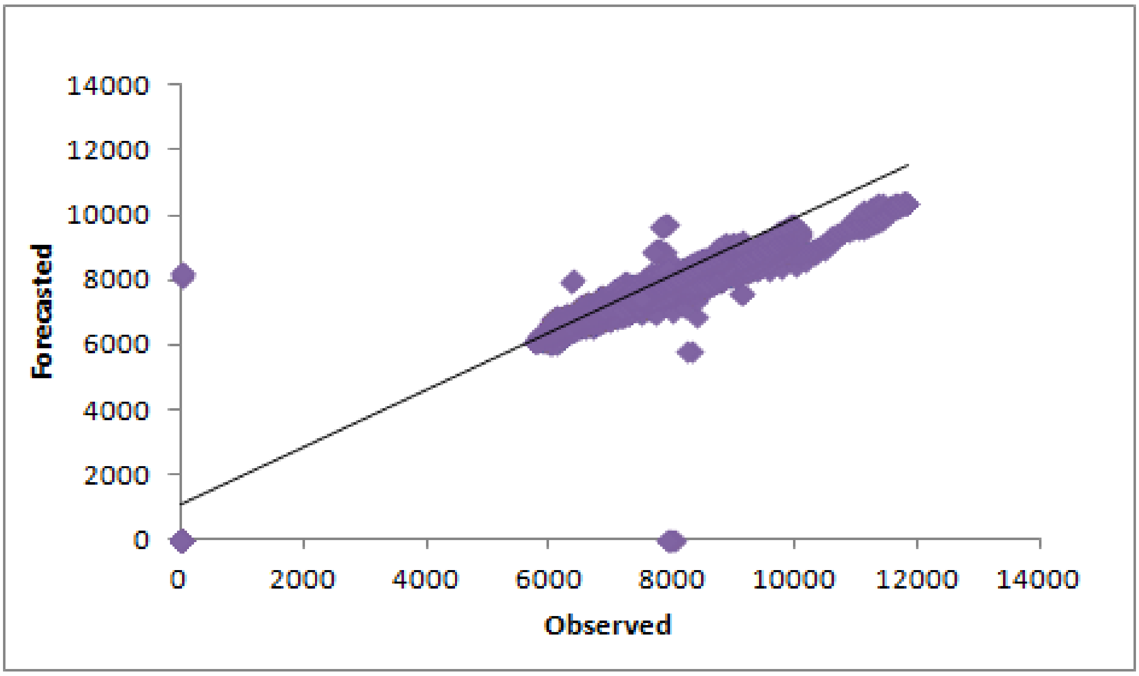
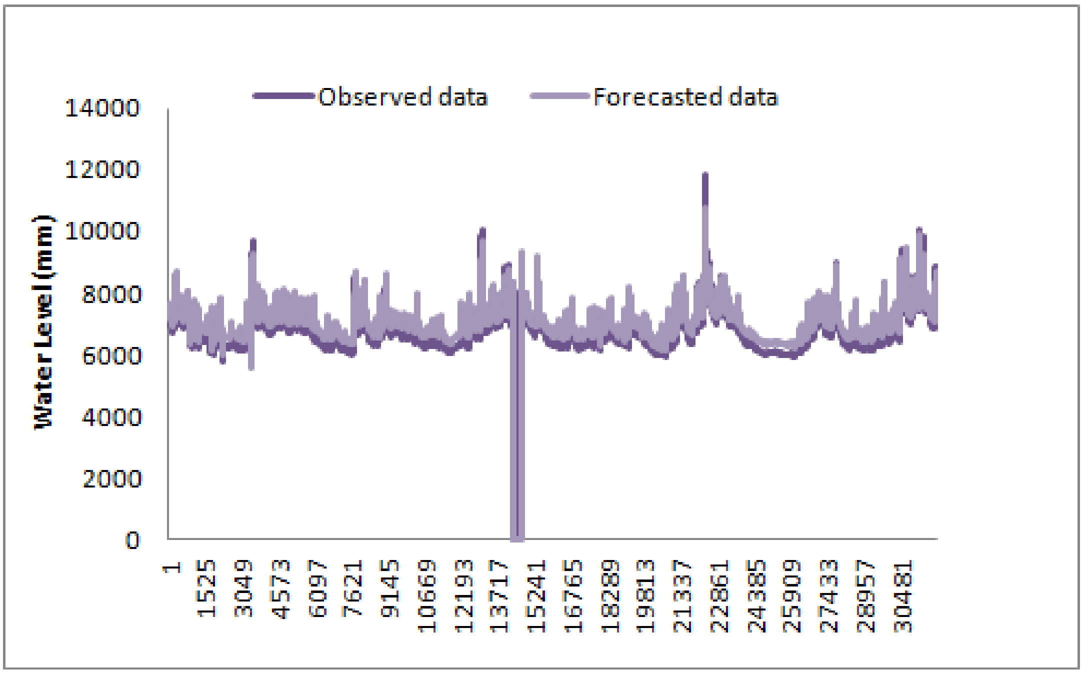


Figure 4.8: Data performance for 3 input with 5000 iteration.



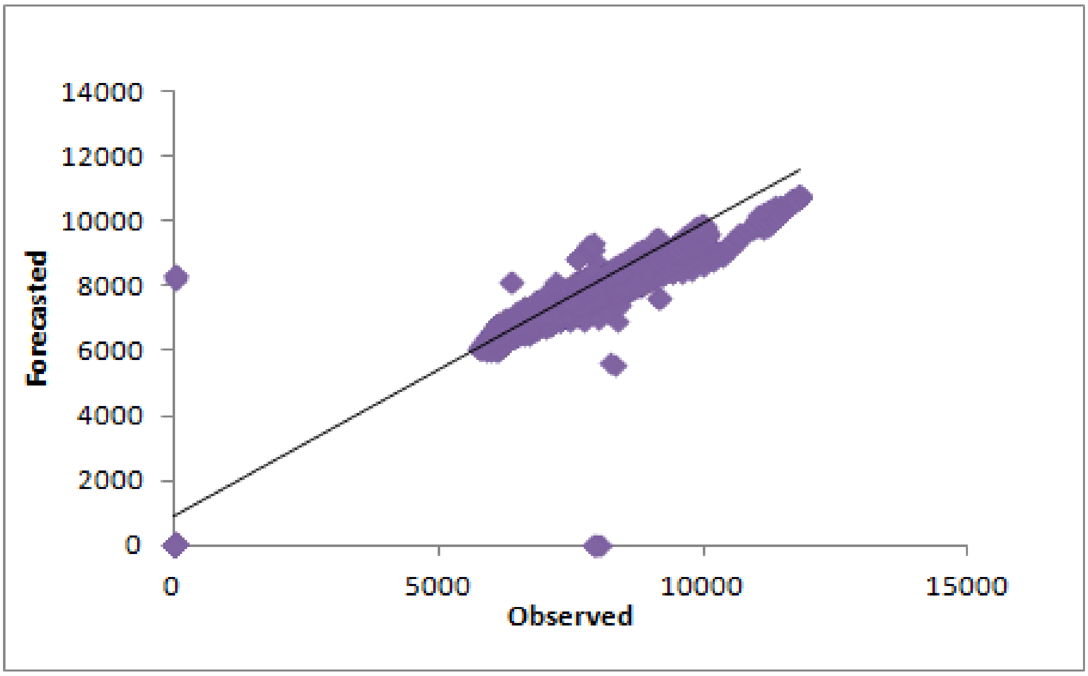
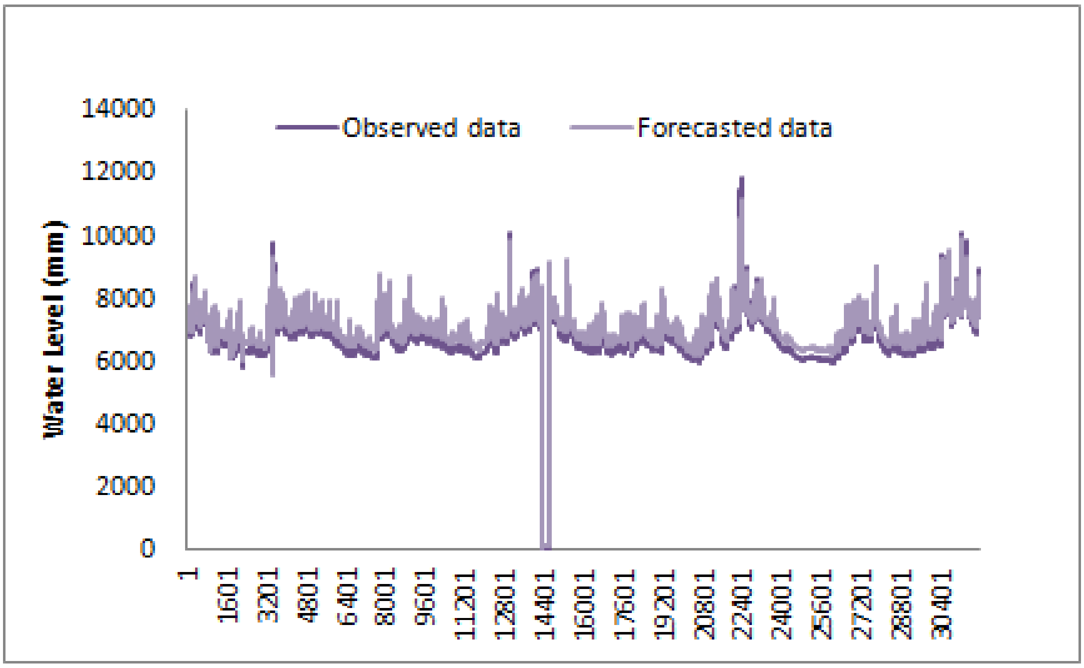


Figure 4.9: Data performance for 4 input with 5000 iteration.



