4-11-2018 11:00 AM

Modeling of Suspended Sediment Concentration Using Conventional and Machine Learning Approaches, in Thames River, Canada

Issam Mohamed
The University of Western Ontario

Supervisor
Shah, Imtiaz
The University of Western Ontario Co-Supervisor
Yanful, Ernest
The University of Western Ontario

Graduate Program in Civil and Environmental Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Engineering Science

Abstract

Water resource management, planning, hydraulic design, environmental conservation, reservoir management and operation all require reliable information and data about Suspended Sediment Concentration (SSC). To predict such data, direct sampling and Sediment Rating Curves (SRC) are commonly utilized. Since direct sampling can be risky during extreme weather events and SRC cannot provide satisfactorily dependable results, engineers are trying to propose new precise fore-casting approaches. Various soft computing techniques have been applied to model different hydrological and environmental problems and have showed promising results in this regard. Although many studies have been performed to simulate the phenomena of SSC at numerous rivers and creeks in the literature, the SSC is a site-specific problem. In this study, Adaptive Neuro Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN) models were proposed and compared with the conventional SRC and linear regression methods. Using different combination of measured data from 1993 to 2016 of SSC and simultaneous Stream dis-charge, Water Temperature, and Electric Conductivity for Thames River at Byron Station, London, Canada, several models were trained. Goodness of each model was evaluated using Mean Absolute Error, Root-Mean Square Error and Nash-Sutcliffe Efficiency Coefficient. Results show that ANN models are of a superior accuracy if compared with other approaches in predicting SSC for this river.

Keywords

Suspended sediment concentration, SRC, Linear regression, ANFIS, ANN.

Acknowledgments

I would like to express my sincere gratitude to my advisor, Dr. Imtiaz Shah for his guidance, patience, and advice in the work of this thesis. I also thank Dr. Ernest Yanful for being my co-adviser in the past two years, and Dr. M. Reza Najafi for his constructive advice in the work of this thesis.

I would like to thank the faculty, fellow graduate students, and staff in the Department of Civil and Environmental Engineering at Western University for their help and encouragement. I am also thankful for the Upper Thames River Conservation Authority's (UTRCA) staff for providing me with the required information throughout my research work, special thanks to Karla Young, a monitoring technician at UTRCA, for her help.

Lastly, I would like to address my deep gratitude to my beloved father, Dr. Hassan Mohamed, for his support, advice and encouragement not only during the development time of this thesis but throughout my entire life. The appreciation is also extended to all my family as they are always supportive on every milestone in my life.

Table of Contents

Abstract i

Acknowledgments ii

Table of Contents iii

List of Tables vi

List of Figures viii

List of Appendices xiii

CHAPTER 1 1

1.1 Introduction 1

1.2 Significance of Studying Suspended Sediments 3

1.2.1 Water Quality 3

1.2.2 Channel Navigation 3

1.2.3 Fisheries and Aquatic Habitat 4

1.2.4 Water Supply Plant 5

1.2.5 Hydroelectric Facilities 5

1.3 Problem Statement 5

1.4 Research Objective 7

CHAPTER 2 LITERATURE REVIEW 9

2.1 General Overview 9

2.2 Fuzzy Logic and Conventional Methods 10

2.2.1 Fuzzy Logic Approach 10

2.2.2 SRC Overview 11

2.2.3 Application of Fuzzy Logic and Conventional Methods in Predicting Suspended Sediment Problem 12

2.3 ANN and Conventional Methods 13

2.3.1 ANN Overview 13

2.3.2 ANN Application in Estimating Suspended Sediments 13

2.4 ANFIS, ANN and Conventional Methods 15

2.4.1 ANFIS Overview 15

2.4.2 ANFIS and ANN Application in Estimating the Suspended Sediment Problem 15

2.5 Genetic Programming (GP) 17

2.5.1 General Overview 17

2.5.2 GP Approach in Sediment Modeling 17

2.6 Other Modelling Techniques 18

CHAPTER 3 METHODOLOGY, AREA OF STUDY AND DATA COLLECTION 20

3.1 Conventional Approaches 20

3.1.1 Sediment Rating Curve 20

3.1.2 Linear Regression 21

3.2 Machine Learning Approaches 23

3.2.1 Artificial Neural Networks 24

3.2.2 Adaptive Neuro-Fuzzy Inference System 38

3.2.3 Study Area and Datasets 44

CHAPTER 4 MODELS DEVELOPMENT AND RESULTS 50

4.1 Data Preprocessing 50

4.1.1 Training Dataset 52

4.1.2 Testing Dataset 54

4.1.3 Models Performance Evaluation 56

4.2 Input Variable Scenarios 58

4.2.1 Training Various Models for S1 59

4.2.2 Training Various Models for S2 69

4.2.3 Training Various Models for S3 77

4.2.4 Training Various Models for S4 86

CHAPTER 5 SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS..... 95

5.1 Summary 95

5.2 Conclusions 96

5.3 Future Work Recommendations 98

References 99

Appendices 111

List of Tables

Table 1-1 Sediment concentration (mg/l) and the corresponding risk to fish and their habitat (Birtwell, 1999) 5

Table 3-1 ANFIS hybrid learning process 44

Table 4-1 Samples of Q data duplications 50

Table 4-2 Different scenarios proposed for this study 59

Table 4-3 Statistical measures for the training and testing phases of SRC model (S1) 60

Table 4-4 Outputs of the SLR performed using S1's training dataset 62

Table 4-5 Statistical measures for the training and testing phases, SLR model (S1) 62

Table 4-6 Statistical measures for the training and testing phases, ANFIS model (S1) 64

Table 4-7 Statistical measures for the training and testing phases, ANN model (S1) 67

Table 4-8 Summary of different performance indicators for all models of S1 68

Table 4-9 Uncertainty estimates for S1 various models 69

Table 4-10 Outputs of the MLR performed using S2's training dataset 69

Table 4-11 Statistical measures for the training and testing phases, MLR model (S2) 70

Table 4-12 Statistical measures for the training and testing phases, ANFIS model (S2) 72

Table 4-13 Statistical measures for the training and testing phases, ANN model (S2) 75

Table 4-14 Summary of different performance indicators for all models used using the S2 inputs 76

Table 4-15 Uncertainty estimates for S2 various models 77

Table 4-16 Outputs of the MLR performed using S3's training dataset 77

Table 4-17 Statistical measures for the training and testing phases, MLR model (S3).......... 78

Table 4-18 Statistical measures for the training and testing phases, ANFIS model (S3)....... 80

Table 4-19 Statistical measures for the training and testing phases, ANN model (S3).......... 82

