



UNIVERSITI
TEKNOLOGI
PETRONAS

FINAL YEAR PROJECT II

Final Report (Dissertation)

Numerical Modelling for Prediction of Suspended Sediment Concentration in Bidor River

September 2013

Prepared by:

Munira Malim Binti Abdul Kadir (15390)

Dissertation submitted in partial fulfilment of

the requirements for the

Bachelor of Engineering (Hons)

(Civil Engineering)

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan.

CERTIFICATION OF APPROVAL

Numerical Modelling for Prediction of Suspended Sediment

Concentration in Bidor River

by

Munira Malim Binti Abdul Kadir

A Dissertation submitted to the
Department of Civil Engineering
Universiti Teknologi PETRONAS

in partial ~~fulfillment~~fulfilment of the requirement for the
BACHELOR OF ENGINEERING (Hons)
(CIVIL ENGINEERING)

Approved by,

(-Dr.:-Muhammad Raza UI Mustafa)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

SEPTEMBER 2013

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work obtained herein have not been undertaken or done by unspecified sources or persons.

(Munira Malim Binti Abdul Kadir)

ABSTRACT

Nowadays, improvement of artificial intelligence, as an estimator for hydrological phenomenon has generated an abundant change in estimations. The estimation of sediment concentrations in rivers is highly important for designing and operation of various water resources since the life of water resources structure are directly concerned to the amount of sediments in rivers. In this study, Radial Basis Function Neural Network (RBFNN) model using Thin Plate Spline function was developed to estimate the suspended sediment concentration of a Bidor River in Perak. There are many studies had been through to predict the suspended sediment concentration in the hyper-concentration river by using soft computing techniques. Such as artificial neural network (ANN), support vector machines (SVM), gene expression programming (GEP), and Wavelet – ANN approach. Usually, ANN technique is a preferable to predict the suspended sediment concentration. Due to the lack of accuracy of the result, the different numerical modelling techniques will be used and at the same time to evaluate the performance of the models and recommend the best model among the developed ones. The data for this study are given by Department of Irrigation and Drainage (DID), where 4-years data (1992 – 1995) was applied for training and 1-year data (1996) was applied for testing. The parameters that will be use during the estimation are rainfall and discharge as the input data and suspended sediment as the output data. The root – mean square error (RMSE) and coefficient of determination (R^2) were expected to evaluate the model's performance. The best model performances will be used as estimator in the future studies on designing the hydrology structures.

ACKNOWLEDGEMENT

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

**With the name of Allah
The Most Gracious the Most Merciful**

First of all, my deepest gratitude to Allah SWT, the Lord of Universe for His blessings and guidance for being able to accomplish my Final Year Project entitled "Numerical Modelling for Prediction of Suspended Sediment Concentration in Bidor River". Special thanks and appreciation dedicated to my parents and family members for their support and encouragement towards me in completing this project.

First and foremost, I wish to thank my Supervisor, Dr Muhammad Raza Ul Mustafa for his valuable advice, continuous guidance, motivational support and positive criticisms that has helped me a lot since the start of this project until the completion of this project. He has driven me beyond my defined limits, which enabled me to discover my new strengths and capabilities. This Final Year Project has given me lots of useful knowledge and experiences.

Secondly, I would like to thanks to my research partner; Ahmad Safwan Bin Abu Bakar, who has been conducting the study together with me, and I would like to express appreciation and gratitude for helping along the whole duration of the study.

Finally, thank you to Civil Engineering Department, Examiners and Coordinators of the Final Year Project for making this program a success. They have been very helpful, tolerant and considerate throughout completing this project. Thank you to all parties involved for their contribution, guidance and support until the completion of this project.

TABLE OF CONTENT

<u>CERTIFICATION OF APPROVAL</u>	<u>i</u>
<u>CERTIFICATION OF ORIGINALITY</u>	<u>ii</u>
<u>ABSTRACT.....</u>	<u>iii</u>
<u>ACKNOWLEDGEMENT</u>	<u>iv</u>
<u>CHAPTER 1</u>	<u>1</u>
<u>PROJECT BACKGROUND.....</u>	<u>1</u>
<u>1.1 BACKGROUND OF STUDY</u>	<u>1</u>
<u>1.2 PROBLEM STATEMENT</u>	<u>2</u>
<u>1.3 OBJECTIVES</u>	<u>3</u>
<u>1.4 SCOPE OF STUDY.....</u>	<u>3</u>
<u>1.5 SIGNIFICANCE OF PROJECT</u>	<u>3</u>
<u>1.6 RELEVANCY OF PROJECT.....</u>	<u>43</u>
<u>1.7 FEASIBILITY OF THE PROJECT</u>	<u>4</u>
<u>CHAPTER 2</u>	<u>5</u>
<u>LITERATURE REVIEW / THEORY</u>	<u>5</u>
<u>2.1 SUSPENDED SEDIMENT CONCENTRATION (SSC).....</u>	<u>5</u>
<u>2.2 SOFT COMPUTING MODELLING.....</u>	<u>6</u>
<u>2.3 SUSPENDED SEDIMENT LOAD.....</u>	<u>9</u>
<u>2.4 SUPPORT VECTOR MACHINE (SVM)</u>	<u>9</u>
<u>2.5 CONCLUSION</u>	<u>10</u>
<u>CHAPTER 3</u>	<u>11</u>
<u>METHODOLOGY.....</u>	<u>11</u>
<u>3.1 RESEARCH METHODOLOGY</u>	<u>11</u>

Formatted: Line spacing: 1.5 lines

Formatted: Font: Not Bold

3.2	TOOLS / SOFTWARE	1312
3.3	STUDY AREA AND DATA GATHERING.....	13
3.4	RADIAL BASIS FUNCTION NEURAL NETWORK (RBFNN)	14
3.5	SUMMARY OF RBFNN MODEL DEVELOPMENT.....	21
3.6	PROJECT ACTIVITIES	22
3.7	GANTT – CHART / STUDY PLAN	23
3.8	KEY MILESTONE.....	25
CHAPTER 4		27
RESULT AND DISCUSSION		27
CHAPTER 5		3635
CONCLUSION AND RECOMMENDATION.....		3635
CHAPTER 6		3736
REFERENCES.....		3736
CERTIFICATION OF APPROVAL		i
CERTIFICATION OF ORIGINALITY.....		ii
ABSTRACT.....		iii
ACKNOWLEDGEMENT.....		iv
CHAPTER 1.....		1
PROJECT BACKGROUND		1
1.1	BACKGROUND OF STUDY	1
1.2	PROBLEM STATEMENT	2
1.3	OBJECTIVES	3
1.4	SCOPE OF STUDY	3
1.5	SIGNIFICANCE OF PROJECT	3
1.6	RELEVANCY OF PROJECT.....	3
1.7	FEASIBILITY OF THE PROJECT.....	4
CHAPTER 2		5

LITERATURE REVIEW / THEORY.....	5
2.1 — SUSPENDED SEDIMENT CONCENTRATION (SSC).....	5
2.2 — SOFT COMPUTING MODELLING.....	6
2.3 — SUSPENDED SEDIMENT LOAD.....	9
2.4 — SUPPORT VECTOR MACHINE (SVM).....	9
2.5 — CONCLUSION.....	10
CHAPTER 3.....	11
METHODOLOGY.....	11
3.1 — RESEARCH METHODOLOGY.....	11
3.2 — TOOLS / SOFTWARE.....	12
3.3 — STUDY AREA AND DATA GATHERING.....	13
3.4 — RADIAL BASIS FUNCTION NEURAL NETWORK (RBFNN).....	14
3.5 — PROJECT ACTIVITIES.....	20
3.6 — GANTT CHART / STUDY PLAN.....	21
3.7 — KEY MILESTONE.....	23
CHAPTER 4.....	25
RESULT AND DISCUSSION.....	25
CHAPTER 5.....	33
CONCLUSION AND RECOMMENDATION.....	33
CHAPTER 6.....	34
REFERENCES.....	34

LIST OF FIGURES

<u>Figure 1: Research Methodology</u>	<u>12</u>
<u>Figure 2: Location of Sungai Bidor</u>	<u>1413</u>
<u>Figure 3: Schematic representation of radial basis function neural network</u>	<u>1514</u>
<u>Figure 4: Time series of suspended sediment concentration and stream discharge of Bidor River.....</u>	<u>18</u>
<u>Figure 5: Time series of comparison between daily suspended sediment concentration and daily stream discharge for training data</u>	<u>19</u>
<u>Figure 6: Time Series of comparison between daily suspended sediment concentration and daily stream discharge of testing data</u>	<u>2019</u>
<u>Figure 7: Summary of RBFNN model development</u>	<u>21</u>
<u>Figure 8: Project Activities</u>	<u>22</u>
<u>Figure 9: Key Milestone (FYP 1).....</u>	<u>25</u>
<u>Figure 10: Key Milestone (FYP 2).....</u>	<u>26</u>
<u>Figure 11: Comparison of observed and prediction SSC for Training Data.....</u>	<u>30</u>
<u>Figure 12: Comparison of observed and prediction SSC for Testing Data.....</u>	<u>31</u>
<u>Figure 13: Scatter plot of prediction versus observed Suspended Sediment Concentration for Training Data</u>	<u>32</u>
<u>Figure 14: Scatter plot of prediction and observed suspended sediment concentration for testing data.....</u>	<u>33</u>
<u>Figure 1: Research Methodology</u>	<u>12</u>
<u>Figure 2: Location of Sungai Bidor.....</u>	<u>13</u>
<u>Figure 3: Schematic representation of radial basis function neural network.....</u>	<u>14</u>
<u>Figure 4: Time-series-of suspended sediment concentration and stream discharge-of Bidor River</u>	<u>17</u>
<u>Figure 5: Time-series-of comparison between daily suspended sediment concentration and daily stream discharge for training data</u>	<u>18</u>
<u>Figure 6: Time Series-of comparison between daily suspended sediment concentration and daily stream discharge-of testing data</u>	<u>18</u>
<u>Figure 7: Project Activities</u>	<u>20</u>
<u>Figure 8: Key Milestone (FYP 1).....</u>	<u>23</u>