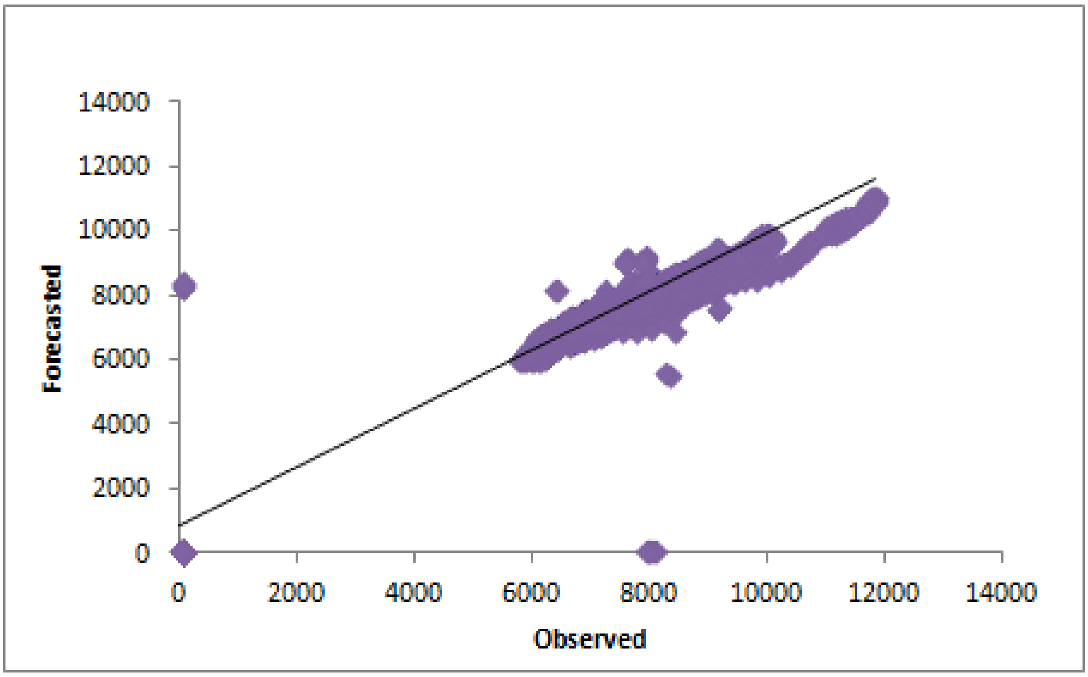
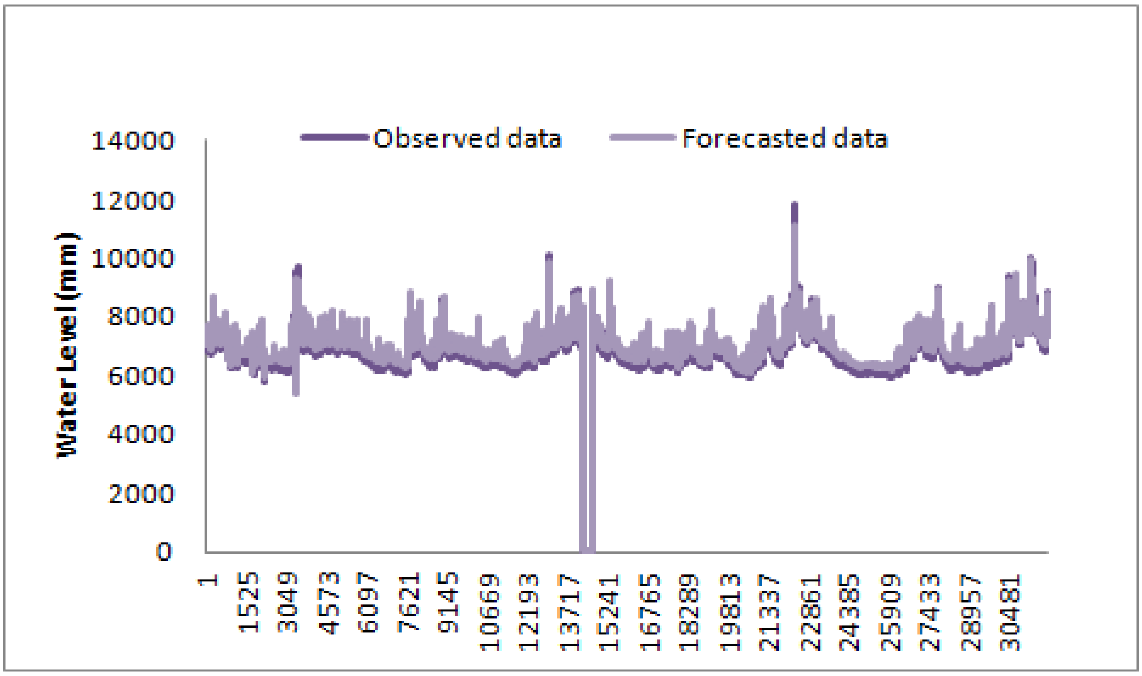


Figure 4.10: Data performance for 5 input with 5000 iteration.



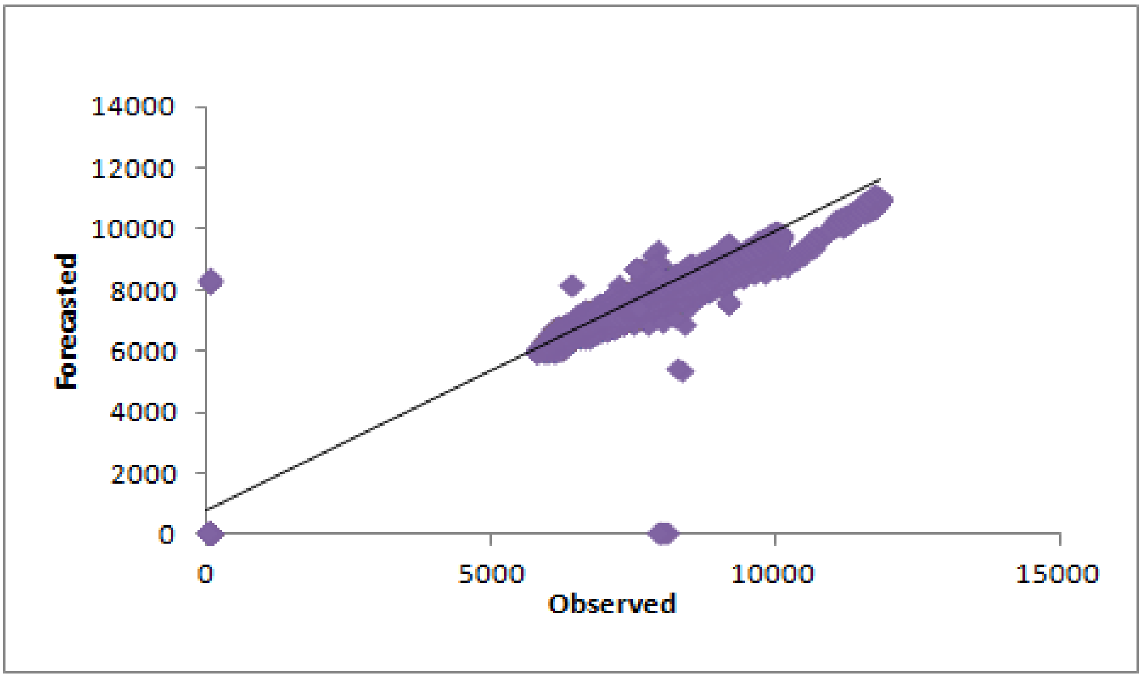
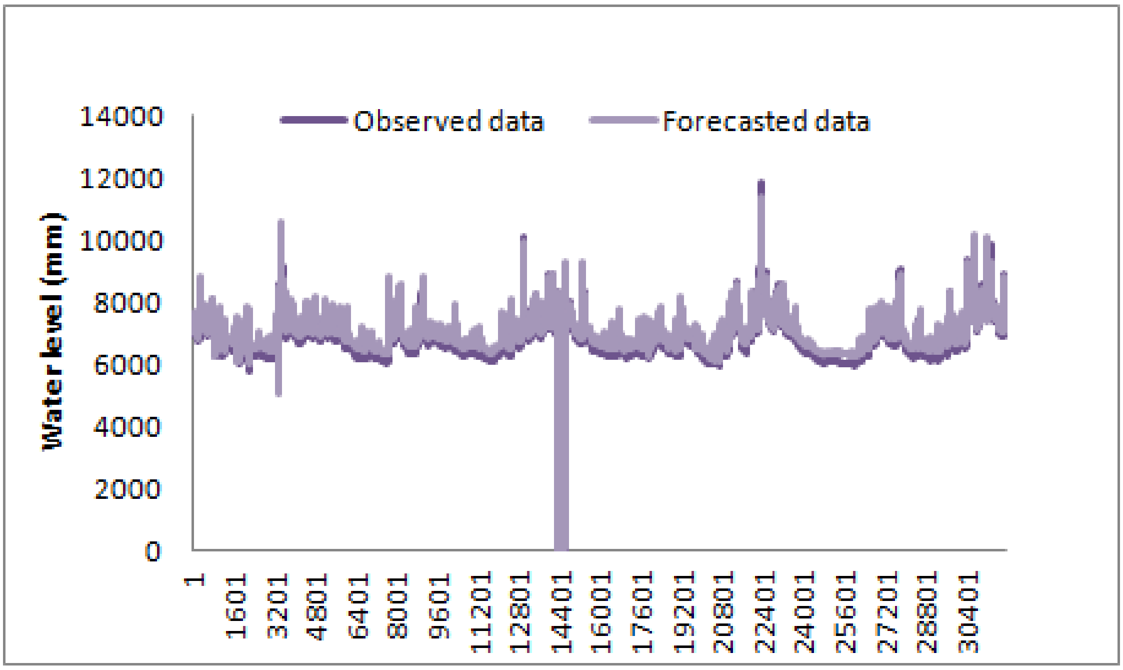


Figure 4.11: Data performance for 6 input with 5000 iteration.



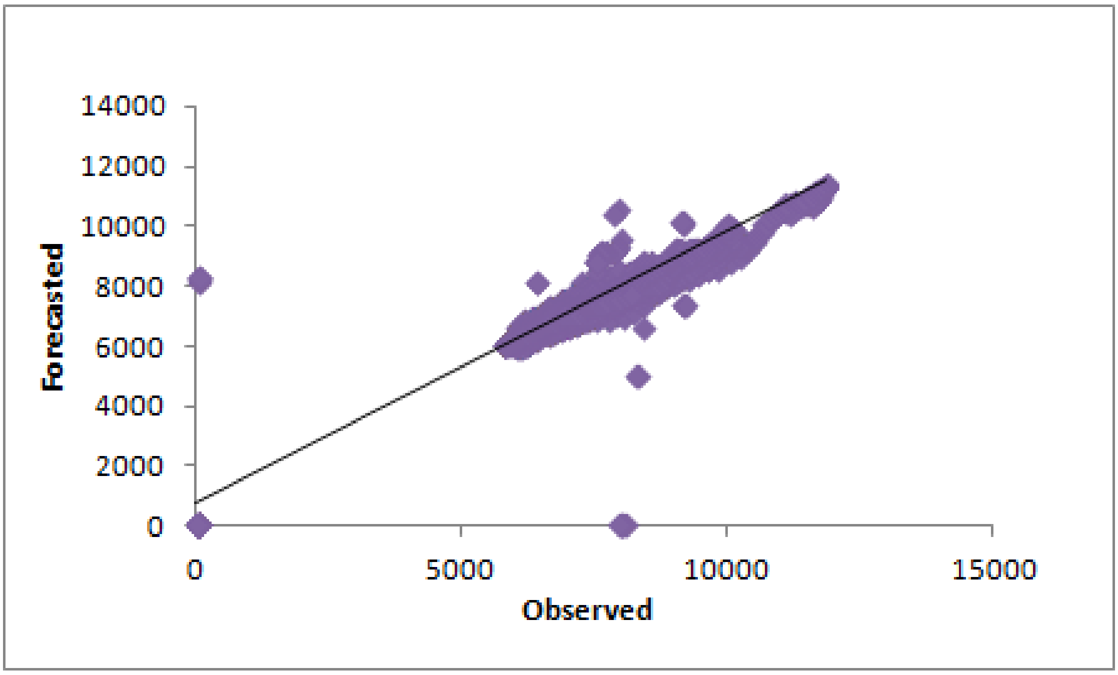


Figure 4.12: Data performance for 7 input with 5000 iteration.

