PREDICTION OF FLOOD EVENT BY USING ARTIFICIAL NEURAL NETWORK AT PADAS RIVER

AMIR ALIFF BIN AMRI

Bachelor of Engineering (HONS) in Civil Engineering

UNIVERSITY MALAYSIA PAHANG

PREDICTION OF FLOOD EVENT BY USING ARTIFICIAL NEURAL NETWORK AT PADAS RIVER

AMIR ALIFF BIN AMRI

THESIS SUBMITTED IN FULFILLMENT FOR THE DEGREE OF BACHELOR

OF CIVIL ENGINEERING (HONS) IN CIVIL ENGINEERING

FACULTY OF CIVIL ENGINEERING
UNIVERSITI MALAYSIA PAHANG.

2015

UNIVERSITI MALAYSIA PAHANG

DECLARATION OF THESIS / UNDERGRADUATE PROJECT PAPER AND COPYRIGHT

Author's full name :AMIR ALIFF BIN AMRI

Date of birth :**15 JULY 1992**

Title :**PREDICTION OF FLOOD EVENT BY USING ARTIFICIAL NEURAL NETWORK AT PADAS RIVER**

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
92071-12-5483	MUHAMMAD@S A KHUSHREN BIN SULAIMAN
(NEW IC NO./PASSPORT NO.)	NAME OF SUPERVISOR
Date:	Date:

NOTES: * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction

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 :

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 : Amir Aliff bin Amri

ID Number : AA11053

Date :

ACKNOWLEDGEMENT

First of all, my appreciation and thanks to the Al-mighty Allah S.W.T for allowing me to and giving me the strength complete this thesis. Appreciation goes out to all my family especially my mother, Nurijam binti Ilma for her endless support and motivation which has given me the push to complete this thesis. Thanks also to all of my friends which has given their support whenever I needed it. Last but not least, a very big thank you to my supervisor Dr. Muhammad @ S.A Khushren for the teaching and guidance throughout the progress of the thesis.

ABSTRACT

For the past few years, the flood that hits Beaufort District became worst with no early warning given out leaving citizens unprepared. Such warning is needed and can be achieved through modelling. In this study water level of Padas River is being used to produce a model by using Artificial Neural Network to predict the occurrence of flood from the rising of water level in Sungai Padas. Artificial Neural Network with feed forward back propagation architecture are used to produce a model to where future water level are produced by using past water level. The model that has been produced in this thesis proved to be practical in doing real time water level prediction as model 3-4000-6 has given satisfactory accuracy in forecasting. Therefore early warnings can be given to unaware citizens on the rising of water level that will surpass the danger level that will lead to flood.

ABSTRAK

Untuk beberapa tahun kebelakangan ini , banjir yang melanda Daerah Beaufort menjadi semakin teruk tanpa amaran awal diberikan meninggalkan rakyat tidak bersedia. Amaran itu diperlukan dan boleh dicapai melalui pemodelan. Dalam kajian ini paras air sungai Padas digunakan untuk menghasilkan model dengan menggunakan Rangkaian Neural Buatan Manusia untuk meramalkan berlakunya banjir dari kenaikan paras air di Sungai Padas. Rangkaian Neural Buatan dengan makanan seni bina pembiakan kembali ke belakang digunakan untuk menghasilkan satu model di mana paras air di masa hadapan dihasilkan dengan menggunakan paras air tahun-tahun yang sebelum ini. Model yang telah dihasilkan di dalam tesis ini terbukti praktikal dalam melakukan ramalan paras air sdi mana model 3-4000-6 telah memberikan ketepatan yang memuaskan dalam ramalan. Oleh itu amaran awal boleh diberikan kepada rakyat mengenai kenaikan paras air yang melebihi paras bahaya yang akan membawa kepada banjir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

According to oxford online dictionary, the definition for flood:

"*An overflow of a large amount of water beyond its normal limits, especially over what is normally dry land*"

Flooding usually occurs when water from river or lakes overflows, in which it breaks the levees and it may also occur when rainwater accumulates on a saturated ground without enough area for infiltration. Some of it develops slowly, while flash floods can erupt in minutes. Climatic changes, technological and economic advancements that have altered the physical infrastructure of the earth, have now caused floods to occur recurrently (Elias, Hamin, & Othman, 2013).

In water management, flood is a concern. Besides heavy precipitation, poor drainage structure and its maintenance becomes a big contributor to the disaster. In central district areas, this is a huge concern as population is dense while vegetation is not. During precipitation, drainage network is the only reliable channel. Occasionally the precipitation is more than the network capacity.

Every year during the monsoon season, most states in Malaysia will receive heavy rainfall which leads to flood in several places especially in central business district areas. In the state of Sabah, Beaufort district is susceptible to flood due to its low elevation and situated near Padas River. Although the area is located along the river channel of Sungai Padas, heavy rainfall cannot be sustained due to small infiltration area which are caused by rapid building development which takes away vegetation. In addition, the shallowness of Padas River estuary due to sedimentation of silt and various human activities, combined with poor drainage system adds up to the problem.

To improve the management of this river, a forecasting model is fundamental to help monitor the river physically. An accurate model can help to forecast imminent flooding and lessen the aftermaths. One of the most feasible approaches in doing so is by using computer modelling.

Computer modelling has been widely used to solve problems in doing research and investigation. It enables man to investigate with short time the changes in real life situation and its effects that takes place over years. Another benefit of computer modelling is that investigation can be done and results can be achieved without having to construct a physical prototype of the research, while saving a tremendous amount of time.. The advancement in computer technology has sprouted a vast amount computer models to predict outcomes where the measurements of the results can be controlled by controlling the parameter.

Artificial neural network is a model that has been used for over 70 years in various fields of research. Networks are used to model a wide range of phenomena in physics, computer science, biochemistry, ethology, mathematics, sociology, economics, telecommunications, and many other areas (Gershenson, 2003). ANN modelling proves to be very gainful in hydrological modelling especially in prediction of rainfall runoff because in the past two decades, ANN has been a practical method for non-linear modelling, for example rainfall-runoff modelling (Farajzadeh, Fard, & Lotfi, 2014), flood forecasting (Nasr, Moustafa, Seif, & El Kobrosy, 2012), wastewater treatment plant performance prediction (Nasr et al., 2012) and etc.