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Line spacing: 1.5 lines

Figure 9: Key Milestone (FYP 2).....	24
Figure 10: Comparison of observed and prediction SSC for Training Data	28
Figure 11: Comparison of observed and prediction SSC for Testing Data	29
Figure 12: Scatter plot of prediction versus observed Suspended Sediment Concentration for Training Data.....	30
Figure 13: Scatter plot of prediction and observed suspended sediment concentration for testing data.....	31

LIST OF TABLES

<u>Table 1: Tools/Software</u>	<u>1312</u>
<u>Table 2: Statistical Analysis on Stream Discharge and Suspended Sediment Load..</u>	<u>16</u>
<u>Table 3: Study Plan for FYP 1</u>	<u>23</u>
<u>Table 4: Study Plan for FYP 2</u>	<u>24</u>
<u>Table 5: Value of Hidden Neurons, Training and Testing Error using trial and error method.....</u>	<u>28</u>
<u>Table 6: Performance statistic of RBFNN modelling during training and testing</u>	<u>29</u>
<u>Table 7: The values of R² from different models.....</u>	<u>3534</u>
<u>Table 1: Tools/Software</u>	<u>12</u>
<u>Table 2: Statistical Analysis on Stream Discharge and Suspended Sediment Load.....</u>	<u>16</u>
<u>Table 3: Study Plan for FYP 1</u>	<u>21</u>
<u>Table 4: Study Plan for FYP 2</u>	<u>22</u>
<u>Table 5: Value of Hidden Neurons, Training and Testing Error using trial and error method</u>	<u>26</u>
<u>Table 6: Performance statistic of RBFNN modelling during training and testing</u>	<u>27</u>
<u>Table 7: The values of R2 from different models</u>	<u>32</u>

- Formatted:** Line spacing: 1.5 lines
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Field Code Changed**
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Font: (Default) Times New Roman, 12 pt
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt, Check spelling and grammar
- Formatted:** Font: (Default) Times New Roman, 12 pt, Not Bold
- Formatted:** Table of Figures, Line spacing: 1.5 lines, Tab stops: 14.48 cm, Right,Leader: ...

CHAPTER 1

PROJECT BACKGROUND

1.1 BACKGROUND OF STUDY

Suspended sediments concentration (SSC) is the ratio of the mass of dry sediment in water – sediment mixture to the mass of the water – sediment mixture. It is typically definite in milligrams of dry sediment per litre of water – sediment mixture. SSC is an important factor that affects stream ecology and morphology. Prediction of suspended sediments in hyper concentration river has a vibrant role in distributing with water resources problems and hydraulic structures such as dams and reservoirs. Suspended sediments are usually transported within and at the same velocity as the surrounding fluid (water or wind). The greater amount of sediment that can be suspended by turbulence are depends on the stronger the flow and/or the finer the sediments. Only the finer fraction (usually silt and clay fraction) of the suspended sediments can be continuously maintained in suspension by the flow turbulence. This fraction is often referred to as "wash load," and is typically not found in significant quantities at the bed surface. This concentration is usually connected to the sediment supply and is difficult to determine theoretically i.e. the higher the discharge the higher the suspended sediment concentration.

Water quality is a worldwide apprehension, and contamination resulting from human activities or natural events affects the lives of human being. In many rivers, these pollution is caused by sediments that be made up of silt and clay materials. Sedimentation in rivers is an important problem to be taken into account in water resources planning. A part of the reservoir volume (dead volume) is gradually filled by the sediment load. The total sediment load contains bed load and suspended materials, and a large part of the sediment load is transported in suspension as wash load. Therefore, the concentration of the suspended sediment particles is an important part of the total amount of suspended sediment (Imen, 2008).

The lifespan of dams and reservoirs can be shortening because of the sediment in the rivers. The sediments used to flow along with the relatively fast – moving river, instead, dropped in the reservoir, when a river is dammed and a reservoir is created.

1.2 PROBLEM STATEMENT

The manual estimator may cause time constrained which is took a long time to predict the SSC. Radial Basis Function Neural Network (RBFNN) will be used to estimate the suspended sediment concentration river to ensure that the prediction is more accurate rather than other modelling from previous studies.

Large amount of sediment load carried out by river flow either in suspension or as bed load. Therefore, the estimation of sediment concentrations in rivers is highly important for designing and operation of various water resources since the life of water resources structure (i.e. hydropower and environmental engineering project such as reservoir, dams, sanitation, water pollution, etc.) are directly concerned to the amount of sediments in rivers.

Prediction of sediment concentrations empower to achieve appropriate designing of hydraulic structures and preventing them from damages. Previously, the numbers of studies have been conducted for modelling suspended sediments concentrated using Artificial Neural Network (ANN) and other computing technique such as Machine-code Linear Genetic Programming (LGP) where the result shown is better than Gene – Expression Programme (GEP) and ANN (Guyen & Kişi, 2011). In year 2011, Rajae (2011) had combined the two (2) methods to estimate the suspended sediments which are Wavelet and ANN with the results show that by combination two techniques gave an accurate result.

At present, used of SVMs is being well thought-out in different disciplines to auxiliary develop upon the performance of ANN models as a potential alternative (Goyal and Ojha. 2011). Lafdani, E. K., et al (2012) has come out with new computing technique to estimate the suspended sediment that is the combination of SVM and ANN models were predictable using the combination of Gamma Test and Genetic Algorithms (GT-GA). In this study, the combination of Radial Basis Function (RBF) with Support Vector Machine (SVM) will be used to estimate the suspended sediments in hyper – concentration river.

1.3 OBJECTIVES

The main objective of this study is to predict suspended sediment concentration in Bidor River with the following specific objectives:

- i. To develop numerical model for prediction of suspended sediment in a concentrated river using Radial Basis Function Neural Network.
- ~~i. To evaluate the performance of the model using several statistical measures. To develop numerical models for prediction of suspended sediment in a hyper concentrated river using different numerical modelling techniques.~~
- ii. ~~To evaluate the performance of the models and propose the best model among the developed ones.~~

Formatted: List Paragraph, Numbered + Level: 1 + Numbering Style: i, ii, iii, ... + Start at: 1 + Alignment: Right + Aligned at: 0.63 cm + Indent at: 1.27 cm

1.4 SCOPE OF STUDY

Prediction of suspended sediment in hyper-concentrated will be performed using discharge, rainfall and sediment data of Bidor river. Therefore, the scope of the study can be defined as:

- ~~i-iii.~~ The study is limited to develop a Radial Basis Function Neural Network models using Thin Plate Spline Function for prediction of suspended sediment using data from Bidor River.
- ~~ii-iv.~~ To evaluate the performance of the prediction model using two statistic measures that is determination coefficient and Root Mean Square Error (R^2 , RMSE).

1.5 SIGNIFICANCE OF PROJECT

Estimation of suspended sediment is important for designing and operation of various water resources, hydropower and environmental engineering project. This study will provide an unconventional numerical modelling technique to predict the suspended sediment in the river. Moreover, future investigations and methods can be carried out to ensure the suspended sediment load degrades to protect the hydro-structure.

1.6 RELEVANCY OF PROJECT

This research is important as it will play a major role in designing the lifespan of hydraulic structure. Prediction of suspended sediment load is considerably vital as it will help to select the appropriate design of Hydraulic structures capacity.

1.7 FEASIBILITY OF THE PROJECT

The required data will be acquired from the Department of Irrigation and Drainage (DID), Ministry of Natural Resources and Environment, Kuala Lumpur, Malaysia which is a convenient move as no fieldwork is required hence reducing the time spent for data collection which allowing more allocation of time for the purpose of data analysis and construction of prediction model.

CHAPTER 2

LITERATURE REVIEW / THEORY

This chapter reviews upon previous published literatures that are related to the research. Other books and papers have been read and were used to assist in this research paper. There are several topics which various papers discuss on the abilities of numerical modelling on estimation the suspended sediments concentration in several locations of rivers.

2.1 SUSPENDED SEDIMENT CONCENTRATION (SSC)

Suspended sediments are usually transported within and at the same velocity as the surrounding fluid (water or wind). The greater amount of sediment that can be suspended by turbulence are depends on the stronger the flow and/or the finer the sediments. Only the finer fraction (usually silt and clay fraction) of the suspended sediments can be continuously maintained in suspension by the flow turbulence. This fraction is often referred to as "wash load," and is typically not found in significant quantities at the bed surface. Its concentration is usually related to the sediment supply and is difficult to determine theoretically i.e. the higher the discharge the greater the SSC.

The suspended sediment concentration can be estimated by calculate using this equation:

$$SSC = Q_{observed} \times C_{estimated} \times 0.0864$$

This process is carried out daily, and the daily loads are summed for the year.