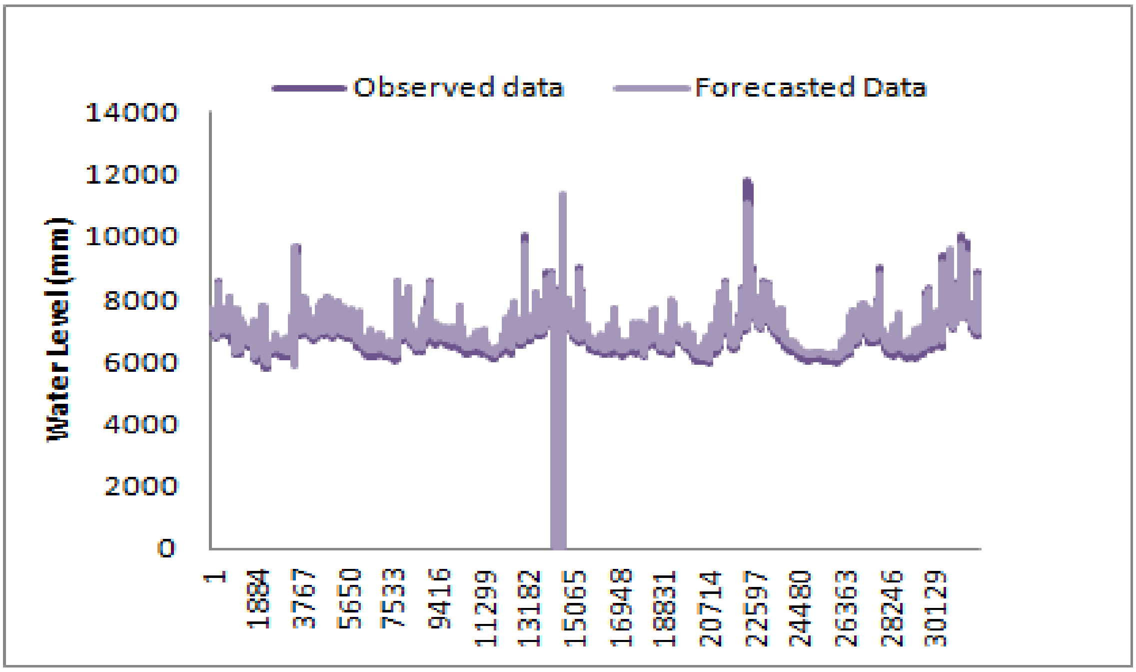
Table 4.2: Data training and data validation results for all inputs trained for 5000 iteration 5000 iteration 6 hour interval time.

Interval 6 Hour							
Input + output	Iteration	Validation DT	Validation DV	NSC		RMSE	
				DT	DV	DT	DV
2+1	5000	0.959	0.824	0.959	0.824	391.708	397.472
3+1	5000	0.962	0.837	0.962	0.837	379.437	382.107
4+1	5000	0.970	0.876	0.970	0.876	333.951	333.365
5+1	5000	0.972	0.885	0.972	0.885	322.773	320.422
6+1	5000	0.974	0.893	0.974	0.893	311.098	309.566
7+1	5000	0.979	0.907	0.979	0.907	281.443	288.332

4.2 3 HOUR INTERVAL TIME.

4.2.1 1000 Iteration.

Other than 6 hour, 3 hour interval times also been used in this thesis. Just like 6 hour interval time, 3 hour interval time also been trained with the same training set in this research. For each 2+1, 3+1, 4+1, 5+1, 6+1, and 7+1 input were also trained with 1000 iteration as the minimum iteration. Furthermore, Figure 4.13 until Figure 4.18 shows data validation results for six inputs that using 1000 iteration trained by utilizing Feed-forward Back Propagation ANN and also shows the error for each input would have. Table 4.3 shows that the validation DT of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 is 0.989, 0.994, 0.991, 0.994, 0.993 and 0.993 respectively. Meanwhile, validation DV of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 are 0.953, 0.970, 0.964, 0.975, 0.970 and 0.971 respectively. Besides, from the iteration it also had showed, the NSC values of data validation for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 0.953, 0.970, 0.964, 0.975, 0.970 and 0.971 respectively. Moreover, the RMSE values of data validation for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 204.971, 163.647, 180.221, 150.271, 163.086 and 161.169 respectively for the data validation. The results indicate that more input used in the training, more accurate the trained pattern we could get. It indicates that for the 2+1 input had more scattered data than 7+1 input data.



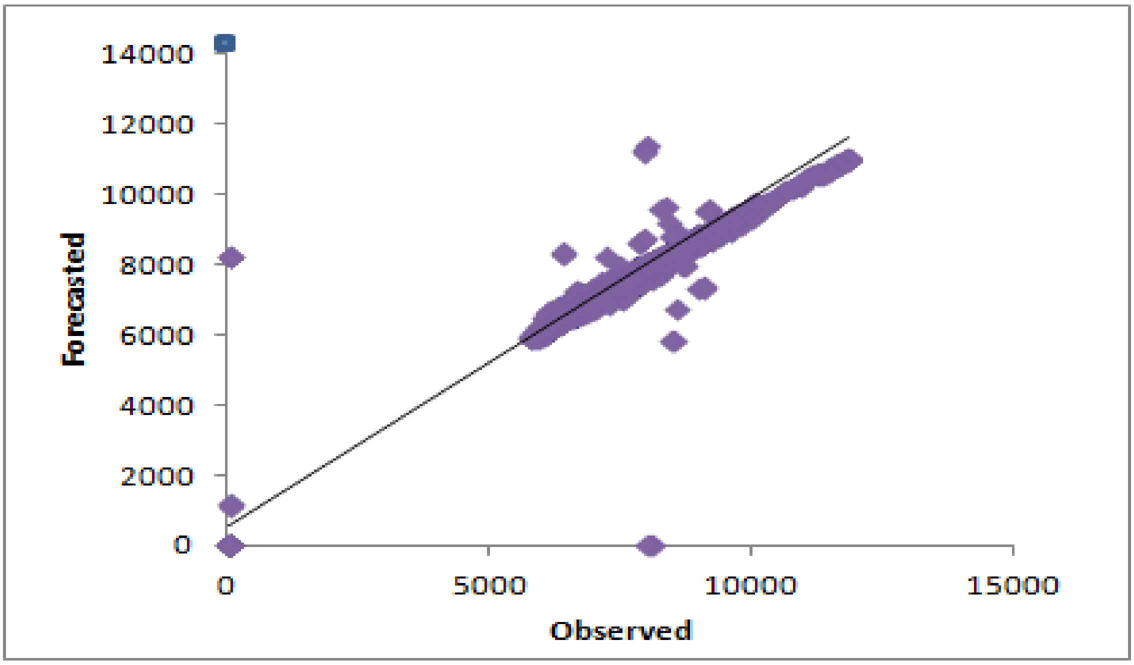


Figure 4.13: Data performance for 2 input with 1000 iteration.

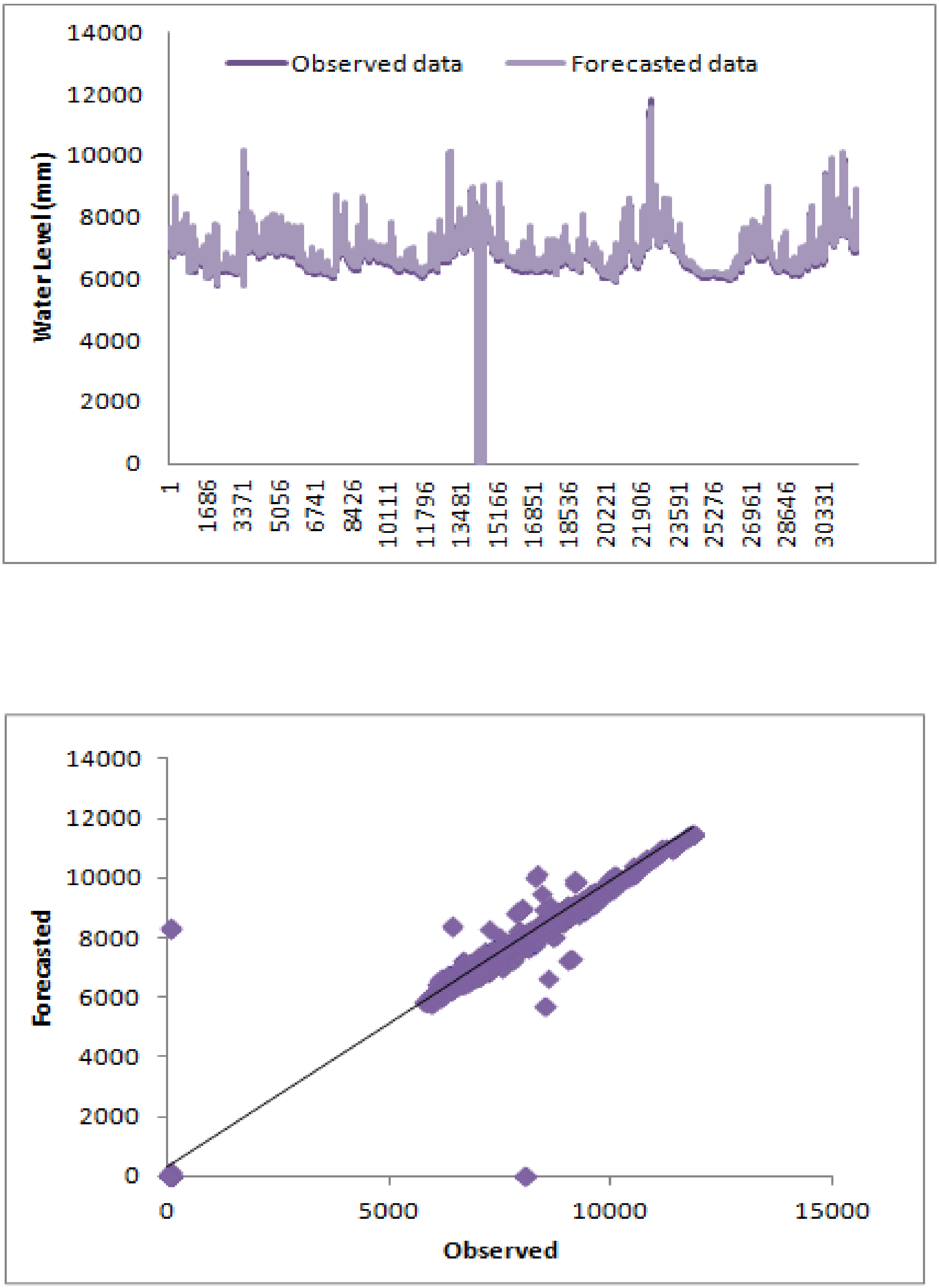
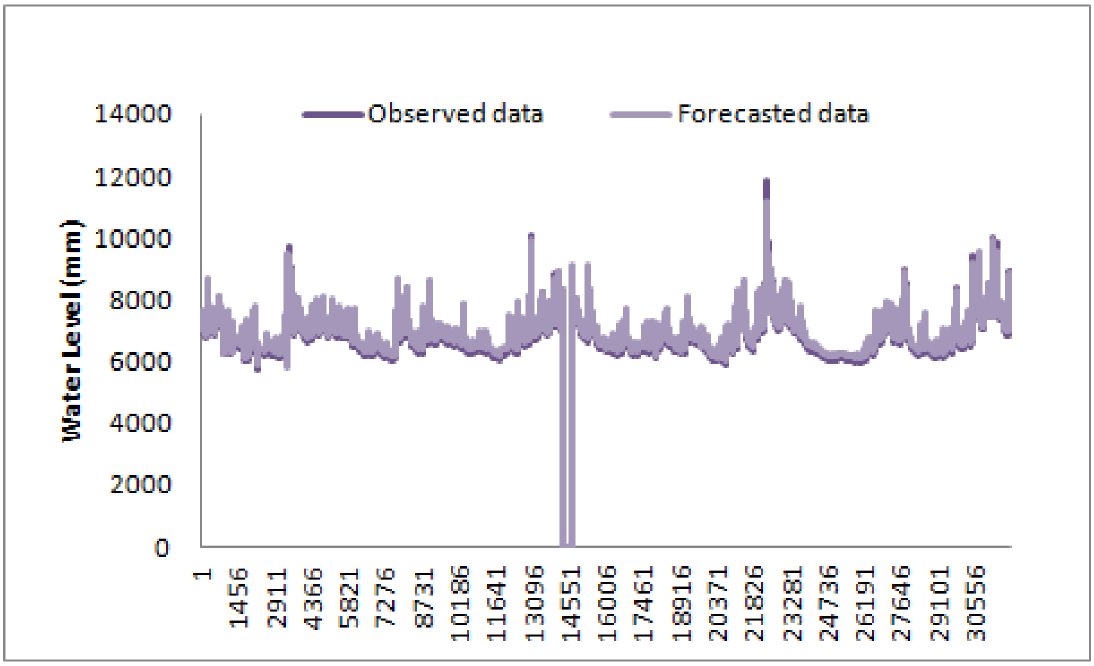