Table 4-20 Summary of different performance indicators for all models used using the S3 inputs.. 84

Table 4-21 Uncertainty estimates for S3 various models 85

Table 4-22 Outputs of the MLR performed using S4's training dataset................................ 86

Table 4-23 Statistical measures for the training and testing phases, MLR model (S4).......... 86

Table 4-24 Statistical measures for the training phase of the ANFIS model (S4)................. 89

Table 4-25 Statistical measures for the training phase of the ANN model (S4)..................... 92

Table 4-26 Summary of different performance indicators for all models used using the (S4) inputs.. 93

Table 4-27 Uncertainty estimates for S4 various models 94

Table 5-1 Final structure of various machine learning approaches' best model 95

Table 5-2 Performance indicators for various models... 97

Table 5-3 Uncertainty estimates for various models 97

List of Figures

Figure 1-1 Average annual suspended sediment load (tonnes) for selected rivers in Canada (Government of Canada, 2013).. 2

Figure 3-1 Example of simple artificial neural network (Krenker *et al*. 2011) 26

Figure 3-2 Biological Neuron Design (https://askabiologist.asu.edu/neuron-anatomy) 26

Figure 3-3 Artificial Neuron Design (Haykin, 2008) 27

Figure 3-4 Step transfer function 28

Figure 3-5 Linear (Purelin) transfer function.. 29

Figure 3-6 Log-Sigmoid transfer function 29

Figure 3-7 Tan-Sigmoid transfer function 30

Figure 3-8 Feed-forward (FNN) and recurrent (RNN) topologies of ANN (Krenker *et al*. 2011) 31

Figure 3-9 Multi-layer feed-forward artificial neural network (Krenker *et al.* 2011) 32

Figure 3-10 Fully recurrent artificial neural network (Krenker *et al.* 2011) 33

Figure 3-11 Architecture of MLP feed-forward ANN (Nastos *et al*. 2011) 37

Figure 3-12 Architecture components of a fuzzy inference system 39

Figure 3-13 Architecture of an adaptive network 40

Figure 3-14 Reasoning from a two-input first-order Sugeno fuzzy model with two rule (Foroozesh *et al.* 2013) 41

Figure 3-15 ANFIS architecture corresponding to Figure 3-13 (Foroozesh *et al.,* 2013) 41

Figure 3-16 Upper Thames River Basin (on the left-had side) and River Bend Subwatershed, the research area of study (UTRCA, 2012) 46

Figure 3-17 Scatter plot of the raw river discharge (m^3/s) data over the period between 1993-2016 ... 48

Figure 3-18 Scatter plot of the raw suspended sediment concentration (mg/l) data over the period between 1993-2016 ... 48

Figure 3-19 Scatter plot of the raw electric conductivity (μS/cm) data over the period between 1993-2016 ... 48

Figure 3-20 Scatter plot of the raw river temperature (^0C) data over the period between 1993-2016 ... 49

Figure 4-1 Example data, normal probability plot ... 51

Figure 4-2 Temperature data (^oC) used for training various models ... 53

Figure 4-3 Discharge data (m^3/s) used for training various models ... 53

Figure 4-4 Electric Conductivity data (μS/cm) used for training various models ... 53

Figure 4-5 Suspended Sediment Concentration data (mg/l) used for training various models ... 54

Figure 4-6 River temperature data (^oC) used for testing various models ... 54

Figure 4-7 Discharge data (m^3/s) used for testing various models ... 55

Figure 4-8 Electric conductivity data (μS/cm) used for testing various models ... 55

Figure 4-9 Suspended sediment concentration data (mg/l) used for testing various models.. 55

Figure 4-10 Applied SRC for the training dataset ... 60

Figure 4-11 Observed and calculated SSC (mg/l), the training period using SRC (S1) ... 61

Figure 4-12 Scatter plot comparing predicted and observed SSC (mg/l) using SRC (S1), testing data ... 61

Figure 4-13 Observed and calculated SSC (mg/l), the training period using SLR (S1) ... 63

Figure 4-14 Scatter plot comparing predicted and observed SSC (mg/l) using SLR (S1), testing data 63

Figure 4-15 ANFIS structure and Fuzzy logic designer toolbox for (S1) training phase 64

Figure 4-16 MF editor and rules for the ANFIS model (S1) training phase 64

Figure 4-17 Observed and calculated SSC (mg/l), the training period using ANFIS (S1) 65

Figure 4-18 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S1), testing data 65

Figure 4-19 ANN best structure using all dataset (S1) 66

Figure 4-20 Best ANN model outputs (S1) 66

Figure 4-21 Observed and calculated SSC (mg/l), the training period using ANN (S1) 67

Figure 4-22 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S1), testing data 67

Figure 4-23 Selected peaks of observed SSC (mg/l) from the testing phase period for S1 and the calculated SSC (mg/l) using proposed approaches 68

Figure 4-24 Observed and calculated SSC (mg/l), the training period using MLR (S2) 70

Figure 4-25 Scatter plot comparing predicted and observed SSC (mg/l) using MLR (S2), testing data 70

Figure 4-26 ANFIS model structure and Fuzzy logic designer toolbox for S2 training phase 71

Figure 4-27 MFs editor for the two main inputs T (on the left-hand side) and Q for the ANFIS model, training phase (S2) 71

Figure 4-28 Rule viewer for the ANFIS model, training phase (S2) 72

Figure 4-29 Observed and calculated SSC (mg/l), the training period using ANFIS (S2) 73

Figure 4-30 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S2), testing data 73

Figure 4-31 ANN best structure using all dataset (S2) 74

Figure 4-32 Best ANN model outputs (S2) 74

Figure 4-33 Observed and calculated SSC (mg/l), the training period using ANN (S2)........ 75

Figure 4-34 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S2), testing data 75

Figure 4-35 Selected peaks of observed SSC (mg/l) from the testing phase period for S2 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon................ 76

Figure 4-36 Observed and calculated SSC (mg/l), the training period using MLR (S3)........ 78

Figure 4-37 Scatter plot comparing predicted and observed SSC (mg/l) using MLR (S3). testing data 79

Figure 4-38 ANFIS model structure and Fuzzy logic designer toolbox for S3 training phase 79

Figure 4-39 MFs editor for the two main inputs Q (on the left-hand side) and C for the ANFIS model, training phase (S3).. 80

Figure 4-40 Rule viewer for the ANFIS model, training phase (S3)...................................... 80

Figure 4-41 Observed and calculated SSC (mg/l), the training period using ANFIS (S3)..... 81

Figure 4-42 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S3). testing data 81