1.2 Problem Statement

For the past few years, the flood that hits the area of Beaufort became worst which cause tremendous amount of property loss. Even though the authorities have taken steps to cope with the problem such as deepening of the river bed of Padas River, flooding still worsen. During this time, the residential area in Beaufort will be cut out of food and resources such as electricity. They are also rendered unable to seek help as the roads are inaccessible.

These citizens will not be left in distress if the authorities were able to give warning of the upcoming flood. Flood estimation is required to give such warning. However, such warning are unable to be announced due to lack of flood estimation equipment and data regarding previous flood level and river channel (Hussin, 2010).

Past study regarding flood event prediction using physical modelling such as regression model (Mediero & Kjeldsen, 2014), hydrodynamic modelling (Gallien, Sanders, & Flick, 2014) and the process in gaining data is costly and time consuming.

Therefore, for such estimation Artificial Neural Network will be used as it is able to process data at high speed and are capable to form a general solution to a problem from a specific set of data (Bishop, 1994) It can perform in the manner of a 'black box model', not requiring detailed information about the system and mathematical algorithm for model build up. This means by using ANN, only a single data type is needed to construct a model. However, does the model are able to be constructed with high accuracy? Does the model can be applied to any given data set other than the ones it is constructed upon? This research is conducted to determine answers for the questions aroused.

1.3 Objectives of Study

The main objective of this study is:

1. To produce a model for predicting flood event at Padas River, Beaufort station.

The main objective is achieved by completing these secondary objectives:

1. To determine if the water level data can be used alone in Artificial Neural Network modelling.

2. To develop a model of high accuracy using Artificial Neural Network.

1.4 Scope of Study

1.4.1 Area of Study

The area of study is Padas River, Beaufort. It is located in the internal division of Sabah, and located 90km south west from Kota Kinabalu, the capital city of Sabah. The coordinates to the central business district is given as 5°20′0″N 115°45′0″E. Beaufort is a quiet provincial town with population estimated to be 64,350 at year 2010. The total area covered by the district is 1735 km^2, which results in the density being 37 citizens per kilometre square.

FIGURE 1. TOWN OF BEAUFORT WITH PADAS RIVER

The Padas River goes through the middle part of the Beaufort distrcit and town and used to serve as a mode of transportation where citizens would go to town by small boats and canoes. It also serves as the source of fresh water fishery as the district is located far from the shoreline.

The water level is taken from the station 5357403 in Sungai Padas river in Beaufort.

FIGURE 2 SHOWS THE BEAUFORT WATER LEVEL STATION OF PADAS RIVER. (PICTURE IS COURTESY OF INFOBANJIR.GOV)

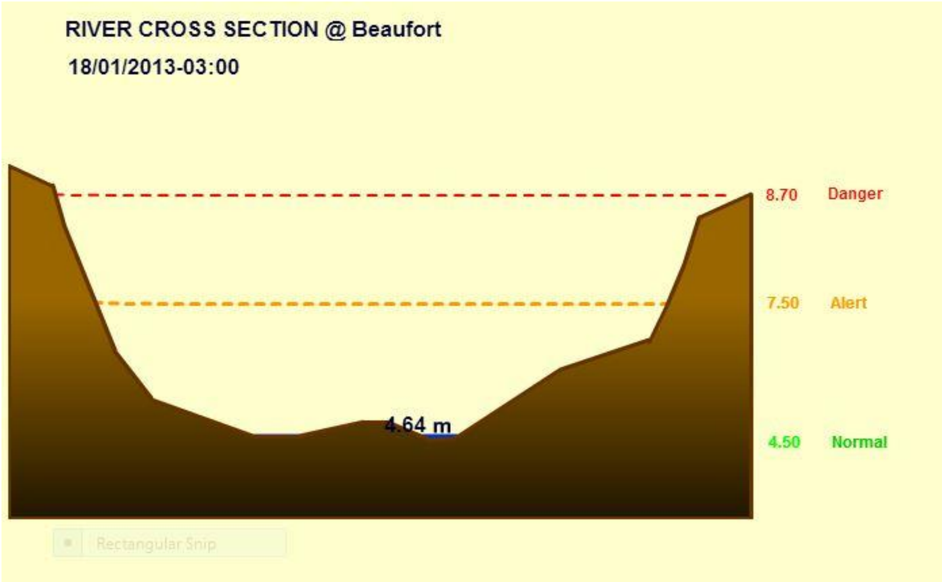


FIGURE 3 SHOWS THE CROSS SECTION OF THE RIVER AT THE BEAUFORT WATER LEVEL STATION. (PICTURE IS COURTESY OF INFOBANJIR.GOV)

The Padas River has a a danger level of 8.7 meters and 7.5 meter alert level. In this study, the water level taken from this station is to be used to be fed as inputs to produce a model.

1.4.2 Modelling Tool

The predicting tool to be used in this research is feed-forward back-propagation Artificial Neural Network to obtain a model for water level data that are obtained from local authorities. Python programming language is used to construct the network system, and it will be made up of 1 input network, 1 output network and several hidden layers.

1.4.3 Water level data

Water level data of Padas River will be obtained as a parameter to construct the forecasting model in the Artificial Neural Network. The data period interval will be hourly and the model constructed will also follow the same period interval.

1.5 Expected Outcomes

This research is expected to produce a high accuracy model for predicting flood at Padas River, Beaufort by using Artificial Neural Network. An hourly model of water level is to be established along with its performance measure.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

The discussion in this chapter is regarding the methods and material/equipment used in the past study done by researchers in flood prediction and artificial neural network. It will serve as a guideline in conducting this research.

2.1 Flood

2.1.1 Introduction

The word flood comes from the Old English word *flod*, and is similar when compared to German language *Flut*, and the Dutch *Vloed*. In Malaysia, flood is one of the common environmental disaster. The period in forming of flood varies. Some floods takes time to elevate, while some can happen without visible signs of heavy rain such as flash floods. In terms of area, flood can happen locally, only affecting neighbourhoods and small communities, or it can grow very large affecting the whole river basin and multiple states in one time.

2.1.2 Factors causing flood

There are many contributors to flood development, and the most deadly effects come from the deed of human that changes the structure of the soil surface.