Figure 4.14: Data performance for 3 input with 1000 iteration.



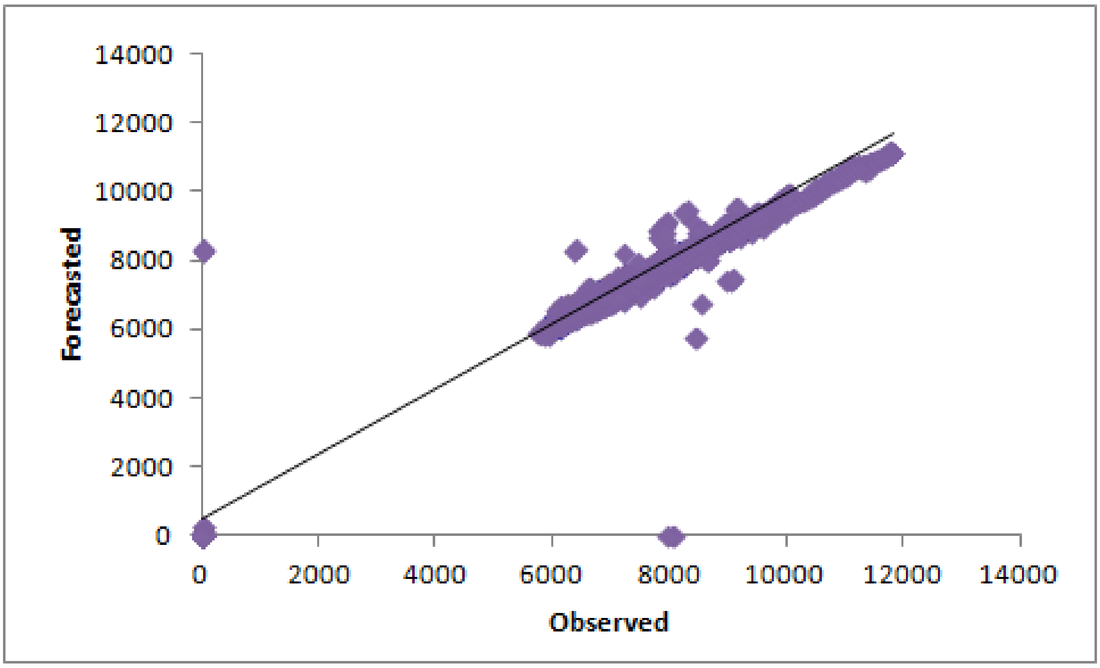
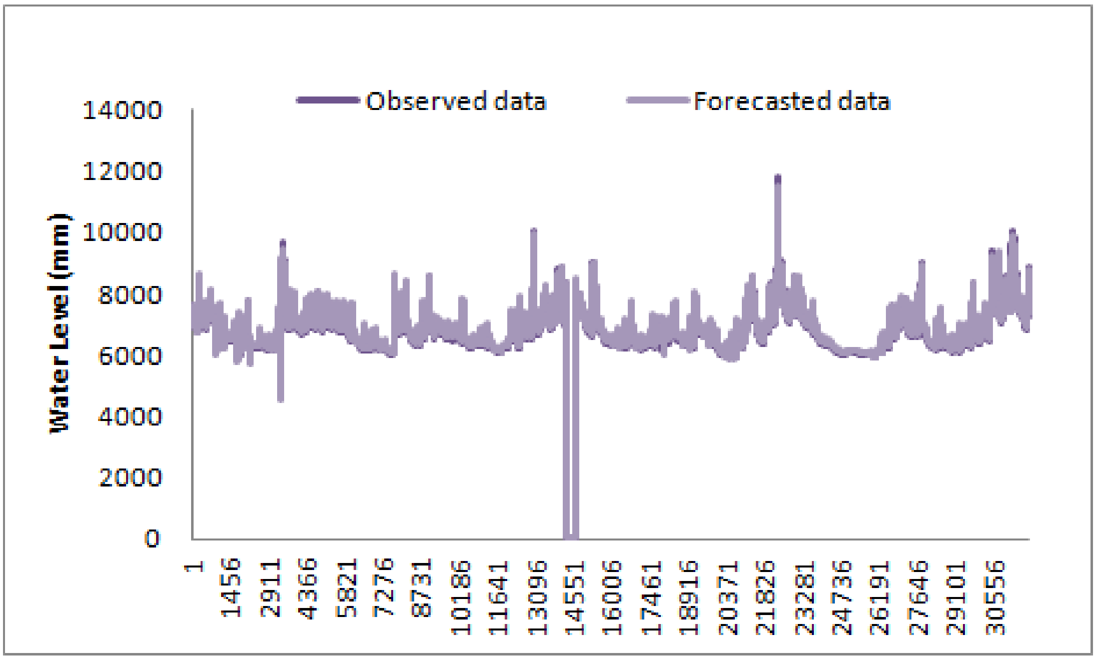


Figure 4.15: Data performance for 4 input with 1000 iteration.



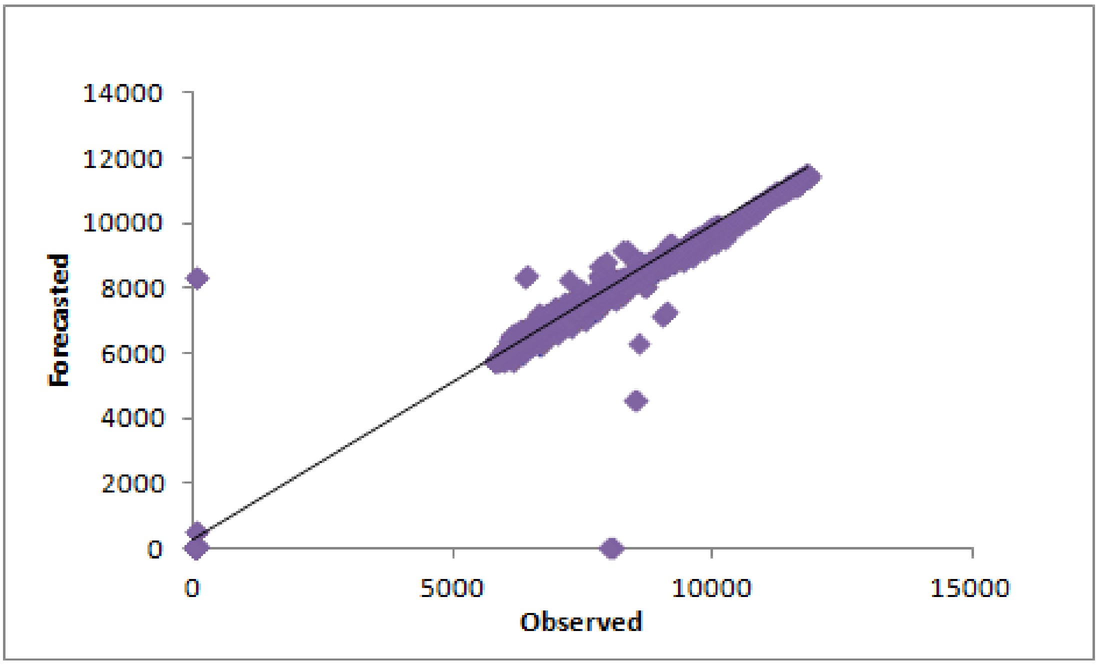


Figure 4.16: Data performance for 5 input with 1000 iteration.