Figure 4-43 ANN best structure using all dataset (S3) 82

Figure 4-44 Best ANN model outputs (S3) 83

Figure 4-45 Observed and calculated SSC (mg/l), the training period using ANN (S3)........ 83

Figure 4-46 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S3), testing data 84

Figure 4-47 Selected peaks of observed SSC (mg/l) from the testing phase period for S3 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon 85

Figure 4-48 Observed and calculated SSC (mg/l), the training period using MLR (S4) 87

Figure 4-49 Scatter plot comparing predicted and observed SSC (mg/l) using MLR S(4), testing data 87

Figure 4-50 ANFIS model structure and Fuzzy logic designer toolbox for training phase (S4) 88

Figure 4-51 MFs editor for two of the three main inputs T (on the left-hand side) and Q for the ANFIS model, training phase (S4) 88

Figure 4-52 MFs editor for the input C (on the left-hand side) and Rule viewer for the ANFIS model, training phase (S4) 89

Figure 4-53 Observed and calculated SSC (mg/l), the training period using ANFIS (S4) 89

Figure 4-54 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S4), testing data 90

Figure 4-55 ANN best structure using all dataset (S4) 90

Figure 4-56 Best ANN model outputs (S4) 91

Figure 4-57 Observed and calculated SSC (mg/l), the training period using ANN (S4) 92

Figure 4-58 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S4), testing data 92

Figure 4-59 Selected peaks of observed SSC (mg/l) from the testing phase period for S4 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon 93

List of Appendices

Appendix A: Raw data table used in this study 111

Appendix B Grubb’s test MATLAB Code 129

Appendix C Processed data after removing the duplications and outliers............................ 132

Appendix D Sample spreadsheet of calculations implemented to find the various statistical measures adopted for this study over the training period for the best scenario case. 144

Appendix E Sample spreadsheet of calculations implemented to find the various statistical measures adopted for this study over the testing period for the best scenario case. 154

CHAPTER 1

1.1 Introduction

One of the major factors that plays an important role in the transformation of the Canadian landscape is water. Every year water erodes great amounts of soil from the landscape in the form of sediments. Sediments transported by river systems and eventually deposited in a lake or sea. For example, the Fraser River, British Columbia, carries an average of 20 million tons of sediment a year into the marine environment (Government of Canada, 2013).

Erosion process is the first stage of the sediment cycle, by which particles and fragments are weathered from the earth's surface. Water, wind, glaciers, and plant and animal activities are all the sort of activities that cause erosion to the ground's material. The most effective agents that cause erosion can be counted as water and wind. Deforestation is assumed to be the main reason for erosion. However, other factors counted in this case, such as climate, soil structure, land topography, vegetation cover and land management.

Fluvial sediment is the term utilized to define the case where water is the main cause for erosion. Natural, or geologic, erosion takes place in a slow manner, over centuries or millennia. Erosion that occurs because of human activity may take place much faster. It is important to understand the role of each when studying sediment transport (Government of Canada, 2013).

Fluvial sediment transportation process initiates when raindrops accumulate and result in sheet flow causing the dislodgeable materials from land surface to be transported. The greater the discharge (i.e. rate of flow), the higher the capacity there is for sediment transport.

Millions of tonnes of sediments are annually move through the Canadian waterways due to the forever erosion cycle, transportation and deposition processes. Figure 1-1 shows average annual suspended sediment load (tonnes) for selected rivers in Canada. Sediment is measured and categorized based on its dynamic characteristics into three major classes:

- suspended load (suspended in the water)
- bed load (rolling or bouncing along the bottom)
- bed material (stationary on the bed)

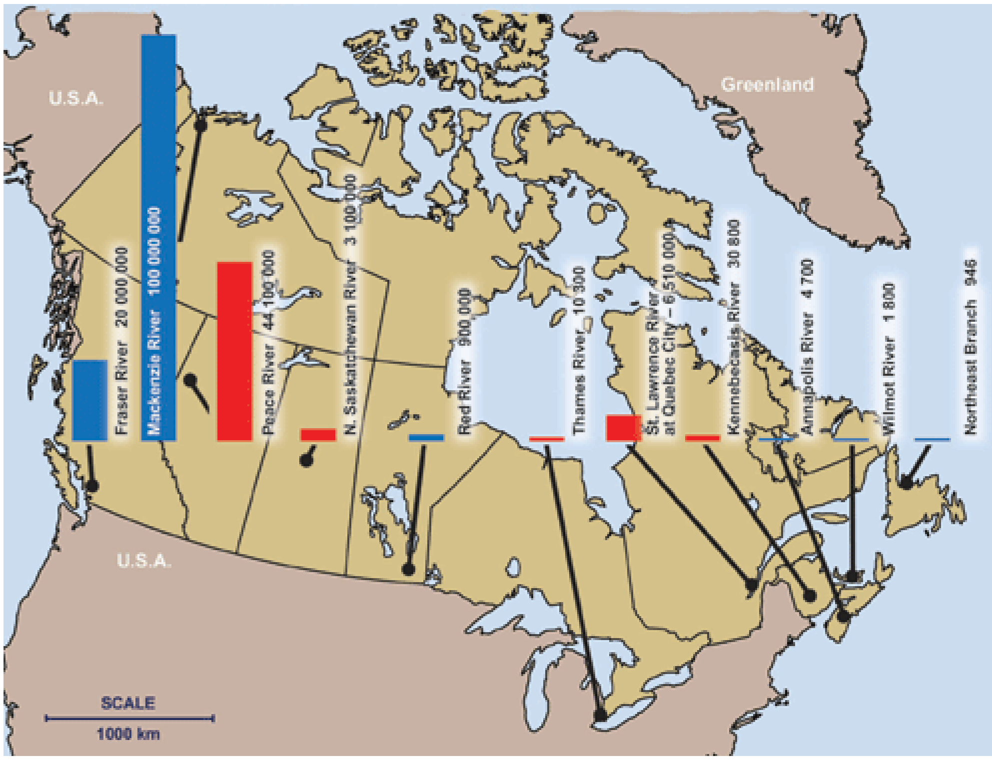


Figure 1-1 Average annual suspended sediment load (tonnes) for selected rivers in Canada (Government of Canada, 2013)

Many researchers have been devoting their studies to the subject of suspended load due to its importance to several topics, for example, erosion around structures, backfilling of dredged channels, pollution, channel navigability, reservoir filling, hydroelectric-equipment longevity, fish habitat, river aesthetics and scientific interests (Demirci & Baltaci, 2013). Sediment transportation process is very complex and often subject to semi-empirical or empirical treatments. On the other hand, most theoretical treatments were developed using idealized and simplified assumptions that the suspended load is only a function of one or two dominant factors, such as water flow rate, average velocity, energy slope and shear stress. Numerous equations have been published accordingly and because of the scarcity in continuous measured data, each equation was developed using limited laboratory and occasional field data. A drastic difference between the calculated results from the various equation and the measured data is often noted. Consequently, none of these published equations have gained universal acceptance in predicting sediment transport rates, especially in rivers. Recently, due to the advancement in the machine learning and modeling software, computer models have been developed to simulate and predict the erosion, sediment transport processes and suspended load.