Insufficient permeable ground surface

Most floods happen on low land areas, where the ground gets saturated quickly during heavy precipitation. Runoff and infiltration cannot happen as quickly as accumulation of water, therefore resulting in flood. This happens when there is not enough permeable surfaces, such as soil to hold up the precipitation. Developed urban areas with most of the areas covered in pavement such as concrete and bitumen roads have little permeable soil that inhibits the water dissipation into the ground, and thus making this area prone to flood.

Urbanization

In urban areas, flooding is called urban flooding where the amount of precipitation overwhelms the capacity of the drainage systems, specifically storm sewers. Urban flooding is a condition where it happens repetitively and impacts communities systemically. Urbanization also affects the structure of the natural river present in the vicinity of its area. Development of buildings will shrink the river section and pollution decreases the flow of water due to impediments.

Shallow river section

River is more effective in depth compared to width (Vishwas, 2002). Most rivers in urban areas are affected negatively by development, on other words, becoming shallower. Kinabatangan river can be taken as an example where plantation of palm oil along it has accumulated silt from the act of deforestation and this has lessen the depth of the river. This ultimately leads to flash flood occurring every time heavy rain falls in the area of Kinabatangan.

2.2 Artificial Neural Network

2.2.1 Biological Neural Network

Artificial Neural Network, ANN is a computational model that was inspired by the neuron cells on the central nervous system. The central nervous system is responsible for processing and integrating sensory information, which can be classified as visual, auditory, somatic sensation, gustatory, olfaction and vestibular. The nervous system consist the spinal cord and the brain, which contains around 10 billion neuron/nerve cells, an electrically active cell.

A neuron cell process and transmit information to another similar cell through chemical and electrical processes. This activity occurs via synapses, where it is the link between neuron cells.

A neuron receives input from the dendrite branch, when it exceeds a threshold it will send out an electrical pulse that will travel through the body known as soma then down to the axon, where the electrical pulse will be transferred to another neuron through the release of neurotransmitters, a chemical released to bind the next neuron cell.

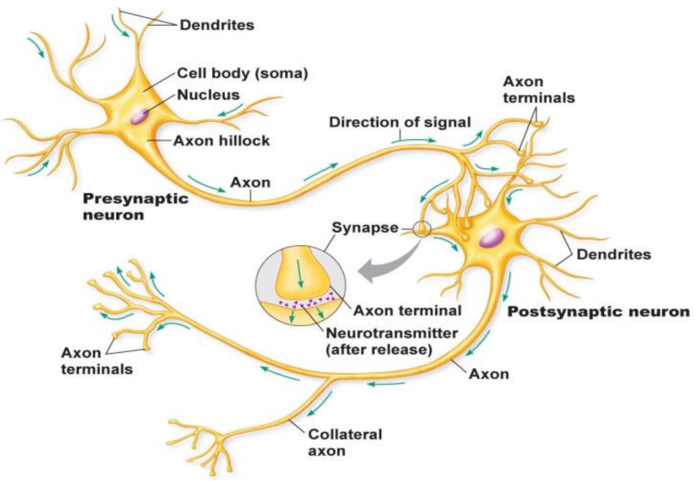


FIGURE 4 A SCHEMATIC ILLUSTRATION OF 2 BIOLOGICAL NEURON. DENDRITES RECEIVES INPUTS AND FIRE ELECTRONIC IMPULSES ALONG THE IN THE DIRECTION SHOWN BY THE ARROWS. OUTPUTS FROM THE AXON PROPAGATES THROUGH A JUNCTION CALLED SYNAPSE.

Every occurring synapse has its own strength, which its counterpart in ANN is weights, and it determines its effect on the post-synaptic neuron. This forges the essential attributes to both neural networks where they are able to alter response when subjected to external signals (Bishop, 1994). Those signals from the synapse might be transmitted to other synapse and activates other neurons.

2.2.2 Artificial Neural Network

2.2.2.1 Brief History on Artificial Neural Network

The dawn of neural networks computing starts in the 1940's by McCulloch and Pitts with their seminal paper, *A Logical Calculus of the Ideas Immanent in Nervous Activity.* They revealed that neural computing can be used universally, to imitate any conventional computing machine (Bishop, 1994). After their first model, a plethora amount of model has been established with differences in activation functions, learning functions, structure and etc., with each suitable for a specific categories of problem.

The first neural hardware called the perceptron was developed by Frank Rosenblatt in the 1950’s that was based on McCulloch-Pitt neuron model. The perceptron model can determine characters that is fed to the inputs as pixelated images (Rosenblatt, 1962). Similar network were also develop by Widrow, called the Adaptive Linear Element (ADALINE) network which uses the Widrow-Hoff learning rule.

There is a mass array of different types of networks that has been developed since the McCulloch-Pitt, but all of them share the same constituent; a set of nodes and the connection between them. The nodes or the neuron performs the computation. They receive inputs, process them and provide outputs. The computation process done by the neurons can be as simple as summing the inputs or it can be another network inside the neuron.

The interaction between neurons prompted the global behaviour of the network, which cannot be distinguished in the elements of the network. This kind of behaviour is affirmed to be emergent (Gershenson, 2003). This result in the network being a potent instrument, one which supersede its element abilities.

2.2.2.2 Application of Artificial Neural Network

Artificial neural network has been used mainly in fields of information processing, such as computer science, pattern recognition, robotics, artificial intelligence, hydrology and etc. It also can even be found in electronics such as data compression in computers and auto exposure in digital cameras. In case of hydrology, most cases have been prediction of flood, water quality and water flow. A.A Masrur Ahmed (2013) developed a feed forward neural network combined with radial basis function neural network to predict the dissolved oxygen from biochemical oxygen demand(BOD) and chemical oxygen demand(COD) in the Surma River, Bangladesh. He revealed that his proposed ANN models can be used successfully to with acceptable accuracy.

Y. Wei et al (2002) used neural network as a predictive method for flood disaster problem. They stated by the results that the use of ANN with function approximating is a valid approach, simple and reliable. The proposed method is applicable to investigate the issue with variables with ambiguous relationship since its mapping approximation process is not connected with the types of variable relationship.

Other than that, J. Farajzadeh et al (2014) uses feed forward neural network combined with time series analysis to model the monthly rainfall and runoff at Urmia lake basin in Iran. She however concludes that the feed forward neural network couldn’t be well trained due to low inflow data that were caused by low precipitation and dam constructions.