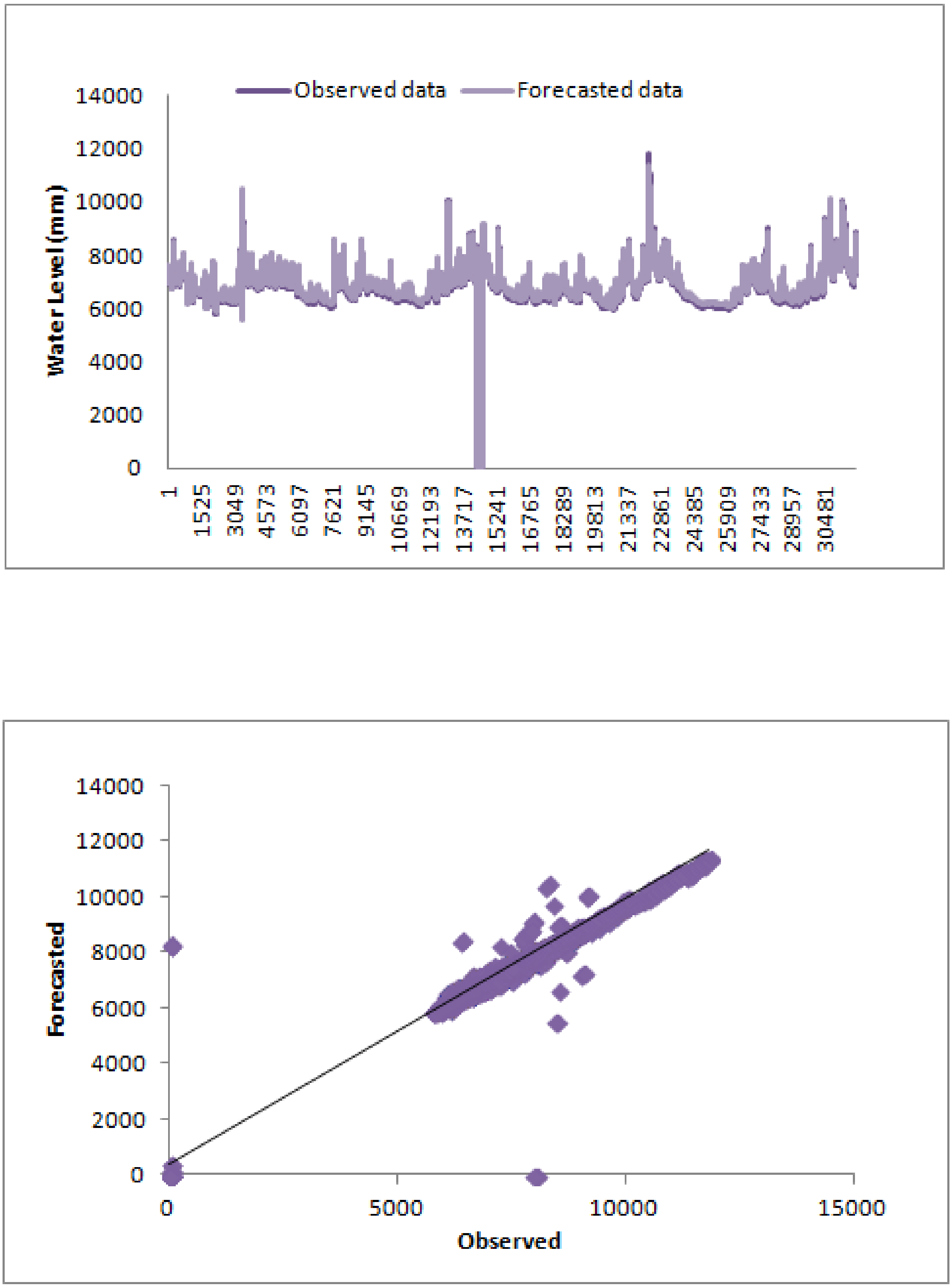
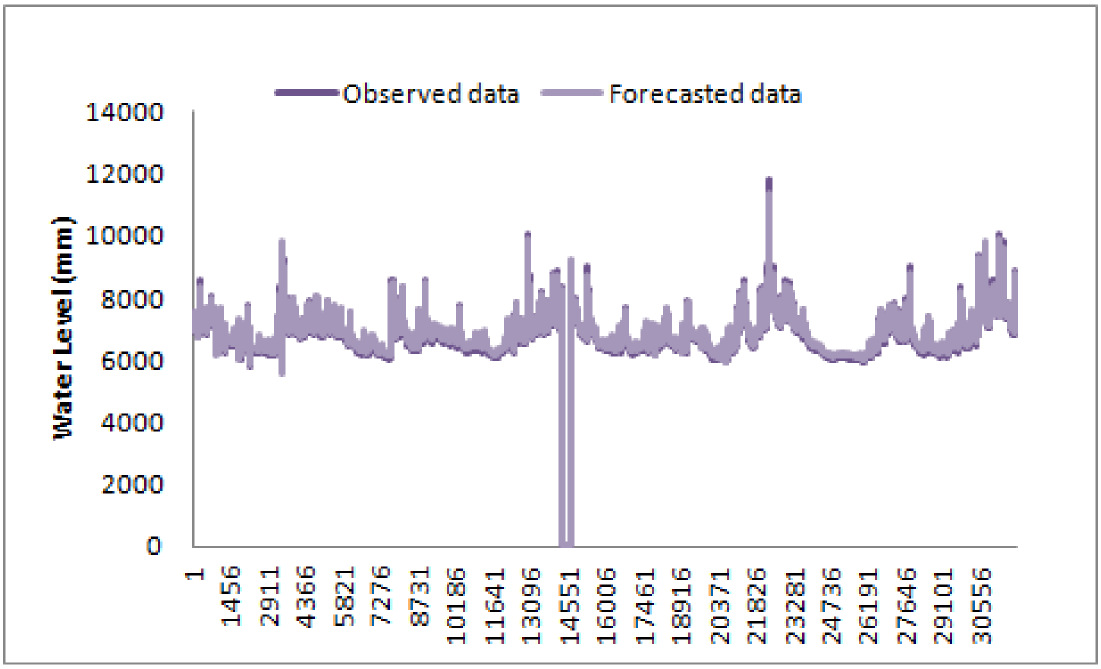


Figure 4.17: Data performance for 6 input with 1000 iteration.



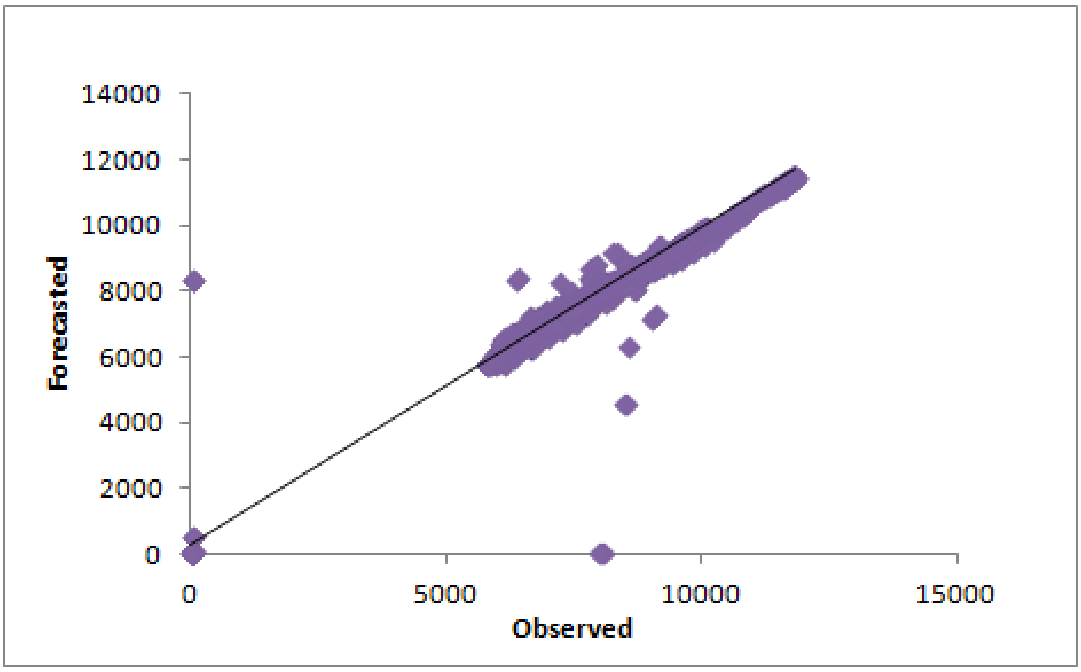


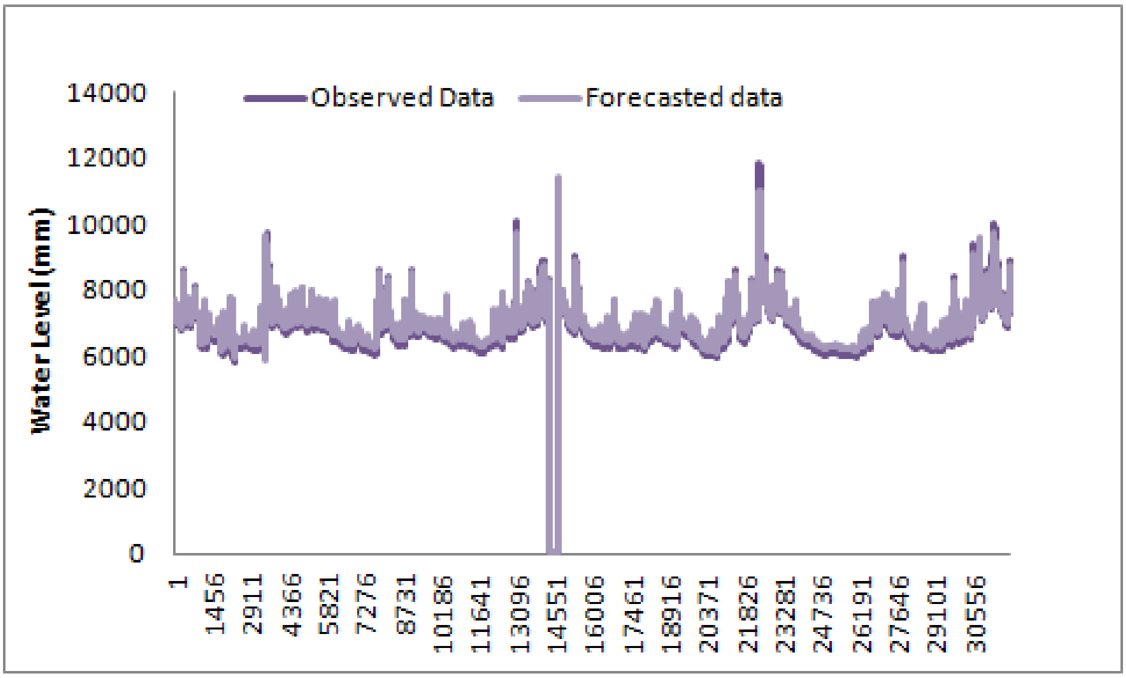
Figure 4.18: Data performance for 7 input with 1000 iteration.

Table 4.3: Data training and data validation results for all inputs trained for 1000 iteration 3 hour interval time.

Interval 3 Hour							
Input + output	Iteration	Validation DT	Validation DV	NSC		RMSE	
				DT	DV	DT	DV
2+1	1000	0.989	0.953	0.989	0.953	200.591	204.971
3+1	1000	0.994	0.97	0.994	0.970	154.046	163.647
4+1	1000	0.991	0.964	0.991	0.964	179.073	180.221
5+1	1000	0.994	0.975	0.994	0.975	145.224	150.271
6+1	1000	0.993	0.970	0.993	0.970	157.128	163.086
7+1	1000	0.993	0.971	0.993	0.971	167.209	161.169

4.2.2 5000 Iteration.