1.2 Importance of Studying Suspended Sediments

Suspended sediment carried in a river has an essential impact in various aspects (e.g. water quality, navigation, fisheries and aquatic habitat). It is a site-specific problem that depends on a several factors (e.g. the catchment area, rain fall intensity, vegetation cover) and should be studied for every river, creek, channel, etc.

1.2.1 Water Quality

Toxic chemicals, including most heavy metals and the majority of the US-EPA Priority Pollutants (including 96% of categories NO 1 and NO 2-most toxic and persistent pollutants) and many other unlisted but environmentally sensitive chemicals, have environmental pathways that are primarily or exclusively associated with sediment and biological substrates (Ongley *et al*., 1992). Toxic chemicals can be adsorbed or attached to sediment particles and therefore flow with them along the stream and would find its way to the recipients open water causing a major effect on the water quality issue. Environmentalists and engineers throughout studying quality, quantity and characteristics of sediment in the stream can help reducing the impact of pollutants on the aquatic environment. Toxic chemicals and its association with sediments is an important matter for the national water quality issues.

1.2.2 Channel Navigation

The deposition of sediments in reservoirs and navigation channels weakens or diminishes the performance of these projects and increases the overall project costs. To maintain navigability of such channels, some of the sediments has to be dredged from the stream.

However, toxic chemicals accompanying the dredging process, water levels and sediment transport rate should be studied carefully to determine how much sediments needs to be removed and how frequent. Sedimentation of navigation channels is a concern in the Fraser River (British Columbia), the Mackenzie River (Northwest Territories), and the Great Lakes-St. Lawrence system (Ontario and Quebec) (Government of Canada, 2013).

1.2.3 Fisheries and Aquatic Habitat

Several reports, such as European Inland Fisheries Advisory Commission (EIFAC), (1964), Hollis, Boone et al. (1964), Lloyd (1987), Newcombe and Macdonald (1991), Waters (1995), Anderson et al. (1996), Caux et al. (1997), examine and provide information on the effect of sediments and turbidity on fish and their habitat. Results of these studies determined that presence of sediments in a given stream would harm/kill fisheries and habitats in multiple ways:

- Certain concentrations of sediment kill fish directly. Table 1-1 shows sediment concentration and the adjacent level of risk to fish and their habitat.

- Fish feeding and schooling practices is directly impacted by the penetrated light into water, which decreases by suspended sediment.

- Stream temperature increased by the adsorbed warmth from the sun through sediment particles, stressing some species of fish.

- Plants, invertebrates, and insects in the stream bed can be dislodged due to high concentration of suspended sediments. This affects the food source of fish and can result in smaller and fewer fish.

- Settling sediments can cause fish eggs to be buried and suffocated.

Table 1-1 Sediment concentration (mg/l) and the corresponding risk to fish and their habitat (Birtwell, 1999)

Sediment concentration (mg/l)	Risk to fish and their habitat
0	No risk
<25	Very low risk
25-100	Low risk
100-200	Moderate risk
200-400	High risk
>400	Unacceptable risk

1.2.4 Water Supply Plant

The presence of sediments in the main water source can cause an extensive damage to the pumps and turbines. Accordingly, the delivery of water supply for domestic, agricultural and industrial uses will be affected. A study determining the sediment amount in the raw water source can help choosing the proper equipment for the water supply plant and thus reducing the project's total cost.

1.2.5 Hydroelectric Facilities

According to Morris et al. (2008), "About 0.5% to 1% of the total volume of 6,800 km^3 of water stored in reservoirs around the world is lost annually as a result of sedimentation". Size of the reservoir and its life expectancy span both are directly dependent on the amount of sediment discharged from the upstream. Sediments that are trapped in the upstream by a hydraulic structure, such as dams, accumulating with time, decreasing the size of the reservoir and consequently affecting the capacity of electricity that might be generated.

1.3 Problem Statement

According to Heng and Suetsugi (2013) the measurement of sediment concentration is deficient in most parts of the world. Several hydrological variables such as bed-form geometry, flow rate, friction factor and discharge have been used to develop different models for predicting sediment concentration in rivers (Karim & Kennedy, 1990; Lopes & Ffolliott, 1994). Direct analysis of the Suspended Sediment Concentration (SSC) and the Sediment Rating Curve (SRC) method are among the two tools used in a wide range to obtain the suspended sediment load. Although direct analysis method is the most reliable method, but it is very costly, time consuming, and in many cases, problematic for inaccessible sections, especially during severe storm events, and cannot be conducted for all river gauge stations (Bayram et al., 2013). On the other hand, because SSC transported in a river is a complex hydrological phenomenon due to presence of several ambiguous parameters such as spatial variability of basin characteristics, river discharge patterns and inherent non-linearity in hydro-meteorological parameters, the conventional SRC method may not be advantageous in estimation of SSC which takes into consideration only Q as an effective variable and assumes that all other effective variables are inherited in such independent variable (Joshi et al., 2015) so does the Regression Models (RM), in which the system is assumed to be static (Ghorbani et al., 2013). In Thames River at Byron station, the sampling for SSC is infrequent, this lack of continuous information about SSC can result in substantial errors in estimates of the SSC using the conventional SRC and RM methods. This imposed the necessity to use the artificial intelligence models for more accurate prediction (Kisi et al., 2012).

During the last two decades, artificial intelligence techniques to estimate and predict various hydrological phenomena has being utilized (Tachi, 2017). Adaptive Neural Fuzzy Inference System (ANFIS) and Artificial Neural Network (ANN) are amongst the two well-known models for prediction and simulation in vital hydrology and hydraulics topics such as prediction of SSC (Angabini et al., 2014). This study therefore aims to develop efficient ANFIS and ANN models in predicting the SSC in River Thames, London, ON, and to compare their results with one another and with the results of the conventional SRC and RM methods. This will be based on the available data concerning the inputs (discharge, temperature and electric conductivity) and the output (SSC) variables.

1.4 Research Objective

The goal of this study is to:

1. Develop the best model to estimate the Suspended Sediment Concentration (SSC) in Thames River, London, ON, using two conventional approaches (i.e. Sediment rating curves and linear regression) and two artificial intelligence modeling approaches (i.e. Adaptive neuro fuzzy inference system and artificial neural networks) by means of various effective input variables combinations.