CHAPTER 3

METHODOLOGY

3.0 Introduction

This chapter will discuss about the methods of collecting data, methods of analyzing data and the modelling to be done. This research will focus on using Artificial Neural Network to analyze data and come up with a model for the purpose of prediction. This chapter will serve an important role in implementing this research study accordingly inside the scope of study.

3.1 Artificial Neural Network

The data analysis tool to be used in this research is Feed-Foward Back Propagation Artificial Neural Network, FF-ANN. The feed forward network, which can be considered as common purpose non-linear functions for aligning two sets of variables, form the basis for most present day applications.

The coding tool to provide this network in this research is python programming tool. It is a widely used, general purpose high-level programming language. It can be used to program multiple models including object-oriented programming, imperative programming and functional programming.

The ability of a network to generalize and its complexity depends on the network topology. The network is design by trial and error. A larger-than necessary network are bound to over-fit the training data, where smaller-than necessary network will be hard to train.

3.1.1 Feed Forward Neural Network

The model of neural network that will be used in this research is the feed forward neural network that was popularized by David Rumelhart. The reason of this decision is because of the success shown by the network in construction forecasting model, and the simplicity in implementation.

The data will be served at the input neuron, where it will be multiplied by a certain weight which will be the strength of the respective signals, and then computed by an activation function. When it reaches a certain threshold, the output neuron will compute the data to give an output. There can be numerous hidden layers, but there can be only a single input and output layer in a single network.

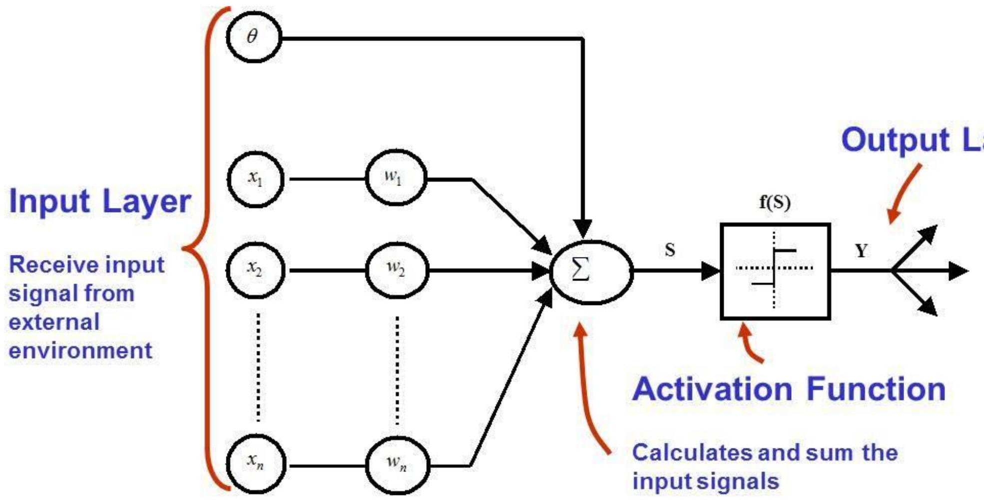


FIGURE 5 *MCCULLOCH-PITTS MODEL OF A SINGLE NEURON THAT SUMS THE INPUTS X_N, THEN TRANSFORMS THE SUMS OF INPUTS USING A NON-LINEAR ACTIVATION FUNCTION TO GIVE A FINAL OUTPUT.*

The weight determines the strength of the input multiplied by it. It can be either in positive or negative value. The computation of the neuron differ according to the weight. The output can be adjusted to the value intended by adjusting the weights multiplying the inputs. However for inputs of hundreds and thousands, the weights are adjusted by using algorithms to obtain the desired output, which is called training or learning. In this research, the algorithm used is back-propagation algorithm, that will be discussed further in the next topic.

3.1.1.1 Non-linear mapping function

In this research, the feed forward network is regarded as a general purpose non-linear function for mapping two sets of variable. The networks can be better understood by learning the mapping structure of it. Figure 4 shows a non-linear function which takes *n* input variables $x,..,x_n$, and then maps them onto *m* output variable $y,..y_m$. In neural networks, variables x are called inputs while variables y represents outputs.

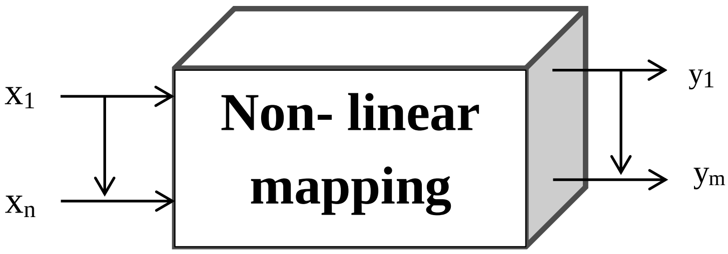


FIGURE 6 A SCHEMATIC ILLUSTRATION SHOWING A NON-LINEAR FUNCTION MAPPING *N* X VARIABLES ONTO *M* Y VARIABLES

3.1.1.2 Analogies of ANN with polynomial curve fitting

When input and output variables are collected and respective vectors are formed, which are denoted as $x = (x_1, x_2, x_3 \dots x_d)$ and $y = (y_1, y_2, y_3 \dots y_d)$, the precise model that maps x onto y is determined by the architecture of the network and the values of weights multiplier, denoted as $w_1 ,w_2 ,w_3 \dots w_d$. Thus the network mapping can be written in the form y = y(x;w), which describes y as a function of x that is multiplied by parameter w.

Let's take into account the nth order polynomial equation

$$y = w^1x^1 + w^2x^2 + \cdots + w^nx^n + w^0 = \sum_{j=0}^{m} w_j x^j \qquad (1)$$

This equation can be considered as a non-linear mapping function which accepts x as an input variables and gives out y as an output variable. The accuracy of the equation is strictly determined by the values of w^0,…w^n, parameter, which are analogous to the weight parameter in the network, where w^0 in the network functions as a bias weight parameter. The bias weight parameter usually holds a value of 1 and is present in the input layer and each of the available hidden layers so that the value of the forecasted output value is in the range of 0 to 1. The value of 1 acts as the threshold value.