Besides than using 1000 iteration, 5000 iteration of 3 hour interval time also been used in this research. For each 2+1, 3+1, 4+1, 5+1, 6+1, and 7+1 input were also trained with 5000 iteration as the maximum iteration. Furthermore, Figure 4.19 until Figure 4.24 shows data validation results for six inputs that using 5000 iteration trained by utilizing Feed-forward Back Propagation ANN and also shows the error for each input would have. Table 4.4 shows that the validation DT of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 is 0.989, 0.992, 0.992, 0.995, 0.995 and 0.995 respectively. Meanwhile, validation DV of input 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 are 0.951, 0.965, 0.965, 0.977, 0.977 and 0.976 respectively. Moreover, the NSC values of data validation for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 0.951, 0.965, 0.965, 0.977, 0.977 and 0.976 respectively. Besides, from the iteration it also had showed the RMSE for 2+1, 3+1, 4+1, 5+1, 6+1 and 7+1 input are 209.575, 177.266, 178.317, 134.801, 143.799 and 145.779 respectively for the data validation. The results indicate that more input used in the training, more accurate the trained pattern we could get. It indicates that for the 2+1 input had more scattered data than 7+1 input data.



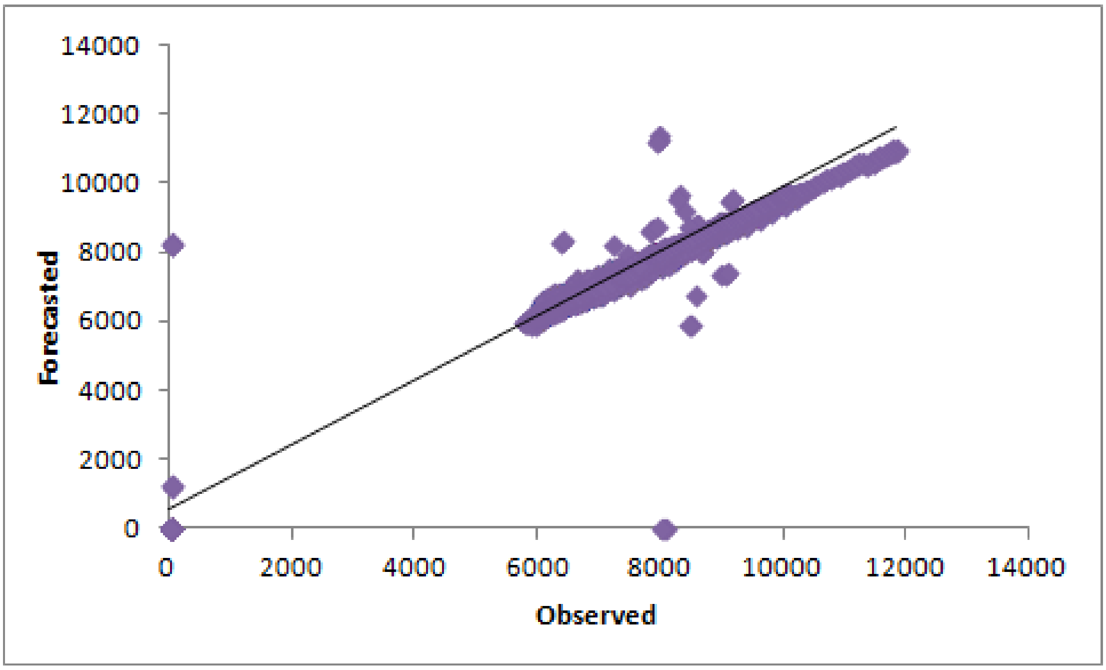
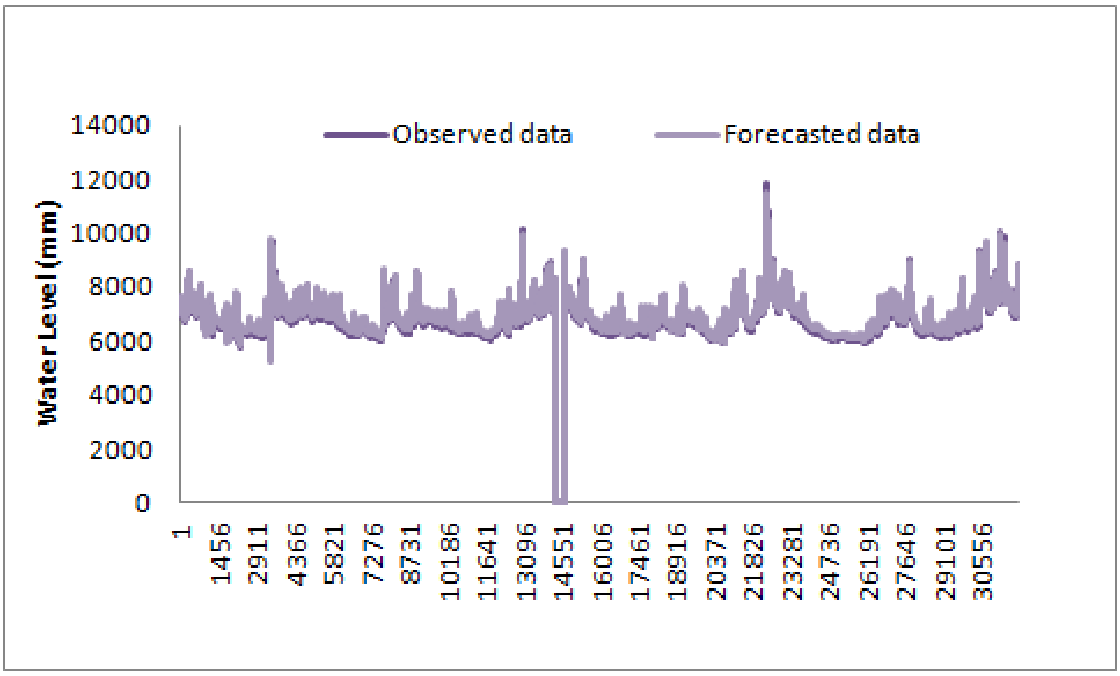


Figure 4.19: Data performance for 2 input with 5000 iteration.



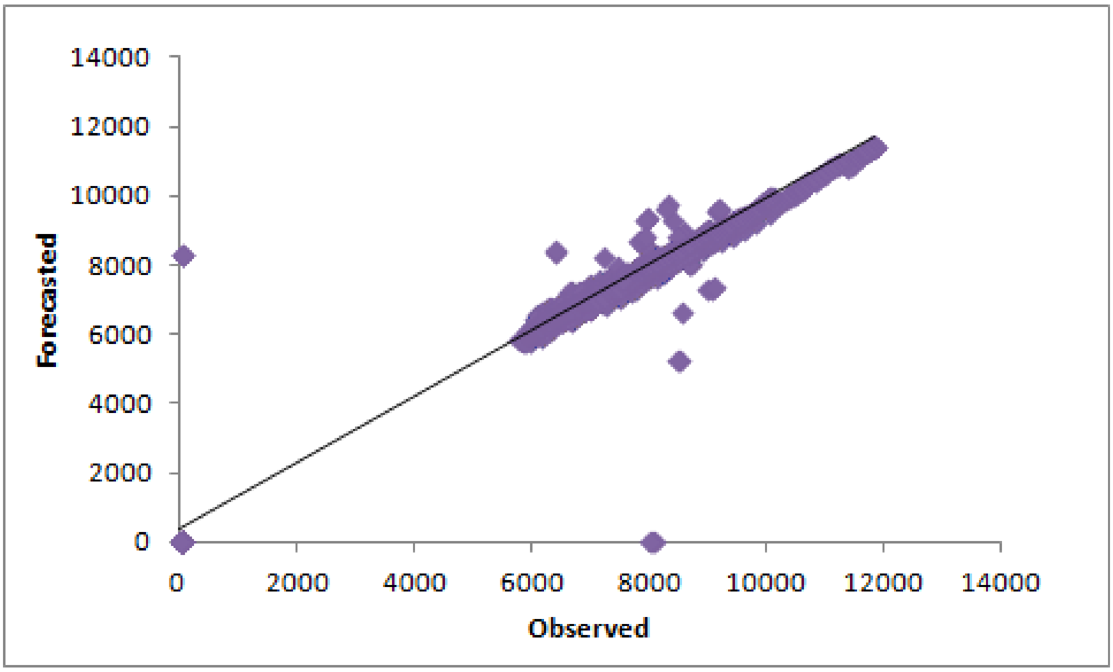
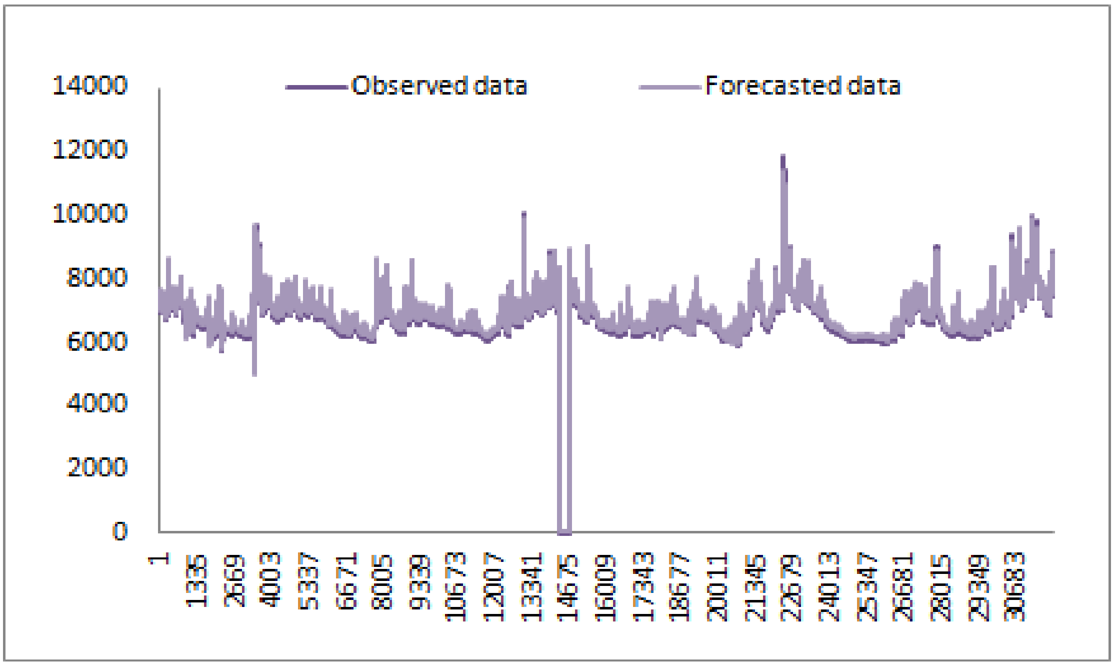


Figure 4.20: Data performance for 3 input with 5000 iteration.