2.2 SOFT COMPUTING MODELLING

Lafdani et al. (2013) investigated the capabilities of support vector machines (SVM) and artificial neural networks (ANN) models to calculate daily suspended sediment load (SSL) in Doiray River at west of Iran by collected 11 data from year 1994 – 2004. There is varied lag time series of Stream flow and Rainfall as input data and Discharge (Qst) as output data in this study. The best input was identified by amalgamation of Gamma Test and Genetic Algorithm (GT-GA). By using this combination to find the best accuracy and reliability of this result and will be compared to the result that achieved as of the traditional correlation coefficient analysis. Radial Basis Function (RBF) is one of four kernels (nu-SVR) models that shown as best kernel for modelling in SVM compared to the other three kernels. The consistency of SVM and ANN assessed based on performance criteria such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Efficiency Index (EI) and correlation coefficient (R^2). The less statistic value, the better the function of the model.

Formatted: Font: (Asian) Chinese (PRC), Superscript

Different model of soft computing technique that also been used toward prediction of daily suspended sediment concentration in hyper-concentration river. Wavelet-ANN is combination of time-dependent spectral analysis viz. time-based illustration of progressions with the adjustable to modelling and expecting in nonlinear behaviour. Liu et al (2013) studied the connection among suspended sediment concentration (SSC) and river discharge (Q) by the wavelet- ANN (WANN) at Kuye River in middle of Yellow River catchments of China. Based on the result, the WANN model showed higher prediction accuracy than the sediment rating curve (SRC) model or the ANN model. The WANN model showed more sturdy performance than the SRC and ANN models, specified by the suitable values of error autocorrelation and input-error correlation (Liu et al, 2013). WANN model is better predicts SSC in a hyper – concentrated stream setting, with extremely nonlinear and non – stationary time series. WANN model was developed with low correlation of 0.601 among the discharge and sediment concentration outstretched the predictive performance (R^2) by 57.5% and decreased RMSE by 46.2% compared with the SRC model.

Formatted: Font: (Asian) Chinese (PRC), Superscript

According to Demirci, M. and Baltaci, A. (2012), ~~using~~ Fuzzy Logic (FL) can be used to approached to predict the suspended sediments concentration from stream

flow data of Sacramento Freeport Station operated by United States Geological Survey. The result is compared together using Multi-linear Regression (MLR) and Sediment rating curve (SRC). SRC had widely used to estimate SSC but can't provide sufficiently accurate results. Fuzzy rule-based model could significantly improve the magnitude of prediction accuracy of suspended sediment in hyper-concentration River. For SRC model, 1215 data used for training and 611 data are divided for testing, while MLR model evaluated 5-year data and same with FL model. The results are used to compare the performance by using the MSE, MAE and R parameter obtained from testing data. The FL model gave the best accuracy in total sediment load estimation rather than MLR and the worst estimation result were obtained in SRC model.

Azamathulla et al. (2012) state four basic operators in gene expression programming (GEP) as alternative approach to modelling the suspended sediment load of river systems. They also used adaptive neuro-fuzzy inference system (ANFIS), regression model together with GEP to predict the suspended sediment. The data provided by REDAC with total 214 sets of data from three different rivers in Malaysia were used in this study. The result from different model will be compared and obtained the best performance by using RMSE, R^2 and average error (AE). Even ANFIS predicting accurate value of SSC but GEP model is suggested for preliminary prediction due to the complexity of ANFIS model since the traditional formulas fail to predict the suspended sediment load accurately. ANFIS has been suggested to be used to predict the SSC in the future.

According to Mustafa, M. R. et al (2012), Multilayer Perceptron (MLP) also can be used to calculate the suspended sediment which is one of the artificial neural network (ANN) training. There are four different training algorithms used in MLP training; Gradient Descent (GD), Gradient Descent with Momentum Descent (GDM), Scaled Conjugate Gradient (SCG) and Levenberg Marquardt (LM). Neurons / nodes are a data processing mathematical model that be made up of a number of units or elements in ANN. These time series of river suspended sediments with four algorithms was competent to trait the performance of the ANN models. The GD and GDM provided some destructive values throughout training with the default learning rate, $\mu = 0.01$. But SCG and LM algorithms were able to reduce significantly the number of negative values. SCG and LM algorithms learned the non-linear pattern

Formatted: Font: (Asian) Chinese (PRC), Superscript

very fit paralleled to the GD and GDM during training in expecting sediments loads at low tides.

As reference in Kisi, et al (2012) is an important factor to model suspended sediment in water resources engineering as it highly change the design and management of the structures. The genetic programming (GP) i.e. GEP has been used and compared the results with ANFIS, ANN and SVM where the data of Cumberland River in U.S. were obtained as of the USGS website designed for period of 10 years. GEP has its own advantage which is why this technique has been used to predict the suspended sediment rather than other techniques that can produce mathematical formula of the relationship between the input parameters, and it can be derived from the expression trees. However, it is hard to interpret the program as large programs are computationally expensive to develop where it will form a nested function such as expand phenomena based on the observation of the program sizes of parse trees started growing. The parameter values are based on the default value of GeneXpro programme.

Besides that, there is another study on GEP that proved the performance on predicting the suspended sediment is improved than ANN and ANFIS. In this study, ANN will used Lavenberg – Marquardt, conjugate gradient and gradient descent training algorithms and result shows that among these three algorithms, Conjugate gradient is better that the others. Kisi and Shiri (2012) used the hydro – meteorological data for Eel River at California, USA as the case study on calculate approximately of daily suspended sediment concentration in rivers. There are three variables that will be used as parameter for training and testing period, i.e. discharge, precipitation and sediment that substantial influence on the estimated SSC with the cross-correlation method was used.

2.3 SUSPENDED SEDIMENT LOAD

All types of sediment such as rock, dirt and clays are carried by each river wherever it goes. Suspended sediment is known as when a river flows over any of these materials, it may pick them up and carry them to the downstream (Zang, n.d.). Sediment carried along by the flow of a river is known as sediment load. The capacity of suspended sediment in the river and its watershed is including geology and ecology, as well as the impact of human activities like development and pesticide use. Suspended sediment can be calculated by sediment concentration multiplied by stream discharge equals to sediment load.

2.4 SUPPORT VECTOR MACHINE (SVM)

SVM is related to the new learning algorithm and has been introduced originally to solve pattern recognition and classification problems and regression techniques that have been consequential from numerical learning reproductions (Vapnik & Lerner, 1963). Therefore, GEP has an advantage on predicting the daily suspended sediment and show almost accurate result rather than other technique, but it still computationally expensive to develop. As a result, SVM has empirically good performance which by where there are successful applications in many field before. According to Vapnik (1995), the principals of numerical learning concept is formulated in the direction of SVM is engaged ~~SVM is engaged~~ in grouping and regression algorithms series.

Formatted: Font: (Asian) Chinese (PRC), (Other) English (U.K.)

2.5 CONCLUSION

This chapter includes all the past researchers that have been conducted by other researchers in this related study. There are several topics that were identified from the past researchers like accuracy, types and abilities of numerical modelling. Furthermore, it also refers to other published materials that have contributed to this same area of study. The review of some existing literature has revealed that several successful researches on suspended sediment estimation in hyper-concentrated rivers using numerical soft computing techniques. Such as artificial neural network (ANN), support vector machines (SVM), gene expression programming (GEP), and Wavelet – ANN approach. Usually, ANN technique is a preferable to predict the suspended sediment concentration. Hence, there are many studies that have been done using ANN technique. Nowadays, they try to conduct new computing techniques to predict the suspended sediment concentration and come out with the best accuracy in prediction on the suspended sediment. For instance, a research done by Lafdani et al. in year 2013 has investigated the abilities of SVM and ANN imitations to calculate daily suspended sediment load (SSL) in Doiray River at west of Iran by collected 11 data from year 1994 – 2004 where the result shows that the abilities of SVM on prediction is more accurate rather than ANN. Further, Liu et al. studied the connection between suspended sediment concentration (SSC) and river discharge (Q) using the wavelet-ANN (WANN) at Kuye River in middle of Yellow River catchments of China.

CHAPTER 3

METHODOLOGY

In this part of research methodology, the activities can be divided into two parts. Basically, the first part is on the training part where the data will be analysing to get the input data. For the second part, will be the testing data where the data will be tested on the MATLAB to find the hidden neuron to compute the output data. These data will be used to verify that this model can be used as a prediction of suspended sediment even though the data might not fully given by the [Department of Irrigation and Drainage \(DID\)](#).

3.1 RESEARCH METHODOLOGY

The research methodology is present as followed:

a. Literature search and review

The research consists of journal articles, books and website articles. Relevant information extracted from these sources is compiled for citation and cross – referencing and this becomes the basis for the further analysis.

b. Data Gathering

This project aims to gather all information regarding the material that may be used during the estimation of the suspended sediment based on the data given by the Department of Irrigation and Drainage (DID) and gather it by using data spread from Microsoft Excel.

c. Data Analysis

By assessing the data that have been gathered, the suitable material will be choosing for the estimation of the suspended sediment at the hyper-concentration river. At the end of the project, we offer some recommendations that might help reduce, rectify or even prevent the effects of the problem's root causes.