2. Compare between various models using several statistical measures, including mean absolute error (MAE), root mean square error (RMSE) and Nash-Sutcliffe efficiency (NSE), along with the uncertainty analysis, to select the best model.

In order to achieve the above objective the following tasks were performed:

- Data collection (including river discharge, river temperature, SSC, water electric conductivity) for the site over the period from 1993 to 2016.

- Data classification and preprocessing using the Grubbs test for multiple outliers.

- To develop a SSC predictive model using the SRC method (river discharge as the only input).

- To develop a SSC predictive model using the Simple Linear Regression (SLR) method (river discharge is the only input).

- To develop a SSC predictive model using The Multiple Linear Regression (MLR) method (a combination of inputs including river discharge, river temperature and water electric conductivity).

- To develop a SSC predictive model using the ANFIS model (a combination of inputs including river discharge, river temperature and water electric conductivity).

- To develop a SSC predictive model using the ANN model (a combination of inputs including river discharge, river temperature and water electric conductivity).

- To compare the accuracy of each individual model in order to select the most accurate predictive model using different statistical parameters, including, Root Mean Square Error (RMSE), Nash-Sutcliffe Efficiency coefficient (NSE) and Mean Absolute Error (MAE) between the measured and computed data results.

- To conduct uncertainty analysis as an additional performance indicator and use its parameters to select the best model.

- To compare the performance of the most accurate model from each approach with the others to select the best predictive model among all models using the same statistical measures.

CHAPTER 2
LITERATURE REVIEW

2.1 General Overview

Sediment loading in rivers plays a major role when it comes to the availability of freshwater resources, it also threatens the aquatic habitat, alters stream geomorphology, and reduces the reservoir capacity behind various hydraulic structures (Foster and Charlesworth, 1996; Kondolf, 1997; Owens *et al.* 2005; Wood and Armitage, 1997). Freshwater ecosystems can be harmed by the sediments which transport in suspension; such suspended sediments reduce the transmission of sunlight, cause increases in the surface water temperatures, and interfere with the aquatic biota's metabolic processes. The reduction of sunlight transmission caused by the presence of suspended sediments may limit or even restrict photosynthesis, hence a huge influence on the aquatic food chain results. Contaminants such as heavy metals, metal nutrients and other pollutants often use the suspended sediments as a conveyance system (Foster and Charlesworth, 1996; Uri, 1999). Recreational activities in various streams can also be reduced due to the effects of sediment on freshwater biota and aesthetics. Higher concentrations of suspended sediment can cause extensive damage to pumps and turbines and accordingly can increase the cost of water treatment processes. Reservoir storage capacity loss occurs due to the sediments that settles out of suspension and starts to fill the reservoir, impede channel navigability and increase flooding rate of recurrence and its harshness (Uri, 1999; Williams, 1989). Not only the excessive existence of suspended sediment in a stream can be looked at as a harmful phenomenon, the very minute suspended sediment concentration (lower than the natural background levels) can also cause as destructive consequences to the stream geomorphological and ecological activities as too high sediment concentration may cause. Suspended sediment concentrations are usually lower in the watercourses connected with dams and reservoirs. Because the more the water body deprived of sediment the greater the potential energy to expand on erosion processes, the streams which have lower sediment concentrations often experience higher channel erosion and channel incision. Channel incision and building up (larger particle sizes) of bedload materials occurs due to the extra energy. Flow energy dissipation causes the channel to erode, therefore; the more frequent and the higher the flow, the larger the channel erosion rate (Wolman and Schick, 1967). Aquatic habitat can also be significantly changed due the deprivation of sediments in various watercourses (Kondolf, 1997).

Many scientists and researchers have been conducting numerous studies to understand and examine the mechanism of sediment transport in natural streams during the past decades. The estimation of sediment discharge is being among the most important area of interest to various researchers due to its great impact on stream training, stream management, stream engineering applications, and different hydraulic structures construction, in addition to the previously discussed environmental effects and how it is a key to water quality treatment. In the literature, practical formulas and several predictive modeling approaches to predict the sediment yield have been developed. Yet there has been slight or not at all success in predicting the sediment load using the classical techniques (Tuan and Shibayama, 2003). SRC are the most common method utilized in estimating the sediment discharge when sampling and measurements on site are limited. SRC are empirical relations between total flow discharge and total sediment discharge, in the form of $S = aQ^b$, where S is the SSC, Q is the discharge, and a and b are constant and vary for every stream (Campbell and Bauder, 1940). However, because of the linearity assumed in this case (SRC) and in the several other time series conventional practices, the results usually are inadequate. With the advance in the artificial intelligent models, in the literature many researchers have been comparing the performance of conventional methods such as SRC and MLR methods with different soft computing techniques (i.e. Fuzzy logic, ANN, ANFIS, Genetic programming (GP), etc.).

2.2 Fuzzy Logic and Conventional Methods

2.2.1 Fuzzy Logic Approach

Fuzzy logic systems define the relationship between the input variables and output variables of a model using fuzzy if-then statements (Adriaenssens *et al.* 2004). The input space is divided into different overlapping fuzzy sets, this results in a fuzzily defined interval. This dividing process of the inputs into fuzzy sets is referred to as fuzzification. Other important components of the fuzzy rule based system (FRBS) are the membership function, domain partitions and IF-THEN inference rules. The membership function is also known is the degree of belongingness to a particular fuzzy set and its value ranges from 0 to 1. For every membership set there are consequents given by the IF-THEN rules. These rules can be fuzzy and then the model is called Mamdani model (Mamdani, 1974) , or it can be linear in which case the model is Takagi-Sugeno-Kang (TSK) model (Takagi and Sugeno, 1985).

2.2.2 SRC Overview

The developing of a SRC is an important matter of hydrology and river engineering. It is a defined as a mathematical relationship between the water discharge and sediment concentration data. Most of the time the SRC is of the type of a power equation as below

$$SSC = aQ^b \tag{1}$$

$$\text{Or, } log\, SSC = log\, a + b\, log\, Q \tag{2}$$

Where SSC and Q stand for suspended sediment concentration (mg/L) and discharge (m^3/s), respectively. In SRC method, the coefficients a and b are constants without physical significance (Kisi and Zounemat-Kermani, 2016); however, Morgan (2009) mentioned that a-coefficient is an erosion severity indicator and the higher the value the easy the material to be transported and the intensive they to be weathered. In the same regard, he stated that the b-coefficient represents the erosive power of a river, high values representing an increase in erosive power due to a slight increase in the river's discharge. The relationship between these variables (a and b) and some river characteristics like grain size of sediment, river channel morphology, erodibility and the stream power of the river basin were also examined by other researchers (Morehead *et al.* 2003; Wang *et al.* 2008; Yang *et al.* 2007)

2.2.3 Application of Fuzzy Logic and Conventional Methods in Predicting Suspended Sediment Problem

The conventional SRC method is a relation between the sediment and the river discharges. Regression analysis is generally used to create such a relationship, and the curves are usually presented in a power equation form (Kisi *et al.* 2006).