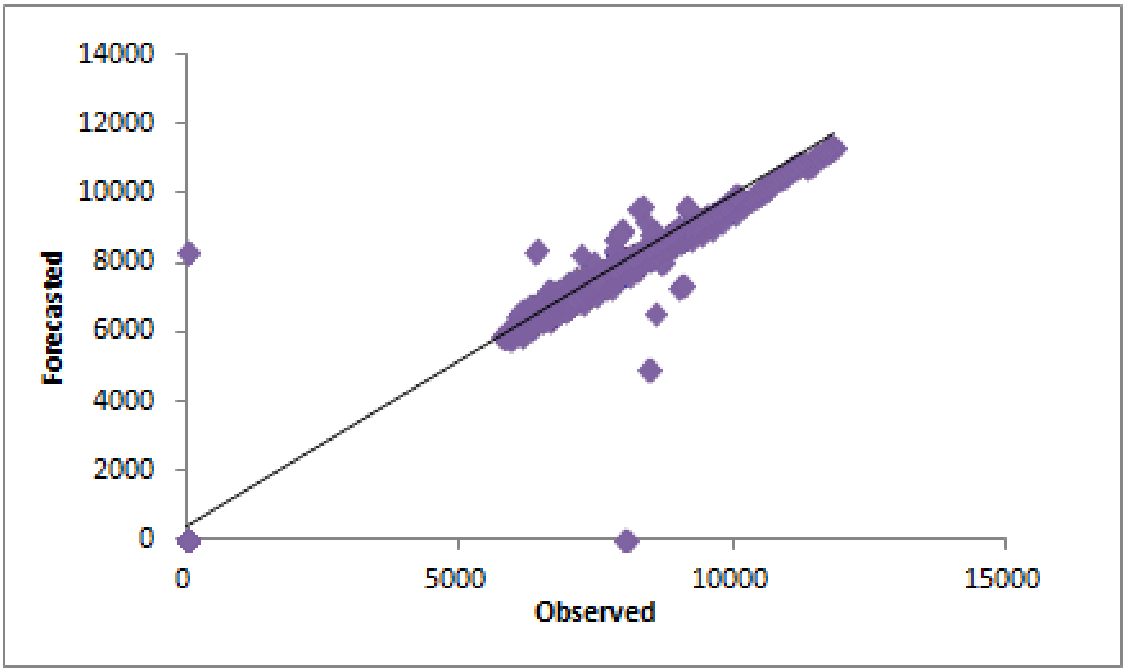
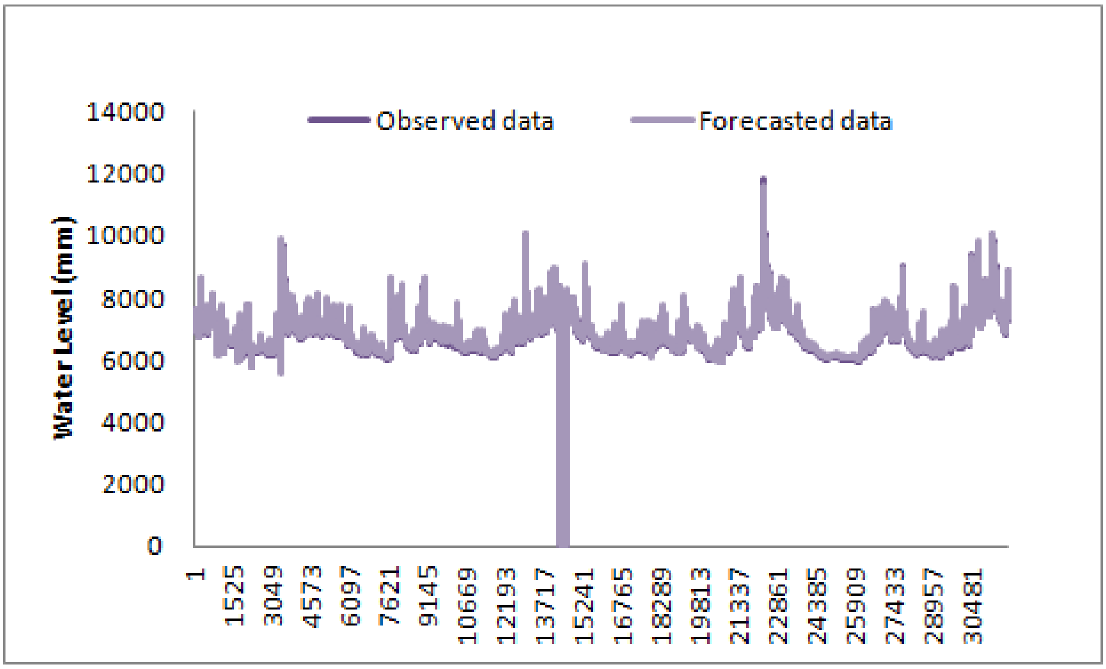


Figure 4.21: Data performance for 4 input with 5000 iteration.



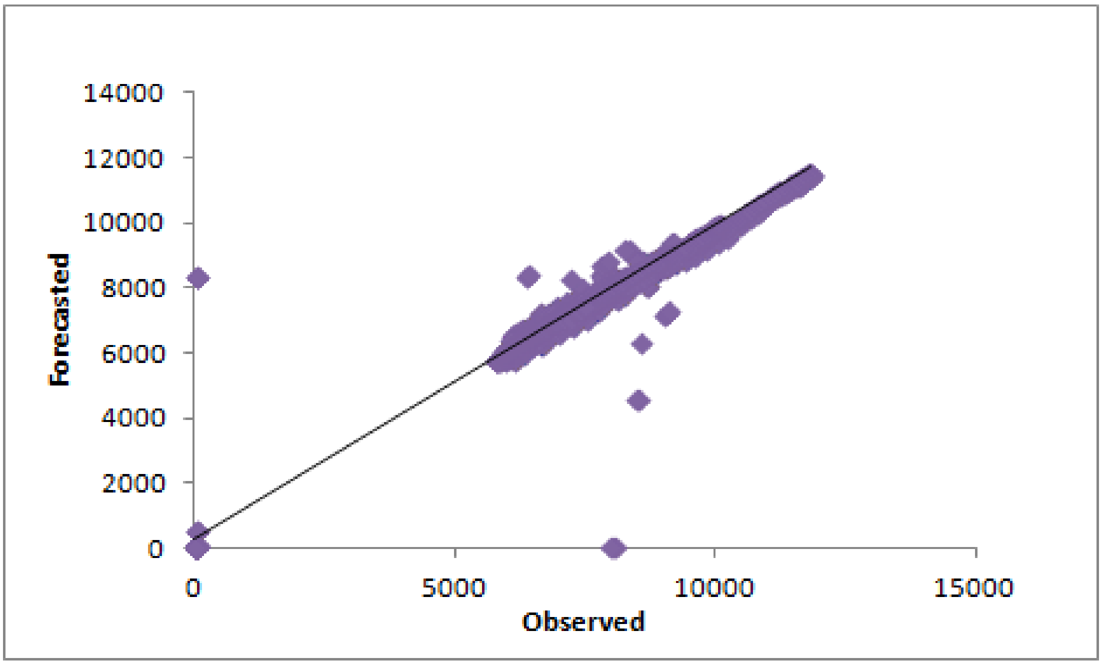
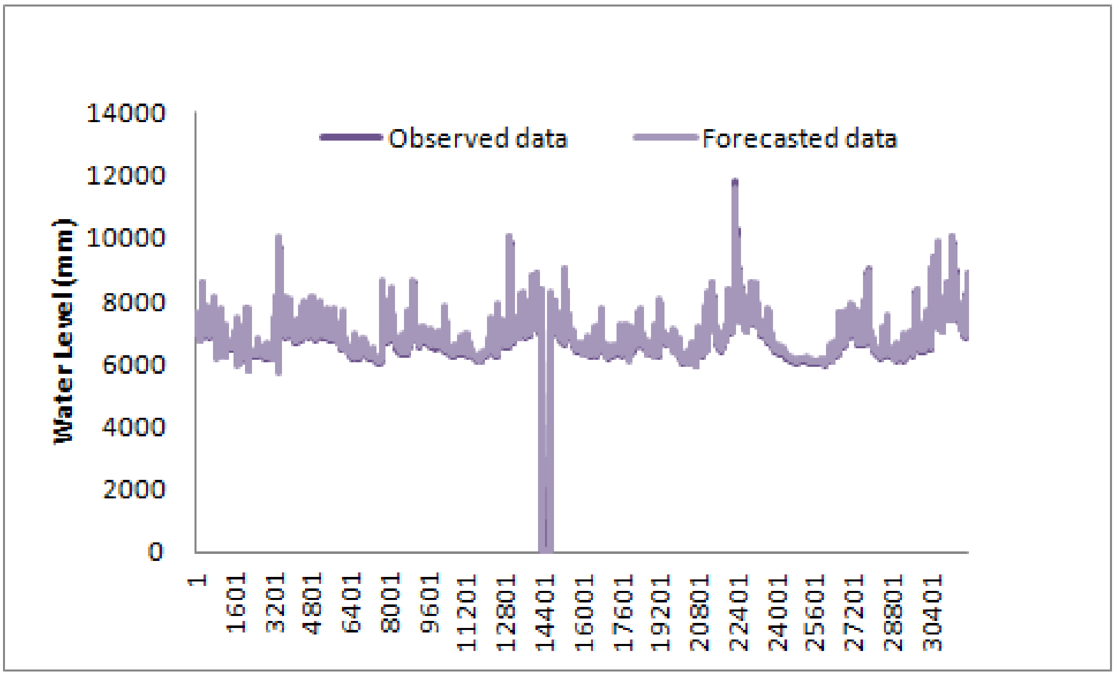


Figure 4.22: Data performance for 5 input with 5000 iteration.