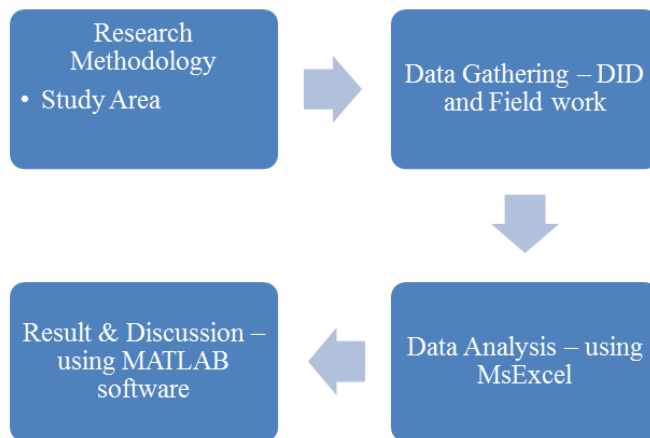


Figure 1: Research Methodology

~~which is formulated using the principals of statistical learning theory.~~

Formatted: Indent: Left: 1.27 cm

TOOLS / SOFTWARE	DESCRIPTION
Microsoft Excel	A spread sheet application that performs graphing tools and pivot tables for data analysis.
MATLAB 8.0	A high – level language and interactive environment for numerical computation, visualization and programming. It's used to analyse data, develop algorithms and create models and

Formatted Table

Formatted: Font:

Formatted: Left, Position: Horizontal: 0.19 cm, Relative to: Margin, Vertical: 2.81 cm, Relative to: Paragraph

Formatted: Font: 12 pt, Font color: Text 1

Formatted: Centered, Position: Horizontal: 0.19 cm, Relative to: Margin, Vertical: 2.81 cm, Relative to: Paragraph

Formatted: Font: 12 pt, Bold, Font color: Text 1

Formatted: Font: Font color: Text 1

Formatted: Font: English (U.K.)

Formatted: Font: 12 pt, Font color: Text 1, English (U.K.)

Formatted: Centered, Space Before: 10 pt, After: 0 pt, Line spacing: 1.5 lines, Position: Horizontal: 0.19 cm, Relative to: Margin, Vertical: 2.81 cm, Relative to: Paragraph

Formatted: Font: 12 pt, Bold, Font color: Text 1, English (U.K.)

applications. Faster solution to explore multiple approaches in with language tools and built – in functions rather than spread sheets or traditional programming language.

Formatted: Font: Font color: Text 1, English (U.K.)

3.2 TOOLS / SOFTWARE

Table 1: Tools/Software

3.3 STUDY AREA AND DATA GATHERING

The research data are obtained from the Department of Irrigation and Drainage (DID) for the river of Sungai Bidor. Sungai Bidor is located in the state of Perak. Bidor is located in the south of Tapah and the north of Sungkai along the North-South Expressway that leads to Padang Besar and Johor Bahru (**Figure 2**). With the prediction of the population that may increase from 32,094 people to 44,564 people in 2020 through main commercial activities in such as plantation, agricultural and business. Due to these situations, producing of the suspended sediment at this area may cause the increasing of the capacity of the river. The data shows in **Table 1**. These data has been analyzed using the Radial Basis Function Neural Network (RBFNN) and will be compared with SVM to get the most accurate performance.

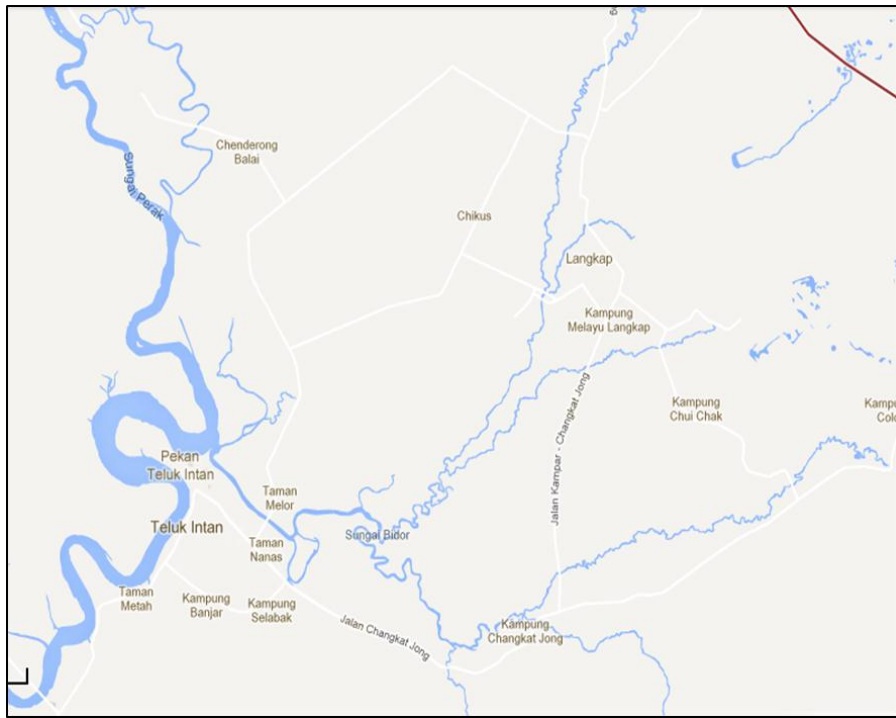


Figure 2: Location of Sungai Bidor

3.4 RADIAL BASIS FUNCTION NEURAL NETWORK (RBFNN)

RBFNN has been developed complete by selecting and partitioning of input data, data normalization, and selection of number of input and hidden neurons. This model has been used on predicting of pore water pressure (Mustafa at el., 2012). Basically, the general structure of RBFNN based on 3 main layers which are; input layer, hidden layer (hidden neuron) and output layers. The data need to be selected properly in order to develop the RBFNN and at the same time partitioned and normalization the data to reduce the complexity of the learning process of RBFNN model to provide the prediction with high accuracy. The architecture of RBFNN model is presented in **Figure 3**.

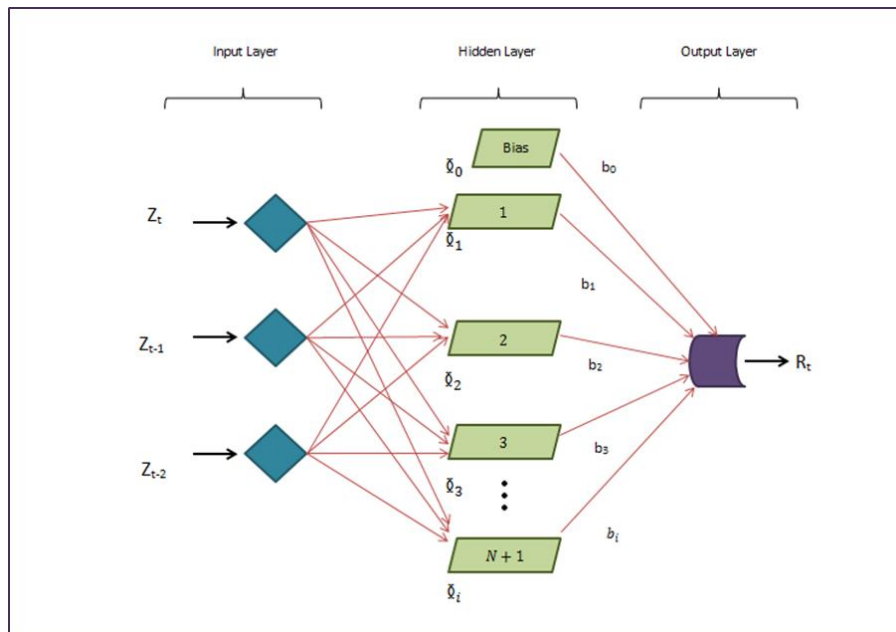


Figure 3: Schematic representation of radial basis function neural network

The three input variables, Z_t , $Z_{(t-1)}$ and $Z_{(t-2)}$, are linked to each of the neurons in hidden layer and as a result, there is only produced one output which is the predicted suspended sediment concentration at present, R_t . The results obtained from the output were analyzed and presented in the form of table.

a. Data Selection

Average of discharge, Q and different interval of time series of rainfall are used as modal input data in estimation of suspended sediment concentration by using the soft computing and conventional methods. The data for this river are taken from the year 1992, 1993, 1994, 1995 and latest by 1996. These years were chosen because of they have the most completed data available and the most recent. The 4-year's earlier data will be used for training, while data at year 1996 will be used as testing data. The data will be gathered as per

shows in **Table 2**. These completed data needed to form a pattern to be projecting by RBFNN model, which use to create accuracy of estimation. Missing data will result to high skewness and scattered data, which as a result increasing the difficulty of the processes of RBFNN model.

While for the validation of the modelling, the set of data for year 2006 will be used where in this year DID didn't provided the complete data. From here, the RBFNN model will be tested to ensure that this model can predicted the suspended sediment even the given data of discharge is not completed.

b. Model input selection

To envisage daily SSC, different interval of time series of rainfall and discharge as input data and suspended sediment load as output data has been taken. The best input combination is the most important step in any modelling. By selected the best input data, hidden neuron can be found by using the trial and error method.

Table 2: Statistical Analysis on Stream Discharge and Suspended Sediment Load

Statistics	Stream Discharge	Suspended
	(m ² /day)	Sediment Load (tons/day)
No. of Data	1409	1409
Mean	21.65	98.23
Max	101.04	453.90
Min	3.47	6.00
SD	13.54	60.73

Training

- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, Bold, English (U.K.)
- Formatted: Font: 14 pt, Bold, English (U.K.)
- Formatted: Font: 14 pt, Bold, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted Table
- Formatted: Font: 14 pt, Bold, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)
- Formatted: Font: 14 pt, English (U.K.)