Uncertainty in suspended sediment curves was examined by McBean and AI-Nassr (1988) and they concluded that the exercise of using sediment load versus discharge is misleading as the goodness of fit implied by this relation is spurious. As a recommendation, they suggested utilizing the regression between sediment concentration and discharge as an alternative.

Lopes and Ffolliott (1993) mentioned that due to the hysteresis effect, the relationship between SSC and streamflow is further complicated. A power equation is generally utilized to represent sediment rating and its transformation since the conventional regression method is not able to capture the hysteresis effect.

The rating relationship establishment process is basically a non-linear problem. The commonly used statistical measures in such cases are curve fitting and regression. However, due to the complexity of the phenomena, these methods are not able to provide sufficiently accurate results, and an improvement can be further applied. The fuzzy rule-based (FRB) approach introduced by Zadeh (1965) is being widely utilized in different areas of science and technology. It is a qualitative modeling scheme and does not require an extensive previous knowledge of the phenomenon to be studied, it uses linguistic functions and the key idea of FRB is it allows something to be partly this and partly that instead of having it being all this or all that, and this degree of "belongingness" of that set or category can numerically range from 0 to 1.0 (O. Kisi *et al.* 2006). The applicability of the FRB approach has been demonstrated in water quality management field over the past two decades and the results generated much of enthusiasm. In this regard Kisi *et al.* (2006) have employed fuzzy logic modeling approach to predict the SSC and have compared the results with those resulted from the SRC method. The study was done based on a 5-year period of continuous streamflow and SSC data of Quebrada Blanca Station operated by the

United States Geological Survey (USGS). Nine different fuzzy logic and two SRC models were compared and the results showed that those who obtained with the fuzzy logic modeling approach are of a higher accuracy than those which found using the conventional SRC method. These results confirm the ability of fuzzy logic approach to provide a superior alternative to the conventional SRC approach. Fuzzy model presented is site-specific and dose not simulate the hysteresis effect.

2.3 ANN and Conventional Methods

2.3.1 ANN Overview

Artificial neural network (ANN) based model is a massive parallel distributed system for information processing inspired by a research on the nature of the biological structure of the human brain and the nervous system. This processing of the information is done by a number of interrelated neurons or nodes (Ghorbani *et al.* 2013). ANN is recognized as one of the most powerful technique of the Artificial Intelligence (AI) of which hydrologist used during the past two decades. It has helped many researches to control all types of data, and to capture different nonlinear phenomena with its capability of identifying and recognizing the complex interrelationships between inputs and outputs. The most common processing types used in prediction application of ANN is feed-forward back propagation neural network (Nagy *et al.* 2002).

2.3.2 ANN Application in Estimating Suspended Sediments

Artificial neural network (ANN) idea and its applicability to model various complex water-resources problems have been widely accepted as an alternative solid modeling tool especially for prediction (Firat and Güngör, 2010). The ANN model maps the input to output without the need to identify the physics of the priori (Ghorbani *et al.* 2013). Some of the ANN application to hydrology include river flow estimation (H. K. Cigizoglu and Kisi, 2003; Dawson *et al.* 2002; Ö. Kisi, 2007), rainfall-runoff modeling (Jeong and Kim, 2005; Srinivasulu and Jain, 2006), monthly precipitation forecasting (Aksoy and Dahamsheh, 2009), municipal water demand forecasting (Firat *et al.* 2009), sediment transport prediction (H. Cigizoglu and Alp, 2003; H. K. Cigizoglu, 2003).

Jain (2001) used the ANN approach to establish an integrated stage-discharge-sediment concentration relationship. The data of two gauging sites at Mississippi River was used to compare the performance of ANN and the conventional SRC. Study concluded that the ANN results are much closer to the observed values than the conventional technique. Nagy, Watanabe, and Hirano (2002) provided ANN model to estimate the natural sediment discharge in rivers in terms of sediment concentration, using Froude's number, water top width ratio and Reynold's number as inputs. The results of the ANN model were compared with the other conventional formulas and models, the study showed that ANN had the preference in prediction of sediment concentration over the other conventional methods, including the regression analysis. Cigizoglu and Alp (2003) developed a feed-forward back propagation three-layer learning ANN algorithm consisting of an input layer, hidden layer and an output layer to simulate the relationship between suspended sediment, precipitation and river flow by using hydro-meteorological data. The study showed significant improvement in forecasting suspended sediment values after adding the river flow values as an input along with the rainfall's instead of having rainfall as an only input. The performance of the ANN models were better than the multi linear regression and the study suggested taking ANN as an important tool in the problem of forecasting the suspended sediment. Kisi (2004) established three different ANN modeling techniques, namely, multi-layer perceptron (MLP), generalized regression neural networks (GRNN) and radial basis function (RBF), using Levenberg-Marquardt algorithm to predict and estimate daily SSC at two stations on the Tongue River in Montana, USA. The study included various combinations of inputs to better predict the daily SSC, e.g. water discharges at both current and previous time steps, sediment concentrations at previous time steps at the station of interest, as well as data from the upstream station. The study concluded that the MLP method generally gives better SSC estimates over the other neural network techniques as well as the conventional statistical method (MLR). In 2007, Zhu, Lu, and Zhou used the approach of ANN to model the monthly suspended sediment flux, (i.e. SSC multiplied by the water discharge), from 1960 to 2011 in the Longchuanjiang River, the Upper Yangtze Catchment, China. Average rainfall, rainfall intensity, temperature and streamflow discharge were taken as input parameters for constructing the various models of the study. ANN was capable to give the best accuracy in predicting the monthly sediment flux among two other conventional models, the MLR and power relation (PR) approaches. Utilizing precipitation, discharge, and antecedent sediment data from three major rivers of USA (Mississippi, Missouri and Rio Grande) as inputs, Melesse *et al.* (2011) emphasized the same idea introduced by the previously mentioned studies that ANN ability in simulation and prediction of daily and weekly suspended sediment load are of a superior accuracy compared with the conventional MLR, multiple non-linear regression (MNLR) and autoregressive integrated moving average (ARIMA) models. Bayram *et al*. (2012) investigated the feasibility of using turbidity as an indication for SSC employing regression analysis (RA) and ANN techniques. Turbidity data collected between March 2009 and February 2010, from six monitoring stations along the stream Harsit's main branch, Eastern Black Sea Basin, Turkey, was used in the study. ANNs were found to be providing acceptable results. Ghorbani *et al.* (2013) adopted the same concept of ANN and ANFIS modeling approaches in their study to examine the ability of the previously mentioned techniques in modeling the suspended sediment load using the daily river discharge data (1994 to 1995) of river Rio Chama, state of New Mexico and Colorado, USA. Their results showed that ANN is of a higher accuracy among all.