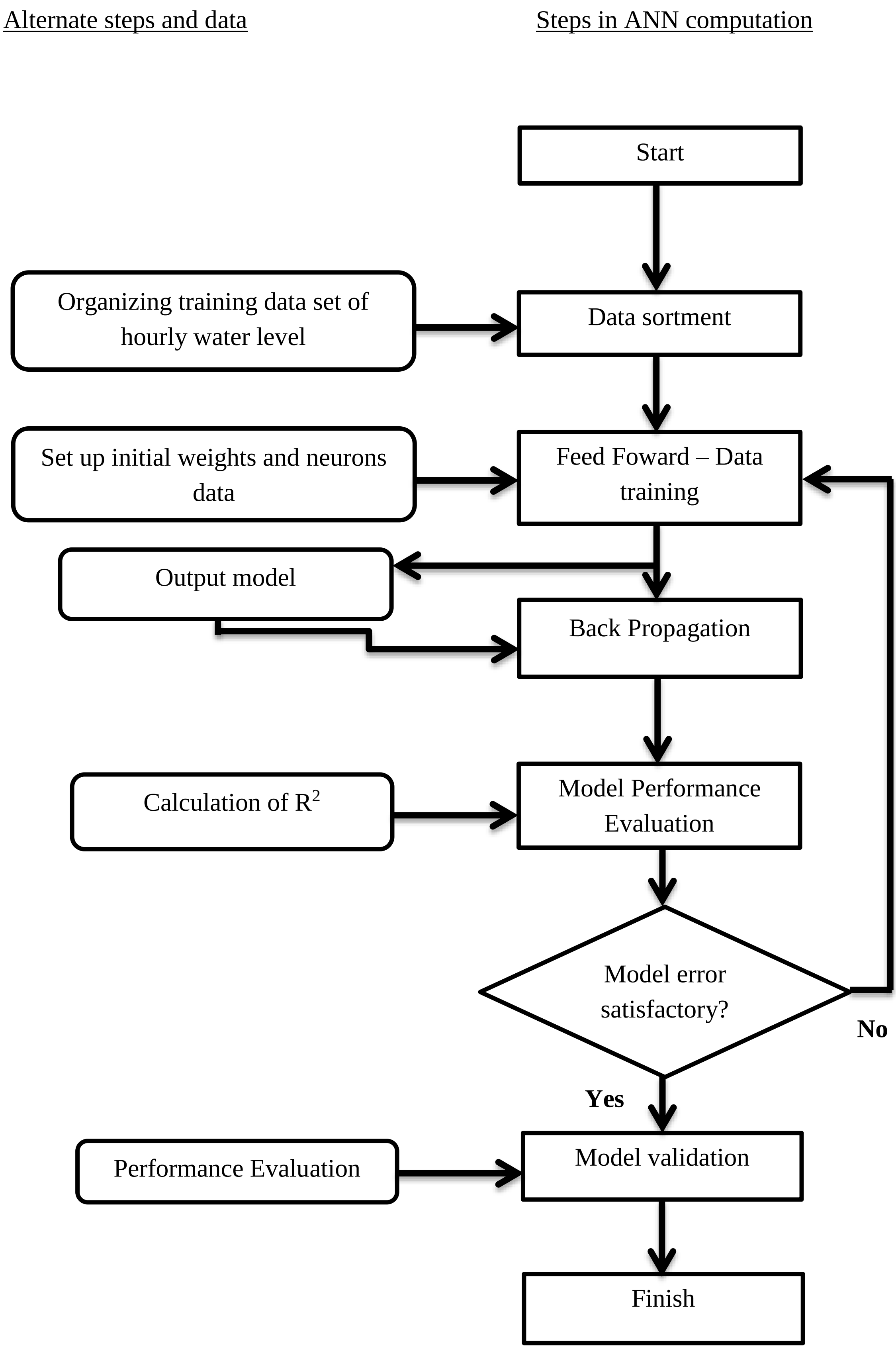
Neural network is superior to polynomial functions as neural network can work with many input variables and output variables as compared to 1 input and output used by the polynomial function. In addition, neural network can operate with plentiful class of functions with high efficiency.

3.1.2 Backpropagation Algorithm

The backpropagation algorithm is the most common used model of ANN, and is used in layered feed-foward network. In this network, the input is fed to the neuron in the input layer, and there will be one or more hidden network layer which includes the activation function neurons and the output is given by the neuron in the output layer.

The backpropagation algorithm will be provided with a set of inputs and outputs to be computed, and the errors, which is the difference between the computed and actual results will be calculated. The algorithm is used to train the network to reduce the errors, until it learns the training data where the actual and the calculated data is matched. The learning or training starts after the network is structured for the chosen application. There are two methods of training : supervised and unsupervised. The backpropagation algorithm uses supervised learning. During learning, the weights will start with a random value, and will be adjusted untill the error becomes the least possible.

Figure 7. Main steps in Back-Propagation Neural Network Modelling



3.2 Data Collection

Data that is to be used in this research is river water level which is meter. A period of minimum 14 years is going to be feeded into the neural network. 75% percent of the data is going to be used for training, and the remaining 25% will be used to evaluate the network model. Broken data wil be disregarded during sorting.

The data will be obtained from Jabatan Pengairan dan Saliran Malaysia, JPS. The data will be prepared and sorted out before feeded into the network by using custom programs that is to be written by using python programming language. Figure 6 shows the steps for data collecting and organizing.