	Variance	183.27	3688.72	Formatted: Font: 14 pt, English (U.K.)
	Skew	1.83	1.78	Formatted: Font: 14 pt, English (U.K.)
	Total	30501.43	138407.40	Formatted: Font: 14 pt, English (U.K.)
	Statistics	Stream Discharge (m²/day)	Suspended Sediment Load (tons/day)	Formatted: Font: 14 pt, English (U.K.)
Testing	No. of Data	357	357	Formatted: Font: 14 pt, English (U.K.)
	Mean	15.03	68.05	Formatted: Font: 14 pt, English (U.K.)
	Max	53.26	239.50	Formatted: Font: 14 pt, English (U.K.)
	Min	6.29	19.00	Formatted: Font: 14 pt, Bold, English (U.K.)
	SD	7.49	35.03	Formatted: Font: 14 pt, English (U.K.)
	Variance	56.10	1226.80	Formatted: Font: 14 pt, English (U.K.)
	Skew	1.75	1.38	Formatted: Font: 14 pt, English (U.K.)
	Total	5364.26	24294.10	Formatted: Font: 14 pt, English (U.K.)

Table 2 shows that the statistical analysis on Stream Discharge and Suspended Sediment Load where the total data is 1409 and 357 used for training and testing data to evaluate the models performances. **Figure 5** and **6** shows the time series of comparison in relation to daily suspended sediment concentration and daily stream discharge for training and testing data.

c. Partitioning Data

Partitioning data for training and testing will be based on the data trend. Based on **Figure 4** below, there are slightly gaps between years 1992-1993

and gaps within year 1996. From that the data that will be used for training and testing has been decided.

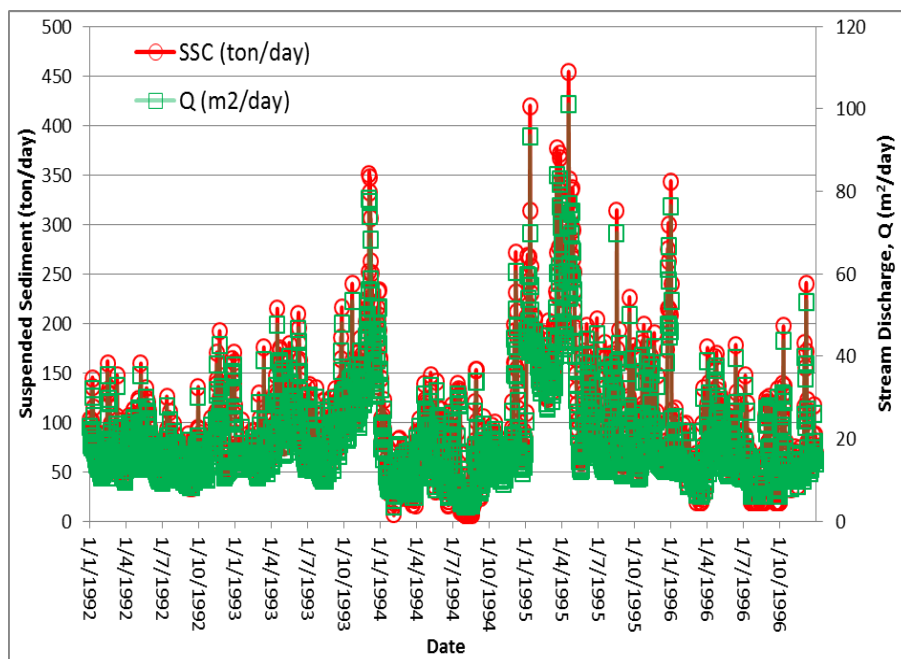


Figure 4: Time series of suspended sediment concentration and stream discharge of Bidor River

Total available data is 1766 data and 1409 data will be used for the drive of training and the remaining 357 will be used for testing purpose, for both stream discharge and suspended sediment concentration. The time series of comparison between daily suspended sediment concentration and daily stream discharge for training and testing data is shown in **Figure 5** and **6**.

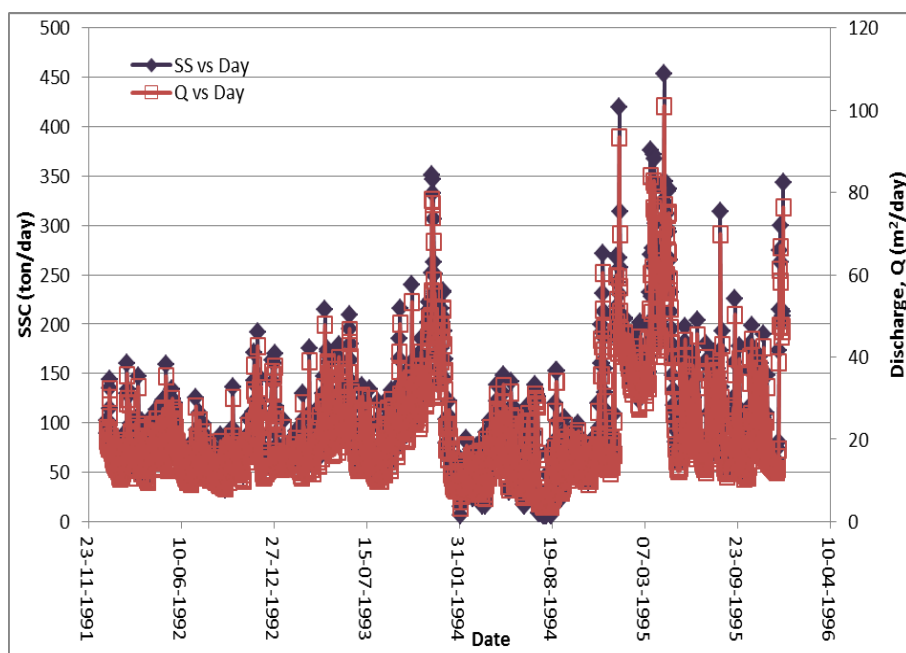


Figure 5: Time series of comparison between daily suspended sediment concentration and daily stream discharge for training data

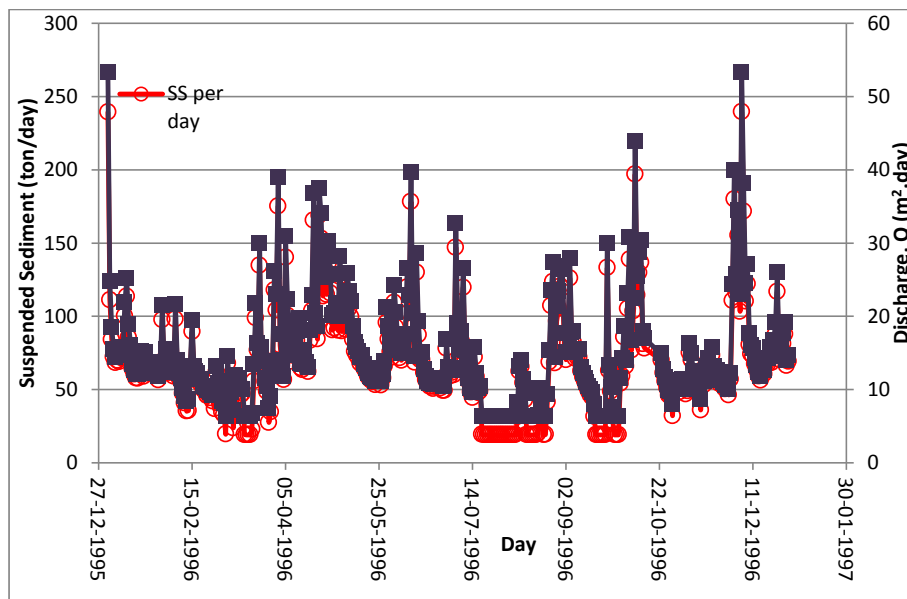


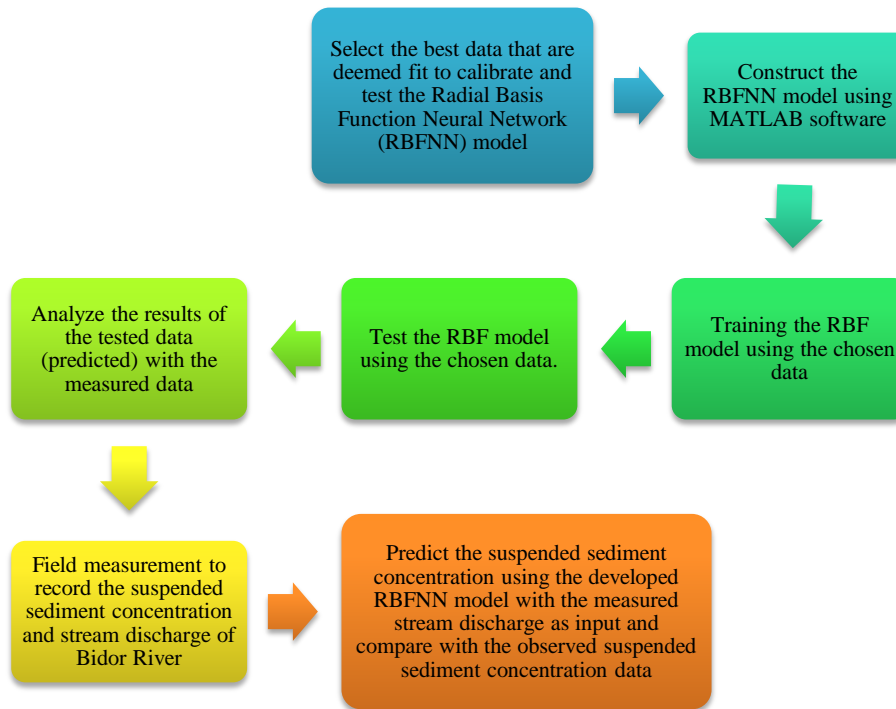
Figure 6: Time Series of comparison between daily suspended sediment concentration and daily stream discharge of testing data

Formatted: Normal, No bullets or numbering

d. Normalization Data

Data normalization is important to signify the data in their single form or commonly known as standard form especially in term of computer science. The data in the context of this study are the daily suspended sediment concentration and daily stream discharge value. Hence, it is important during the processes of RBFNN model to produce the estimation in a short time. In this study, the data were normalized between 0 and 1.