2.4 ANFIS, ANN and Conventional Methods

2.4.1 ANFIS Overview

The adaptive neuro-fuzzy inference system (ANFIS) is a hybrid system first developed by Jang (1993). The ANFIS technique integrates both ANN and the fuzzy logic principles, using ANN learning ability to generate the fuzzy IF-THEN rules that have learning capability to approximate nonlinear functions, which in turn leads to the inference. While the ANN captures the fundamental dependency in the form of the trained connection weights, the ANFIS does so by establishing fuzzy language rules (Ghorbani *et al.* 2013).

2.4.2 ANFIS and ANN Application in Estimating the Suspended Sediment Problem

Many researchers and engineers adopted the concept of ANFIS to simulate and predict the problem of suspended sediment load in various rivers, Kisi (2005) evaluated the ability of ANFIS and ANN to model the streamflow-suspended sediment relationship for the two stations of Quebrada Blanca at Jagual and Rio Valenciano near Juncos operated by the US Geological Survey. Daily time series streamflow and suspended sediment concentrations data from the 1994 and 1995 water years were used to build the various models. SRC and MLR approaches were also applied in this study, however; the comparison showed a better performance of ANFIS model among the other techniques. In 2007, Lohani, Goel, and Bhatia developed stage-discharge-sediment concentration relationships using data from two gauging sites at River Narmada in central India. Fuzzy logic, ANN and SRC methods were used in the study. Fuzzy logic had superiority in performance comparing it with the other methods employed. However, ANFIS technique was not considered in their study. Using hydro meteorological data, Cobaner, Unal, and Kisi (2009) estimated current SSC by ANFIS and ANN approaches. They used different combinations of current daily rainfall, streamflow and past daily streamflow, suspended sediment from Mad River Catchment near Arcata, USA. the potential of ANFIS method was compared with those of the three different ANN techniques and two different SRC. The results revealed that ANFIS was able to provide better performance than the other techniques in predicting the current SSC. Rajaee *et al.* (2009) examined the ability of ANN, ANFIS, MLR and SRC models for simulation of SSC using daily river discharge and SSC data belonging to Little Black River and Salt River gauging stations in the USA. The results illustrated that ANFIS model presented better performance in SSC prediction in compression to other models. Results from the study conducted by Kisi *et al.* (2009) emphasized the superiority of ANFIS among ANN and SRC techniques in predicting monthly suspended sediment. The study used monthly streamflow and suspended sediment time series data belonging to Kuylus and Salur Koprusu, in Kizilirmak Basin in Turkey. Recently, two additional studies using ANFIS and ANN approaches were conducted to simulate the suspended sediment load in the River Dalaki, Iran by Rezaei and Fereydooni (2015) and Tahmoures, Nia, and Naghiloo (2015). Whereas the former study used different combination of sediment discharge and flow discharge data for the period between (1989-2009) to predict the monthly suspended sediment load, the latter used daily stream flow and rainfall SSC with various grouping orders for ten years (1998-2008) to estimate and forecast the SSC. However, both studies concluded that ANFIS approach was capable to give better results than other approaches including ANN.

2.5 Genetic Programming (GP)

2.5.1 General Overview

Genetic programming (GP) first developed by Koza (1992), is an evolutionary computation method that needs no prior knowledge about the shape or structure of the solution to automatically solve the problem. GP is a *systematic, domain-independent* technique causing computer programs to solve problems automatically giving it a high level statement of what is expected to be done (Poli *et al.* 2008). GP involves a repeated random search for solutions by exploring the existing space of computer programs – eliminating points that are below the average quality – by applying the principle of natural evolution such as crossover, mutation and reproduction to form a new generation of points. This process is repeatedly continuing until the best solution is reached. Unlike other programs which is usually expressed in the form of lines of codes, these programs are represented in a form of a syntax trees. The syntax tree consists of nodes and links, where the nodes represent the instructions – called the functions, and the leaves, which are the terminals, represent the independent variables and random constants. Five major steps for basic type of GP are required to be specified before the operation of GP. These include the determination of (i) the terminal set; (ii) the primitive functional set; (iii) the fitness measure; (iv) controlling the run parameters; and (v) the termination criterion and method of designating the result of the run (Burke and Kendall, 2005).

Banzhaf *et al.* (1998) stated that the GP has a superiority over other soft computing techniques because of the following: it is used to examine, classify and integrate a large amount of data in computer readable form; it is used in conditions where small performance progresses are easily and regularly measured; when the correlation between the variables are poorly understood; when the available dataset is small; when it is hard to reach the ultimate result to the problem; and lastly, when conventional mathematical models cannot provide the required analytical solution.

2.5.2 GP Approach in Sediment Modeling

The GP technique has been swimmingly and extensively used as a hydrological modelling tool especially for estimating sediment yield. (Kizhisseri *et al.* 2006) employed GP approach to explore a better interrelationship between the temporal pattern of fluid field and sediment transport, using two datasets of numerical model results and Sandy Duck field data. The results from the numerical models are encouraging though GP found hard to find correlation from the field data. The time series data sets between 1977 and 1981 of suspended sediment and daily stream flow from two stations on Tongue River in Montana, USA were used by Aytek and Kişi (2008) to investigate the ability of GP technique to model the discharge-suspended sediment relationship in compare with the other conventional methods. The results suggested that the GP approach may provide a superior alternative to the SRC and MLR techniques, and that the GP approach has more practicability to use than the other available modeling systems.