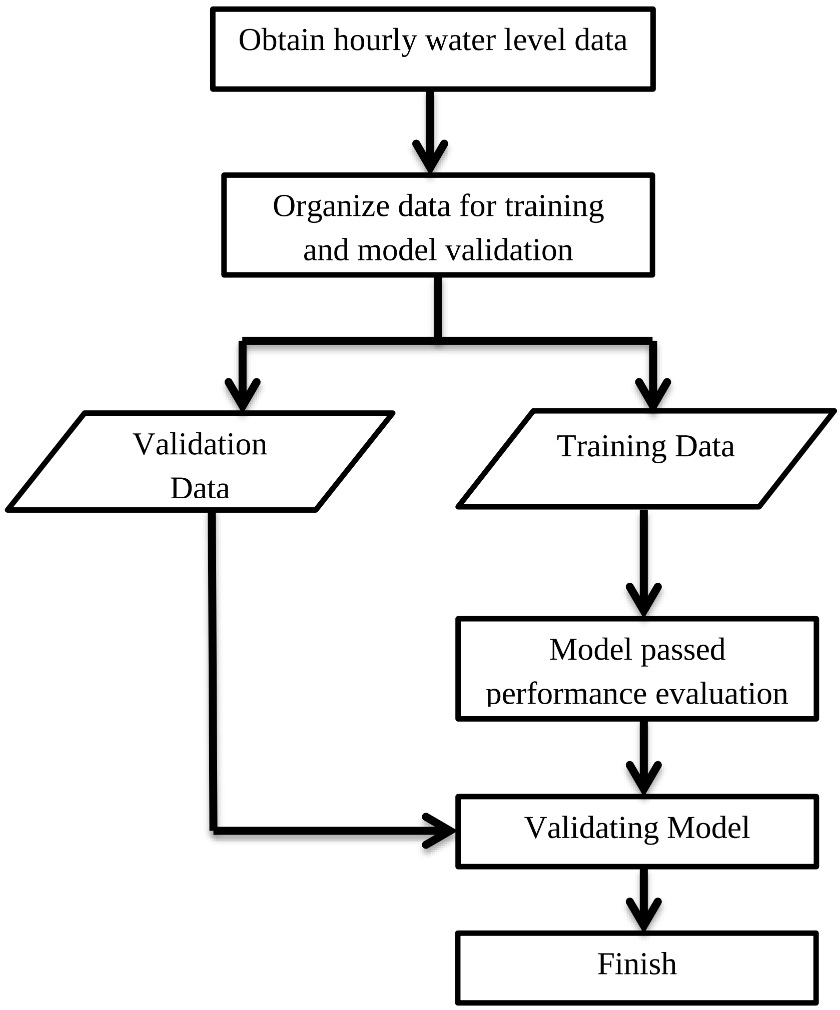


Figure 8. Steps in data collection and organization.

3.3 Data Processing

The data collected will be split into two parts;

75% to data training (DT)

25% to data validation(DV)

Both data used in training and validation are sorted to give 3 and 6 hours lead forecasting time. This will be done with custom programmes written Python programming language to ease the data processing flow.

After fed into ANN, the accuracy of the models and results produced by ANN will be evaluated using RMSE and Nash-Sutcliffe model efficiency coefficient to analyse the accuracy.

Models produced are named according to its respective architecture as in the table below:

3 hours lead time		Model name
No. of Iteration	No. of input	
1000	2	3-1000-2
	3	3-1000-3
	4	3-1000-4
	5	3-1000-5
	6	3-1000-6
4000	2	3-4000-2
	3	3-4000-3
	4	3-4000-4
	5	3-4000-5
	6	3-4000-6

TABLE 1.1 MODEL NAMES FOR 3 HOURS LEAD TIME DATA

6 hours lead time		Model name
No. of Iteration	No. of input	
1000	2	6-1000-2
	3	6-1000-3
	4	6-1000-4
	5	6-1000-5
	6	6-1000-6
4000	2	6-4000-2
	3	6-4000-3
	4	6-4000-4
	5	6-4000-5
	6	6-4000-6

TABLE 1.2 MODEL NAMES FOR 6 HOURS LEAD DATA

CHAPTER 4

RESULT AND DISCUSSION

4.1 Introduction

The data validation for 5 types of architecture, with 2 different of time interval and 2 numbers of iteration are presented and discussed in this chapter. Data validation consist of 25% of the 14 years of hourly water level data that is fed into ANN to produced a model.Comparison are done between architectures, time intervals and numbers of iteration to identify which combination will produce the most accurate model.

The models produced are named based on the architecture which indicates the data interval input, iteration done and the number of inputs fed into the ANN. For example, for model with interval of 3 hours, 1000 iterations and 2 inputs will be named as 3-1000-2 model.

4.2 Results

3 hours interval Data Validation								
No. of Iteration	No. of input	RMS	CE	Error under 500mm	Error between 500mm and 1000mm	Error between 1000mm and 1500mm	Error between 1500mm and 2000mm	Error more than 2000mm
1000	2	415.66	0.932	74%	26%	0%	0%	0%
	3	271.494	0.971	97%	3%	0%	0%	0%
	4	256.818	0.974	98%	2%	0%	0%	0%
	5	297.429	0.965	97%	3%	0%	0%	0%
	6	311.457	0.962	96%	4%	0%	0%	0%
4000	2	433.966	0.926	70%	30%	0%	0%	0%
	3	267.458	0.972	97%	3%	0%	0%	0%
	4	261.584	0.973	98%	2%	0%	0%	0%
	5	285.474	0.968	97%	3%	0%	0%	0%
	6	241.993	0.977	96%	4%	0%	0%	0%

TABLE 2. TABULATED RESULTS FOR 3 HOUR INTERVAL DATA INPUT.

Table 1 shows the tabulated data validation for models with 3 hour interval data input, with architecture of inputs ranging from 1 to 6 inputs that are fed into the ANN, with 1000 and 4000 iteration. The highlighted cells indicates the architecture that produces model with the highest accuracy with CE of 0.977 for 4000 iteration and 0.974 for 1000 iteration. Both of the model passed the requirement of CE > 0.95, with error under 500mm is 98% for 1000 iteration and 96% for 4000 iteration. The root mean square error gained is also low with the value of 241.993 mm. This shows that the result produced from the model does not differ very largely from the observed result.

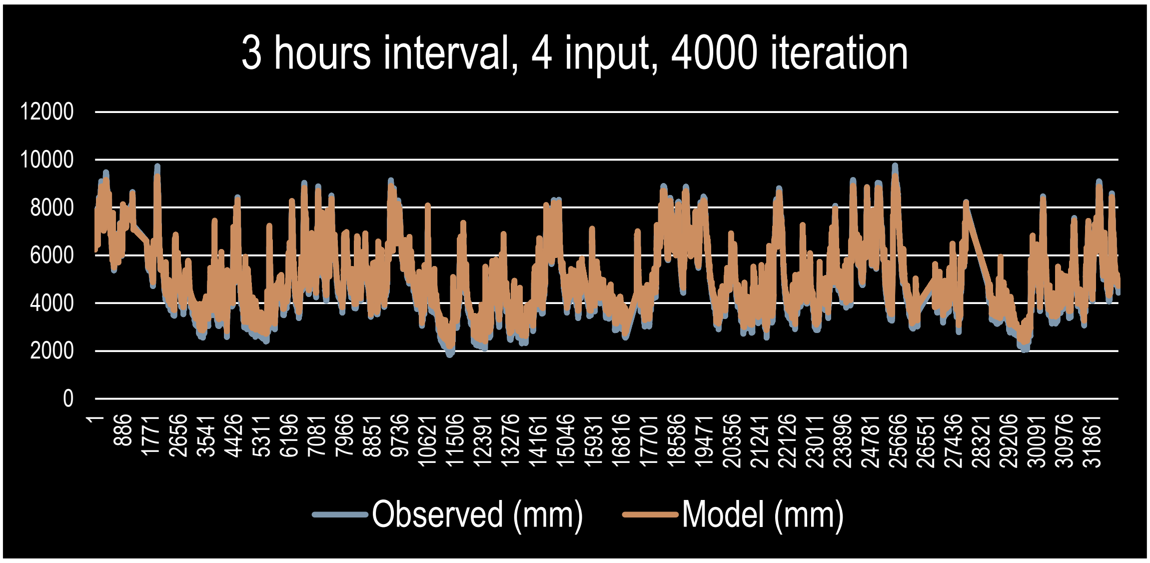
This accuracy allows the practical usage of both highlighted models to be used in the real world. The highest model among those two which is the 3-4000-6 model, can be used to predict real time water level to help in giving warning on alerting water level.