3.5 SUMMARY OF RBFNN MODEL DEVELOPMENT



Formatted: English (U.K.)

Formatted: Heading 2, Left, Indent: Left: 0 cm, Hanging: 1 cm, Space After: 12 pt, Line spacing: single, Numbered + Level: 1 + Numbering Style: 1, 2, 3, ... + Start at: 1 + Alignment: Left + Aligned at: 1.9 cm + Indent at: 2.54 cm

Formatted: Font color: Auto, Not Highlight

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto, Highlight

Figure 7: Summary of RBFNN model development

Formatted: Font: (Default) +Body (Calibri), Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Text 1

Formatted: Font color: Text 1

Formatted: Caption, Centered, Line spacing: single

3.53.6 PROJECT ACTIVITIES

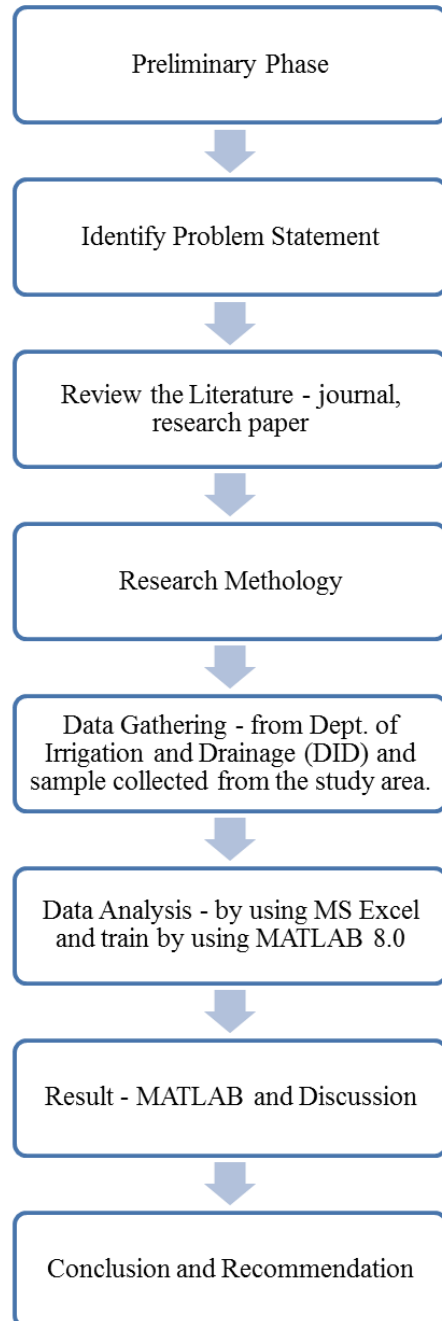


Figure 887: Project Activities

3-63.7 GANTT – CHART / STUDY PLAN

Table 3: Study Plan for FYP 1

NO	ACTIVITIES	WEEK														
		WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6	WEEK 7	WEEK 8	WEEK 9	WEEK 10	WEEK 11	WEEK 12	WEEK 13	WEEK 14	
1	<u>Topic Chosen</u>	-	-	-	-	-		SEMESTER BREAK	-	-	-	-	-			
2	<u>Understanding the research topic</u>	-	-	-	-	-			-	-	-	-	-			
3	<u>Research for journals and materials related to topic</u>	-	-	-	-	-			-	-	-	-	-			
4	<u>Consultation with Supervisor</u>	-	-	-	-	-			-	-	-	-	-	-		
5	<u>Preliminary research work</u>		-	-	-	-			-	-	-	-	-	-		
6	<u>Extended Proposal</u>	-	-	-	-	-				-	-	-	-			
7	<u>Start co-operating with external parties involved</u>								-	-	-	-	-	-		
7	<u>Project Defence and Progress Evaluation</u>	-	-	-	-	-				-	-	-	-			
8	<u>Draft Report</u>	-	-	-	-	-			-	-	-	-	-	-		
9	<u>Final Report</u>	-	-	-	-	-			-	-	-	-	-	-	-	

Formatted: Font: (Default) Times New Roman, 10 pt, Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, 10 pt, Not Bold, Font color: Auto

Table 4: Study Plan for FYP 2

NO	ACTIVITIES	WEEK															
		WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6	WEEK 7	WEEK 8	WEEK 9	WEEK 10	WEEK 11	WEEK 12	WEEK 13	WEEK 14	WEEK 15	
1	<u>Project work continues</u>	-	-	-	-	-	-		SEMESTER BREAK								
2	<u>Progress report</u>	-	-	-	-	-	-										
3	<u>Project work continues</u>	-	-	-	-	-											
4	<u>Pre-SEDEX</u>	-	-	-	-	-											
5	<u>Draft Report</u>																
6	<u>Dissertation</u>	-	-	-	-	-											
7	<u>Technical Paper</u>																
8	<u>Oral Presentation</u>	-	-	-	-	-											
9	<u>Project Dissertation (Hard Bound)</u>	-	-	-	-	-											

Formatted: Font: (Default) Times New Roman, Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, Not Bold, Font color: Auto

Formatted Table

Formatted: Font: Bold, Font color: Black

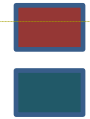
Formatted: Indent: Left: 0.1 cm, Line spacing: Multiple 1.15 li, Position: Horizontal: 5.04 cm, Relative to: Page, Vertical: 0.19 cm, Relative to: Paragraph

Formatted Table

Formatted Table

Formatted Table

Legend:



Key Milestones

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt, English (Malaysia)

3.73.8 KEY MILESTONE

Final Year Project (FYP) is divided into 2 parts i.e. FYP 1 and FYP 2. Figure below is the milestone for both FYP 1 and FYP 2.



Figure 998: Key Milestone (FYP 1)

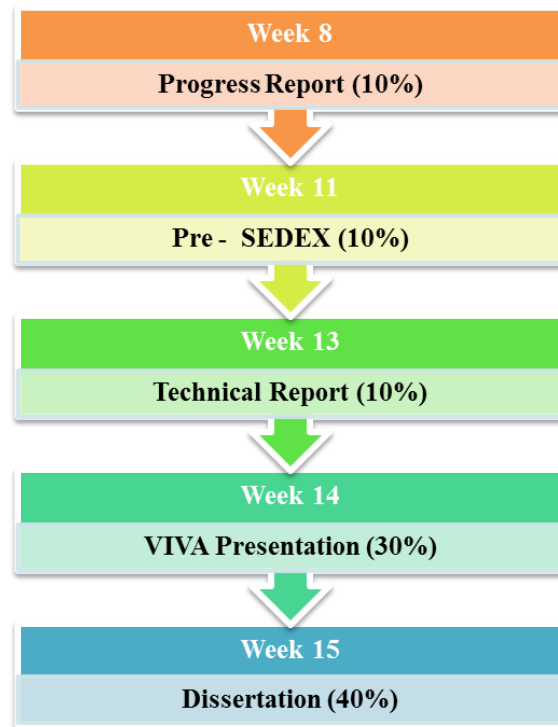


Figure [10109](#): Key Milestone (FYP 2)

CHAPTER 4

RESULT AND DISCUSSION

A total of 1766 data has been used as training and testing data to perform the RBFNN modelling. Before start the modelling process, the data will be analysis where the result shown in **Table 2**. From that, the minimum and maximum value of sediment load was determined where, in this study the value of $z_{\min} = 6$ ton/day and $z_{\max} = 453$ ton/day that getting from the training data. This value used to normalization of raw data by using the equation (1),

$$r_p = 2 \times \frac{(z_p - z_{\min})}{(z_{\max} - z_{\min})} - 1 \quad \dots\dots\dots (1)$$

where, r_p is the normalized data series, z_p is the original data series and z_{\min} , and z_{\max} are the minimum and the maximum value of the original data series, separately. To get the values z_p , the equations as per shows below:

$$z_p = r_p(z_{\max} - z_{\min}) + z_{\min} \quad \dots\dots\dots (2)$$

The data input was selected by using the RBFNN modelling that been created in MATLAB Software. Before normalization the data, the number of hidden neuron must be defined as shown in **Table 5**.