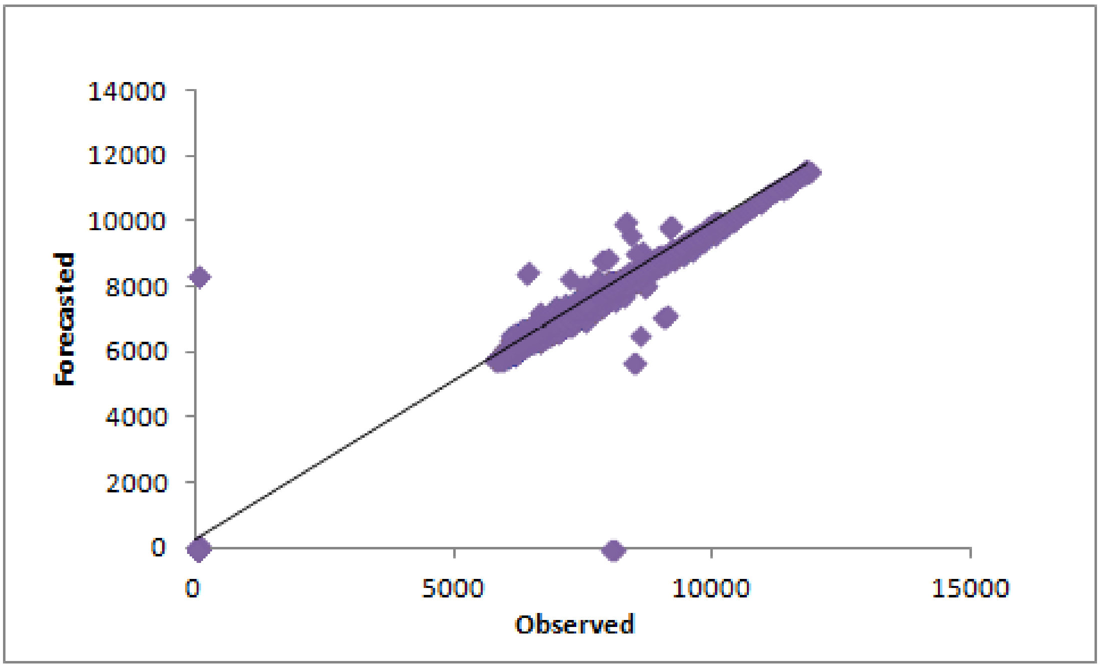
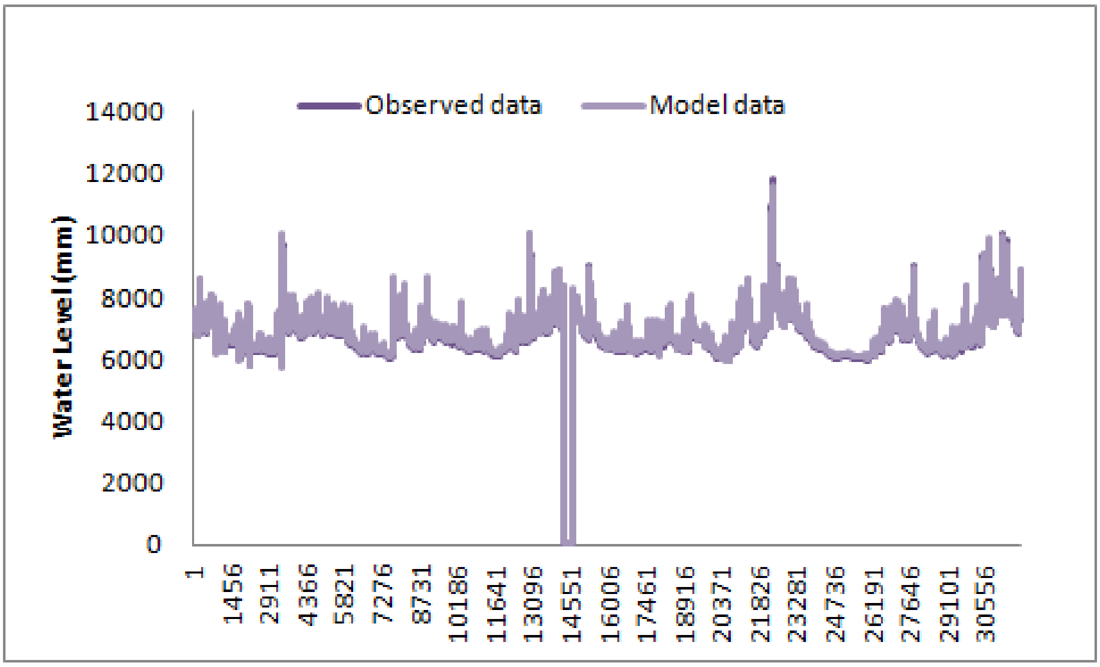


Figure 4.23: Data performance for 6 input with 5000 iteration.



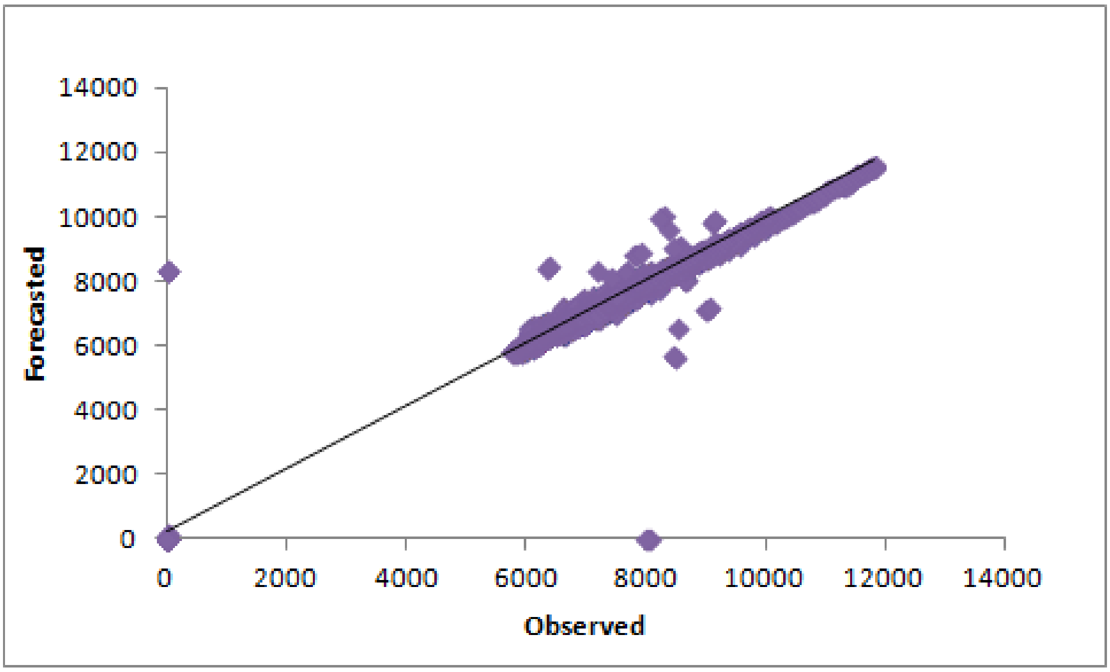


Figure 4.24: Data performance for 7 input with 5000 iteration.

Table 4.4: Data training and data validation results for all inputs trained for 5000 iteration 3 hour interval time.

Interval 3 Hour							
Input + output	Iteration	Validation DT	Validation DV	NSC		RMSE	
				DT	DV	DT	DV
2+1	5000	0.989	0.951	0.989	0.951	206.762	209.575
3+1	5000	0.992	0.965	0.992	0.965	169.858	177.266
4+1	5000	0.992	0.965	0.992	0.965	171..650	178.317
5+1	5000	0.995	0.977	0.995	0.977	145.099	134.801
6+1	5000	0.995	0.977	0.995	0.977	133.071	143.799
7+1	5000	0.995	0.976	0.995	0.976	137.045	145.779

CHAPTER 5

CONCLUSION

5.1 INTRODUCTION

5.1.2 Conclusion 6 Hour and 3 Hour Interval Time.

For 1000 iteration, the results indicate that the NSC value of 2+1 input data validation is 0.821, while for 7+1 input data is 0.905. Results from 5000 iteration indicate that the NSC value of 2+1 input data validation is 0.824, while for 7+1 input data is 0.907. It is clearly shows that the NSC values for both iterations are become higher nearest to the value of 1.0.Furthermore, for 1000 iteration, the RMSE value of 2+1 input data is 400.678, while for 7+1 input data is 292.357. The results from 5000 iteration indicate that, the RMSE value of 2+1 input data validation is 397.472, while for 7+1 input data is 288.332. Results from both iteration defined that the values of RMSE for both iteration are become lesser as it showed that, the error been produced from the data validation is getting lower.

As a conclusion, the results of this interval time demonstrated that the NSC had function efficiently in improved the ANN data training performance when compared with ANN data validation using NSC. Based on this interval time, models input between 5 and 7 had performed better data validation than those with models input between 2 and 4. Models with more input in high iteration had performed better than models with less input and low iteration.

REFERENCES

Dacup, J.A.and Silverman, D. 2000. Artificial Neural Networks and Long Range Precipitation Prediction in California. *Journal of Applied Meteorology*. 39:57- 66.

Department of Irrigation and Drainage, Kedah. 2013. Map of basin area of Sungai Muda, Kedah (online), http://gis.water.gov.my/gismuda/ (1 June 2015).

Elsafi, S.H., 2014.Artificial Neural Network (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*. 53: 655-662.

Eric Roberts' Sophomore College. 2000. The Intellectual Excitement of Computer Science (online). http://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/index.html

Jebur, M.N., Pradhan, B. and Tehrany, M.S. 2014. Flood Susceptibility Mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *Journal of Hydrology*. 512: 332-343.

Lim, H.S., MatJafri, M.Z. and Rajab, J.M.2013. Combining multiple regression and principle component analysis for accurate predictions for column ozone in Peninsular Malaysia. *Atmospheric Environment*. 71: 36-43.

Nastos, P.T., Paliatsos, A.G., Koukouletsos K.V., Larisi, I.K. and Moustris, K.P. 2014. Artificial Neural Networks Modeling for Forecasting The Maximum Daily Total Precipitation at Athens, Greece. *Atmospheric Research*.144:141-150

Population of City. 2013. Kuala Muda (online). http://www.citypopulation.de/php/malaysia-admin.php?adm2id=0204

Sulaiman, M. 2012. Reliable Water Level Forecasting Model Utilizing Artificial Neural Network, Johor River, Malaysia.

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

Universiti Malaysia Pahang
Lebuhraya Tun Razak, 26300 Gambang
Kuantan, Pahang Darul Makmur
Tel: +609-549 2999 | Faks/Fax: +609-549 2998
e-Mel/e-Mail : fkasa_admin@ump.edu.my

Fakulti Kejuruteraan Awam & Sumber Alam
Faculty of Civil Engineering & Earth Resources

UMP.13.02/13.11/1/3 Jld 20 (31) 26 Februari 2015

KEPADA SESIAPA YANG BERKENAAN,

Tuan/Puan,

PENGESAHAN PELAJAR MENJALANI PROJEK KURSUS

Dengan segala hormatnya, saya merujuk kepada perkara di atas.

2. Adalah dimaklumkan berdasarkan pengajian kursus, pelajar Fakulti Kejuruteraan Awam & Sumber Alam, Universiti Malaysia Pahang perlu mendapatkan data dan maklumat bagi menjalani projek kursus. Butiran projek dan pelajar adalah seperti berikut:

Tajuk Projek	Flood Estimation By Using Artificial Neural Network
Nama Kursus	BAA4914 – Final Year Project
Nama Penyelia/ Pensyarah	Dr. Muhammad @ S.A. Khusren bin Sulaiman (0199020060)
Nama Pelajar	1) Nurul Qamar Bt Roszaidi (AA11025) 2) Wan Nurul Hafizah Bt Abd Razak (AA11034) 3) Amir Aliff Bin Amri (AA11053) 4) Muhammad Afiq Bin Mustafa (AA11058) 5) Nurul Murshida Bt Mohd Sabri (AA11197)

3. Pihak FKASA amat berbesar hati sekiranya pihak tuan/puan dapat membantu dan menyalurkan maklumat dan data yang akan digunakan oleh pelajar bagi tujuan pembelajaran bagi masa hadapan negara amatlah dihargai oleh pihak kami.

Sekian, terima kasih.

"BERKHIDMAT UNTUK NEGARA"
Saya yang menjalankan tugas

ERNIE NURAZLIN BINTI LIZAM
Penolong Pendaftar Kanan
Fakulti Kejuruteraan Awam & Sumber Alam (FKASA)
Universiti Malaysia Pahang (UMP)
Tel : +6095492973/2999
Faks : +6095492998