6 hours interval Data Validation								
No. of Iteration	No. of input	RMS	CE	Error under 500mm	Error between 500mm and 1000mm	Error between 1000mm and 1500mm	Error between 1500mm and 2000mm	Error more than 2000mm
1000	2	875.269	0.699	29%	40%	27%	4%	0%
	3	631.581	0.843	42%	51%	7%	1%	0%
	4	645.173	0.837	38%	55%	6%	0%	0%
	5	609.001	0.854	38%	55%	6%	0%	0%
	6	693.668	0.811	34%	56%	9%	1%	0%
4000	2	838.634	0.724	31%	43%	24%	3%	0%
	3	652.412	0.833	40%	51%	9%	0%	0%
	4	620.859	0.849	41%	54%	5%	0%	0%
	5	647.499	0.835	41%	51%	8%	0%	0%
	6	751.798	0.778	31%	53%	16%	0%	0%

TABLE 3. TABULATED RESULT FOR 6 HOUR DATA INPUT

Table 2 shows the tabulated data validation for 6 hours interval data input with the same architecture as 3 hour data input. However the model produced from feeding inputs with 6 hours interval did not produce model with accuraccy as high as 3 hour interval data input. The highest accuracy model is the 6-1000-5 model with CE of 0.854, not reaching to be more than 0.95. Also can be seen in this model that the distribution of error is large reaching up to the range between 1000 – 1500mm. Also the root mean square error gained is large with the lowest value of 609.001mm, which shows difference more than 500mm.

For practical uses this model cannot be used due to the low accuracy of the model produced.



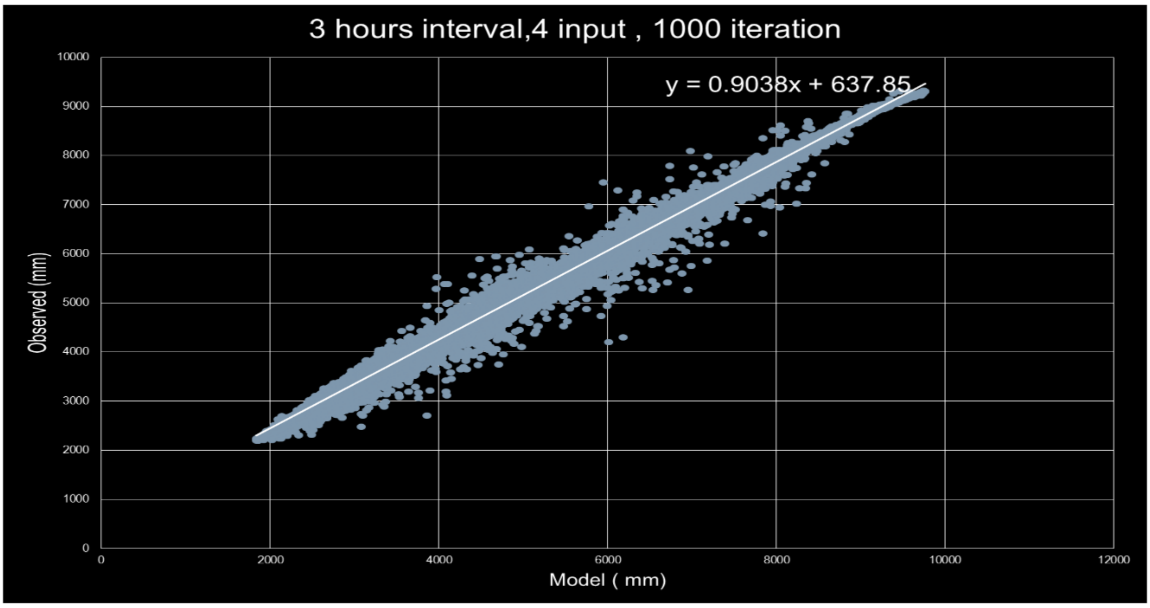


FIGURE 7 SHOWS THE ACCURACY PATTERN OF 3-1000-4 MODEL

Graph 1 shows the accuracy pattern of model 3-1000-4 along the trendline. It can be deduced that the result follows the trendline closely and does not scatter far away from the trendline. Also the pattern of result produced from the model follows closely to the observed data.