Table 5: Value of Hidden Neurons, Training and Testing Error using trial and error method

No. of Trial	Hidden Neuron	Training Error	Testing Error
1	4	0.016016	0.012004
2	5	0.014795	0.0088947
3	6	0.014563	0.007096
4	7	0.013818	0.010439
5	8	0.011913	0.0091901
6	9	0.0055975	0.0056614
7	10	0.0066494	0.0071963
8	11	0.0028625	0.003511
9	12	0.0016113	0.0019323
10	13	0.001105	0.0013529
11	14	0.0039881	0.0034766
12	15	0.0029207	0.0035477
13	16	0.0025996	0.0030364
14	17	0.0017339	0.0023398
15	18	0.0012838	0.0020942
16	19	0.00076738	0.0012665
17	20	0.00067009	0.00071254
18	21	0.001586	0.0024601
19	22	0.00036579	0.00040911
20	23	0.00076217	0.0012244
21	24	0.0008064	0.0010137
22	25	0.0010735	0.0011902
23	26	0.00022786	0.00021596
24	27	0.00035342	0.00033039
25	28	0.00023771	0.00037321
26	29	0.00037548	0.00037453
27	30	0.0001156	0.00016037
28	31	0.00010439	0.00011022
29	32	0.00012204	0.00014381
30	33	0.0001253	0.00014275
31	34	0.00010486	0.00015412
32	35	0.000076965	0.00028979
33	50	0.00017984	0.00035239
34	70	0.000018324	0.00010027
44	100	9.9027E-06	0.000060877
	Min	0.00036579	0.00040911

The method to determine the number of neuron in the hidden layer is trial and error method. **Table 5** shows the details of trial and error method in defining the number of neuron in hidden layer. The minimum error shows from the **Table 2-5** is 0.00036579 for *training error* and 0.00040911 for *testing error* which located at neuron no. 22. From this data, the values r_p can be obtained by using the MATLAB software by using the RBFNN modelling. The values of r_p used to calculate the values of z_p based on equation (2). Then, the performance of statistic for RBFNN modelling during training and testing will obtain and shown in **Table 6**.

Formatted: Font: Bold

Table 6: Performance statistic of RBFNN modelling during training and testing

	RMSE	MAE	CE	R ²
TRAINING	8.56	5.48	0.9798	0.9801
TESTING	9.06	6.00	0.9360	0.931

Formatted Table

The performance statistics of RBFNN modelling during training and testing can be obtained by using the standard statistics measures such as the mean square error (MSE), the root mean square error (RMSE), the mean absolute error (MAE) and the coefficient of efficiency (CE) based on equation below:

$$MSE = \frac{\sum_{s=1}^N (\hat{u}_s - u_s)^2}{N} \dots\dots\dots (3)$$

$$RMSE = \sqrt{\frac{\sum_{s=1}^N (\hat{u}_s - u_s)^2}{N}} \dots\dots\dots (4)$$

$$MAE = \frac{\sum_{s=1}^N |\hat{u}_s - u_s|}{N} \dots\dots\dots (5)$$

$$CE = 1 - \frac{\sum_{s=1}^N (\hat{u}_s - u_s)^2}{\sum_{s=1}^N (\hat{u}_s - \bar{\hat{u}})^2} \dots\dots\dots (6)$$

Where, \hat{u}_s and u_s are the predicted and observed values of target separately, $\bar{\hat{u}}$ is the mean of predicted target values, and N is the number of observations for which the error has been computed. Ideally, the value of MSE, RMSE and MAE should be zero and CE should be one. Based on **Table 4**, the MSE, RMSE and MAE for training value are slightly higher due to the numerical range are within limited from 10,000 ton/day to 100 ton/day as lowest limit rather than the value of testing data. The most important parameter is coefficient of efficiency (CE). From **Table 4**, the values of CE for training data is 0.9798 but for testing data the value that shown is 0.931.

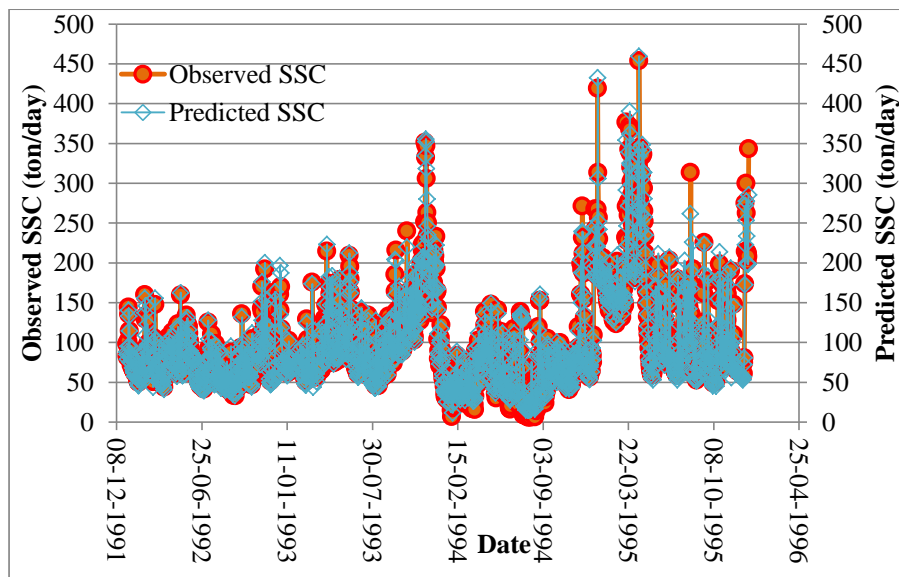


Figure 111110: Comparison of observed and prediction SSC for Training Data

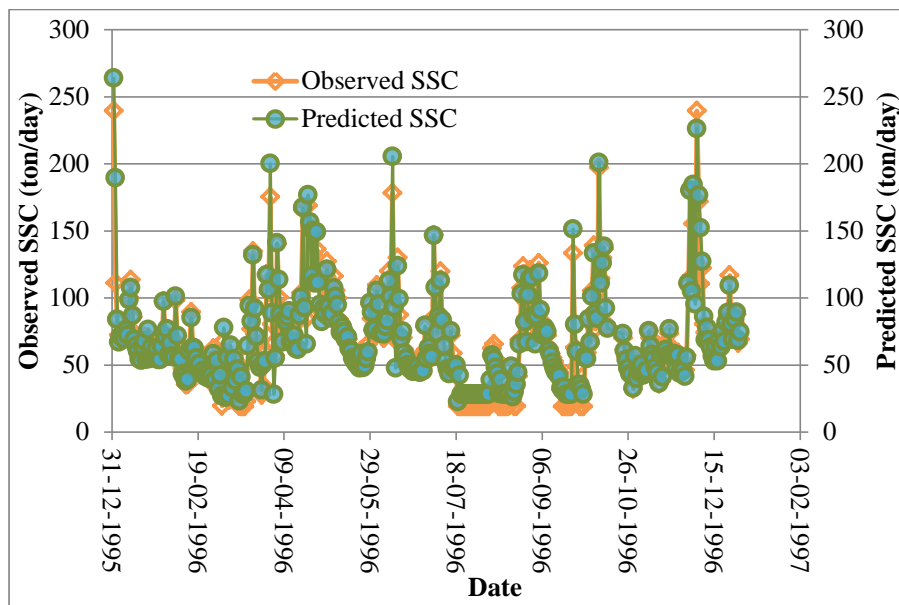


Figure 121211: Comparison of observed and prediction SSC for Testing Data

Based on **Figure 10-11** and **1112**, shown that the comparison of observed and prediction SSC for training and testing data. It can be identified that the most of the predicted values are closely to the observed pattern. From **Figure 1011**, it shows that the ~~input and target both~~-values gave to the model input corresponded to the output. The predicted value followed the trend of the observed data where it shows the values are closed within the data trend. From here, the model can learn the exact pattern of observed data very well which the model was trained well during the training stage. While from **Figure 11-12** the testing only used input data at the model where the predicted output result based on the training stage and there are no target values in testing stage. It is similar with training stage which is predicted value closed to the observed value and shown that this model follow the exact pattern in observed data. These study shows that is no outliers that could reduce the accuracy of prediction. For the time being, the result shown based on the comparison of observed and prediction of SSC shows that the modelling can be used to predict the suspended sediment accurately rather than other modelling.

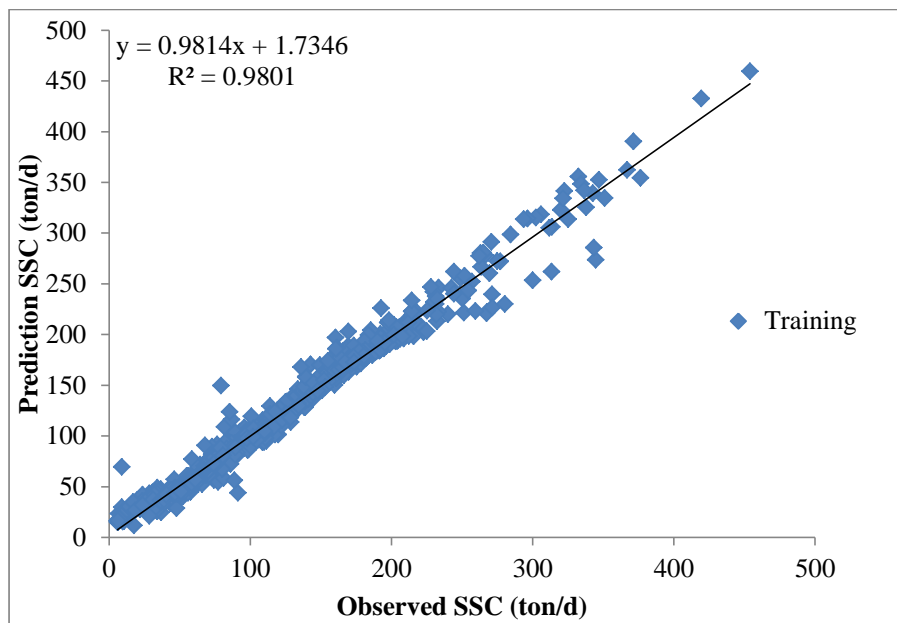


Figure 131312: Scatter plot of prediction versus observed Suspended Sediment Concentration for Training Data

~~Based on~~ **Figure 1213** show the scatter plot of prediction and observed of suspended sediment concentration for the training data where, it can be gotten that the consequence, which is the predicted suspended sediment concentration, is highly correlated with the observed value. **From Figure 13** shows that there are few data are not plot in line where the higher and lower values are visible. This could imply that the RBFNN model developed is able to learn pattern and produce accurate prediction. From the scatter graph shown that the R^2 value for training data is 0.9801 , where it is closed to 1.

Formatted: Not Highlight

Formatted: Not Highlight

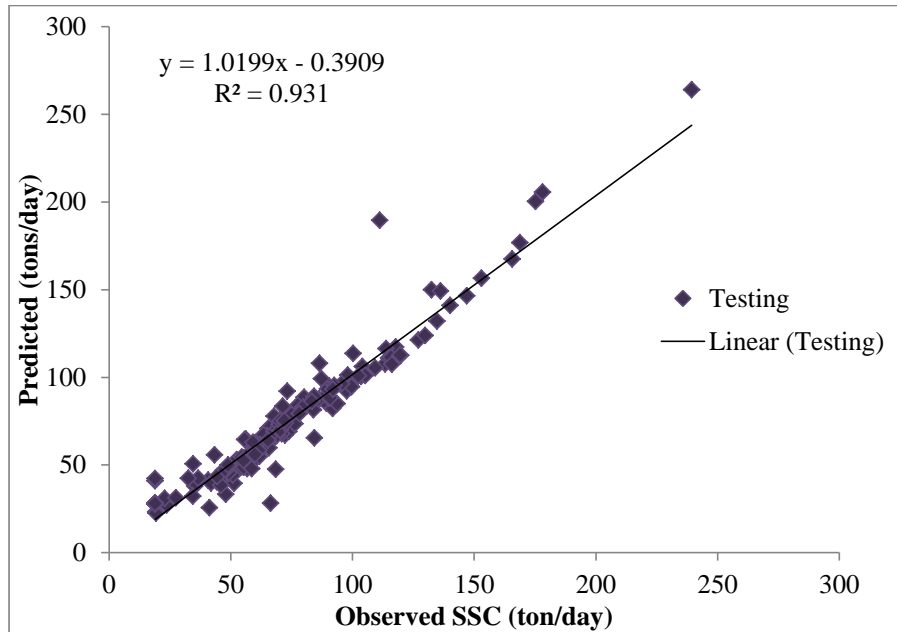


Figure 14: Scatter plot of prediction and observed suspended sediment concentration for testing data

Based on Figure 13, show the scatter plot of prediction and observed of suspended sediment concentration for the testing data. displays the plotting of observed and predicted suspended sediment concentration for testing data. The R^2 value of testing data is slightly lower than training data, where there are few outliers that can be identified. This could imply that the RBFNN model developed is able to learn pattern and produce accurate prediction. The R^2 value for testing data is 0.931. There are few outliers that can be identified. Most of the outliers exist at the peak of the time series. The higher and lower values shown in the Figure 14 was visible where there are outlier in due to not in the line. The reason this model result compared with other models was due to the values of R^2 which can show the result near to 1.

Formatted: Not Highlight

Formatted: Not Highlight

From previous studies, the R^2 values from different models are comparable to the RBFNN model. The result of the R^2 can be shown in the Table 7. Most of the R^2 value from the previous studies is lower compared to the R^2 for RBFNN model.

Table 7: The values of R^2 from different models

Numerical Soft Computing	R^2	Authors
FL	0.917	Demirci, M. and Baltaci, A. (2012)
GEP	0.74	Azamathulla et al. (2012)
WANN	0.846	Liu et al (2013)
RBFNN	0.931	Malim, M. (2013)

Formatted Table

From **Table 7**, the value of R^2 from previous studies has been indicated where the lowest value is $R^2 = 0.74$ by using the GEP model. While the higher and closed to 1 is RBFNN model with $R^2 = 0.931$. From here the RBFNN model show that is the most accuracy of prediction the suspended sediment where the value close to 1. The reason this model result compared with other models was due to the values of R^2 which can show the result near to 1.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

As for the conclusion, the RBFNN modelling using Thin Plate Spline function was developed to predict the suspended sediment in Bidor River. At this point, after been through series of assessment, it can be concluded that this project can be delivered by referring the previous technique and overcome the leak of the previous works. The RBFNN model can be used as predictor the suspended sediments, ~~where the result is closed to the observed result.~~ The performance of statistic shown the accuracy produced by the RBFNN model was comparable with the previous studies.

As for recommendation, this study will be continues by combine the RBFNN model together with the SVM since the period of this study quite short. More advance research will be carried out in the future. Moreover, the project can be deemed as success only if the time management can be performed efficiently, which will be more efficient when the task can be delivered in a shorter duration. Besides, more time can be consumed on analysis of the final data which is the most important in making sure the objective of this project can be achieved successfully.

CHAPTER 6

REFERENCES

- 1.1 Colby, B.R., Hembree, C.H., and Jochens, E.R., (1953). Chemical quality of water and sedimentation in the Moreau River Drainage Basin, South Dakota: U.S. Geological Survey Circular 270, 53 p
- 1.2 Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Trans Inf Theory ; 36(5).
- 1.3 Demirci, M., & Baltaci, A. (2012). Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. Neural Comput & Applic.
- 1.4 Guven, A., & Kişi, Ö. (2011, 01 01). Estimation of Suspended Sediment Yield in Natural Rivers Using Machine-coded Linear Genetic Programming. Retrieved from Springer Link:
<http://link.springer.com/article/10.1007%2Fs11269-010-9721-x>
- 1.5 Imen M., (2008). Estimation of daily suspended sediments using support vector machines, Hydrological Sciences Journal, 53:3, 656-666.
- Kisi. O., Dailr, A. H., Cimen, M. and Shiri, J. 2012 —Suspended sediment modeling using genetic programming and soft computing techniques Journal of Hydrology 450-451 (2012): 45-58
- 1.6 Kunkel, A. (2002). Suspended Sediment in South San Francisco Bay. Retrieved from
<http://www.crrw.utexas.edu/gis/gishydro03/Classroom/trmproj/Kunkel/Term%20Paper.htm>
- 1.7 Lafdani, E. K., Nia, A. M., & Ahmadi, A. (2013). Daily suspended sediment load prediction using artificial neural networks and support vector machines. Journal of Hydrology, 50-62.

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Left, Indent: Left: 0.63 cm, Hanging: 0.87 cm, Line spacing: 1.5 lines, No bullets or numbering

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Default Paragraph Font, Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Left

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Left, Indent: Left: 0.63 cm, Hanging: 0.87 cm, Line spacing: 1.5 lines, No bullets or numbering

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Liu, Q.-J., Shi, Z.-H., Fang, N.-F., Zhu, H.-D., & Ai, L. (2013). Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet-ANN approach. *Geomorphology*, 181-190.

Formatted: Font: (Default) Times New Roman, 12 pt, English (U.K.)

Formatted: Left

Melesse, A.M, Ahmad, S., McClain, M.E., Wang, X. and Lim, Y.H., 2011 —Suspended sediment load prediction of river systems: an artificial neural network approach. *Agricultural Water Management* 98(2011): 855-866

1.8 Mustafa, M., Rezaei, R., Saiedi, S., & Isa, M. (2012). River Suspended Sediment Prediction Using Various Multilayer Perceptron Neural Network Training Algorithms - A Case Study in Malaysia. *Water Resour Manage*, 1879-1897.

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Left, Indent: Left: 0.63 cm, Hanging: 0.87 cm, Line spacing: 1.5 lines, No bullets or numbering

Nagy, H., Watanabe, K., and Hirano, M. (2002). "Prediction of Sediment Load Concentration in Rivers using Artificial Neural Network Model." *J. Hydraul. Eng.*, 128(6), 588–595. Retrieved from <http://ascelibrary.org/doi/pdf/10.1061/%28ASCE%290733-9429%282002%29128%3A6%28588%29>

Formatted: Left

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Left, Indent: Left: 0.63 cm, Hanging: 0.87 cm, Line spacing: 1.5 lines, No bullets or numbering

1.9 Rajaei, T. (2011). Wavelet and ANN combination model for prediction of daily suspended sediment load in rivers. *Science of the Total Environment*, 2917-2928.

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

1.10 Retrieved from <http://en.wikipedia.org/wiki/Bidor>

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

1.11 Retrieved from <http://www.utar.edu.my/dssc/newvillageproject/Bidor.pdf>

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

1.12 Retrieved from http://en.wikipedia.org/wiki/Suspended_load

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

Ulke, A. *et al.*, (2013). Predicting Suspended Sediment Loads and Missing Data for Gediz River, Turkey. Retrieved from <http://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29HE.1943-5584.0000060>

Formatted: Font: (Default) Times New Roman, 12 pt, (Asian) Chinese (PRC), (Other) English (U.K.)

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

Formatted: Default Paragraph Font, Font: (Default) +Body (Calibri), 11 pt

1.13 Vapnik, V., and Lerner A., (1963). Pattern recognition using generalized portrait method. *Automation and Remote Control*, **24**, 774–780.

Formatted: Left

Formatted: Left, Indent: Left: 0.63 cm, Hanging: 0.87 cm, Line spacing: 1.5 lines, No bullets or numbering, Tab stops: Not at 11.76 cm