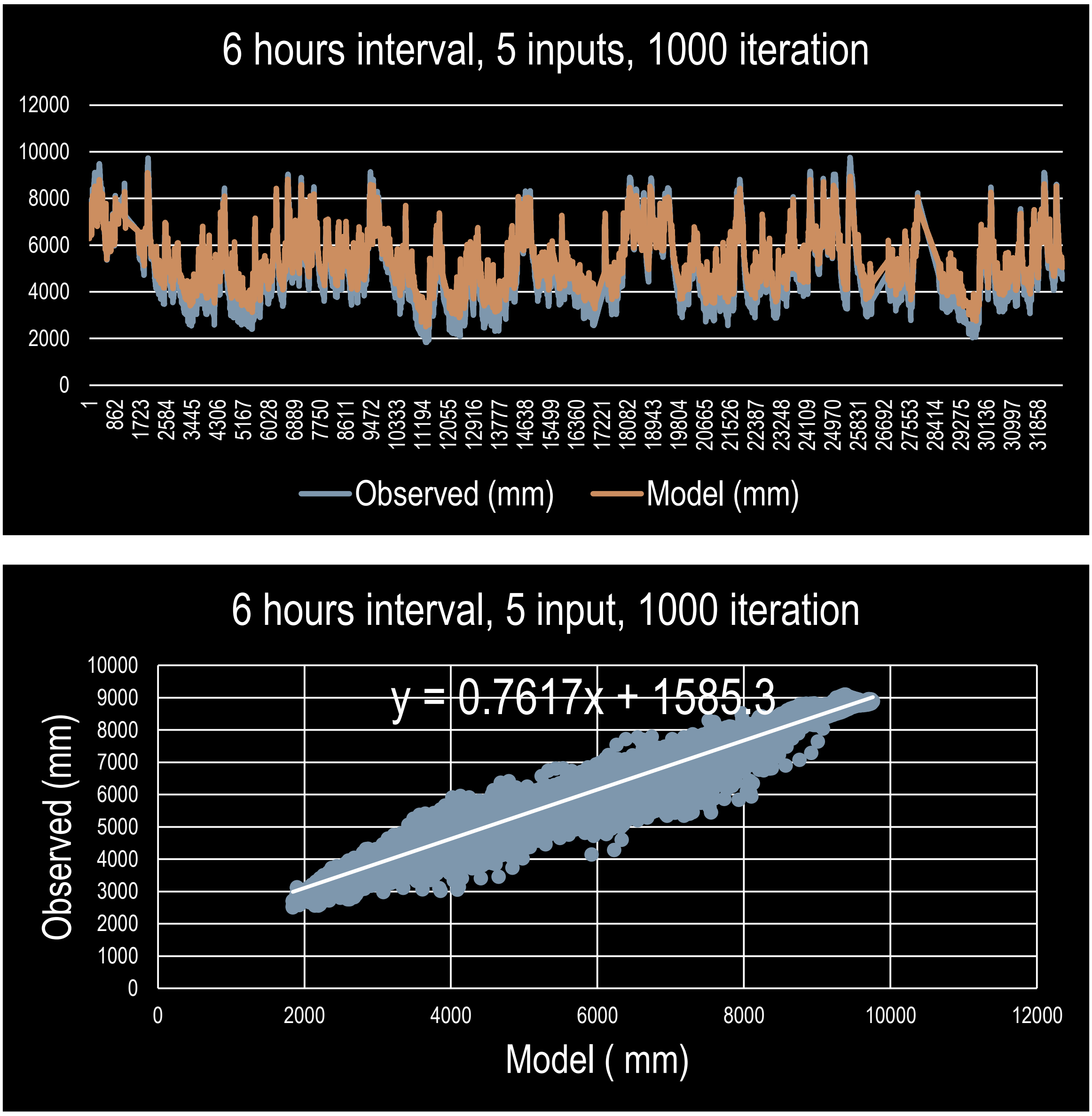


FIGURE 8 SHOWS THE ACCURACY PATTERN FOR MODEL 6-1000-5 MODEL

Graph 2 shows the accuracy of the model 6-1000-5 model. The result produced from this model scatter further away from the trendline when compared to the previous 3-1000-4 model that shows the shallow accuracy of the model. The result pattern produced from this model also deviates far from the observed data.

CHAPTER 5

CONCLUSION

This study has shown that ANN can be used to produce a model in forecasting water level to a high degree given the range of input and iteration and by using a single type of data. The most accurate model produced by the research consist of 5 number of inputs, with 3 hours interval lead time and 1000 iteration in ANN. This model that gives out the lowest RMS error which is 256.818 and the highest NSC coefficient efficiency of 0.974 is the 3-4000-6 model which can be practically used in real life to predict the water level, therefore helping in giving out warning to on incoming flood due to the rise of water level above danger level.

However, the model that uses 6 hours interval lead time in forecasting did not produce satisfactory result and are not practical to be used in real life forecasting. Thus this study shows that in Sungai Padas, a 3 hour lead time of forecasting can be done with satisfactorily high accuracy results in predicting its water level.

This study is hoped to contribute to the society by giving another fundamental and easier alternative of flood forecasting technique or adding up to existing forecasting technique to maximise the accuracy of forecasting so as to reduce the negative impact of flood on humans.

This study also shows that the architecture of ANN and the data structure to be fed into ANN can be expanded and be diversify in terms of architecture, data type and volume to increase its effectiveness.

RERFERENCES

Elias, Z., Hamin, Z., & Othman, M. B. (2013). Sustainable Management of Flood Risks in Malaysia: Some Lessons from the Legislation in England and Wales. *Procedia - Social and Behavioral Sciences*, *105*, 491–497. doi:10.1016/j.sbspro.2013.11.052

Farajzadeh, J., Fard, A., & Lotfi, S. (2014). Modeling of monthly rainfall and runoff of Urmia lake basin using “Feed-forward neural Network” and “Time series Analysis” model. *Water Resources and Industry*, *8*, 38–48. doi:10.1016/j.wri.2014.10.003

Gallien, T. W., Sanders, B. F., & Flick, R. E. (2014). Urban coastal flood prediction: Integrating wave overtopping, flood defenses and drainage. *Coastal Engineering*, *91*, 18–28. doi:10.1016/j.coastaleng.2014.04.007

Gershenson, C. (2003). Artificial neural networks for beginners. *arXiv Preprint cs/0308031*, 1–8. Retrieved from http://arxiv.org/abs/cs/0308031

Hussin, M. H. (2010). FLOOD ESTIMATION AND RIVER ANALYSIS. *Flood Estimation and River Analysis of Sungai Isap, Kuantan*, (November), 24.

Mediero, L., & Kjeldsen, T. R. (2014). Regional flood hydrology in a semi-arid catchment using a GLS regression model. *Journal of Hydrology*, *514*, 158–171. doi:10.1016/j.jhydrol.2014.04.007

Nasr, M. S., Moustafa, M. a. E., Seif, H. a. E., & El Kobrosy, G. (2012). Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. *Alexandria Engineering Journal*, *51*(1), 37–43. doi:10.1016/j.aej.2012.07.005

Rosenblatt, F. (1962). *Principles of neurodynamics;perceptrons and the theory of brain mechanisms.* Washington,. Retrieved from http://hdl.handle.net/2027/mdp.39015039846566