The suitability of a GP in estimating sediment yield considering various meteorological and geographic features, including; river length, drainage density, yearly average rainfall, erodible area and watershed area of Arno River basin in Italy, which is disposed to frequent floods, was investigated by Garg (2011). The results of the study showed that GP can efficiently capture the trend of sediment yield, even with a limited set of data. In a study carried out by Kisi and Shiri (2012), using data sample consisted of 11 years of daily records of river discharge, precipitation, suspended sediment load and maximum, minimum and mean air temperatures, the accuracy of three different soft computing methods, ANN, ANFIS ,and Gene Expression Programming (GEP) were examined and compared to estimate daily SSC in the Eel River near Dos Rios, in California, USA. The comparison of the results indicated that GEP model performed better than the ANN and ANFIS model. Guven and Kişi (2011) modeled daily suspended sediment in the Tongue River in Montana, USA using GEP, ANN and Linear Genetic Programming (LGP) techniques. Using the same daily discharge and suspended sediment data of the study conducted by Aytek and Kişi (2008) mentioned above, the results showed that GEP has given better performance than ANN, however, LGP models were of a superiority among all other models.

2.6 Other Modelling Techniques

A comparative study between three different machine learning techniques, namely, ANN, ANFIS and coupled wavelet and neural network (WANN) and the conventional SRC method to best estimate the daily suspended sediment load using stream discharge and suspended sediment load data obtained from two gauging stations of the Flathead River at Flathead British Columbia, Montana and Santa Clara River, California, was done by Olyaie *et al.* (2015). Results of the study showed that a superior performance of the WANN method amongst the other employed techniques. Sobieszczyk *et al*. (2015) from the USGS and US Department of Interior applied regression analysis and $\log_{10}$-transformed and untransformed data approaches to model SSC using continuous turbidity and (or) streamflow data recorded between October 2011 and September 2014 for Wilson and Trask Rivers, northwestern Oregon. Turbidity and streamflow were evaluated separately using simple linear regressions and together using multiple linear regressions. At the two sites, turbidity and streamflow were extremely interrelated, so simple linear regressions were chosen. Results show that turbidity can be used as surrogate for the SSC. Shamaei and Kaedi (2016) introduced for the first time the stacking method, (i.e. a powerful machine learning technique associate results obtained from various predictive models using a meta-model based on cross validation), to predict the suspended sediment using datasets from the two stations of Rio Valenciano and Quebrada Blanca, in the USA. Stacking method was found to be a great technique to greatly improve RMSE and R^2 statistical measures if compared with the results of LGP, ANN and ANFIS found by Guven and Kişi (2011) and Kisi (2005) for the same stations. Adaptive neuro-fuzzy embedded fuzzy c-means clustering (ANFIS-FCM) approach to predict SSC was first employed by Kisi and Zounemat-Kermani (2016). Using daily discharge and SSC data from two stations, Muddy Creek near Vaughn and Muddy Creek at Vaughn, operated by the USGS three other different models were built, namely, ANN, ANFIS and SRC. All the four models were then compared with each other using different statistical measures to select the best simulative model. ANFIS-FCM model performed the best in estimating SSC.

The present study is aiming to compare the accuracy of two different soft computing modelling techniques, i.e. ANFIS and ANN and three conventional methods of SRC, SLR and MLR in estimating the daily SSC using several statistical indicators including; RMSE, MAE and NSE. Different combinations of models were built based on the available weekly datasets of simultaneously measured SSC, stream discharge, water temperature and water electric conductivity for River Thames at Byron station, London, ON, Canada.

CHAPTER 3
METHODOLOGY, AREA OF STUDY AND DATA COLLECTION

3.1 Conventional Approaches

This section presents the commonly used classical approaches that have been conducted by several researchers in the literature to simulate the suspended sediment concentration problem in numerous streams.

3.1.1 Sediment Rating Curve

Large numbers of hydrologists and engineers have used the rating-curve approach to estimate the suspended sediment concentrations due to the lack of labor force or automatic devices for frequent sampling, and laboratory analysis facilities to analyze various samples. The rating curve method basically relates the sediment concentration to the discharge in a form of a graph or equation. It can then be used to simulate the relationship between SSC and Q utilizing the documented streamflow and SSC data.

The first documented example on the use of the sediment rating curves approach is traced back to a study conducted by Campbell and Bauder (1940). They developed a silt rating curve by plotting daily suspended sediment load against daily stream discharge on logarithmic scale for the River Red in Texas, USA (Kisi *et al.* 2006).

As it was mentioned in the section 2.2.2 of this study, the relationship between SSC and Q is usually represented as a power function (equation (1)). The values of the constants a and b, which are different and unique for every river, creek, tributary, and basically every water stream, can be obtained by plotting the relationship between $log\, Q$ on the X-axis and $log\, SSC$ on the Y-axis. From this linear relationship, gradient, i.e. slope of the line, represents b value, while log a is the y-intercept.

3.1.2 Linear Regression

Linear regression is one of the basic types of regression and is the one which is widely used in the prediction analysis. Regression analysis in general is a statistical methodology aiming to find a functional relationship (model or equation) among dependent and independent variables. Regression can be univariate means that only one dependent variable is being studied, while it is called multivariate when dealing with two or more dependent variables. In the case of studying the SSC in rivers, the study only dealing with one dependent variable which is always the SSC. By fitting a linear equation to the data, the univariate linear regression analysis attempts to model the relationship among variables. In the case where there is only one independent variable used to predict the response variable the regression called simple linear regression. When there is more than one independent variable, the system is called multiple linear regression and in this case the linear fitting is attempting by providing the effect on the dependent variable from a one-unit change in the corresponding independent variable, holding all other independent variables constant. The goal of linear regression is to adjust the values of slope and intercept to find the line that best predicts SSC from Q. More precisely, the goal of regression is to minimize the sum of the squares of the vertical distances of the points from the line. This can be achieved by the Least Squares Method. In linear regression to determine the best-fitting straight line several assumptions are made in the calculation of inferential statistics; these assumptions are listed below:

- The relationship between the variables is linear.
- The variance around the regression line is the same for all the values of the independent variable (Q).
- Normal distribution of the errors of predictions, i.e. the deviations from the regression line are normally distributed.
- The residuals are not correlated with one another.
- The residuals are not related to the independent variables.

In this study Microsoft Excel 2016 MSO spreadsheet was used to model the regression part of the study.

3.1.2.1 Simple Linear Regression (SLR)

The simple linear regression is expressed by the following form the equation:

$$Y = \beta_0 + \beta_1 X + \varepsilon \tag{3}$$

Where:

- Y = the dependent variable (output),
- X = the independent variable (input),
- β_0 and β_1 = the regression coefficients or regression parameters, and
- ε = an error to account for the difference between the predicted data using Eq. (3) and measured data.

The predicted value form of Eq. (3) is

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \tag{4}$$

Where:

- $\hat{Y}$ "read Y hat" = the fitted or predicted value, and
- $\hat{\beta}$ = estimates of the regression coefficients.

The regression coefficient in this case can be found after plotting the linear relationship between SSC and Q. Values of $\hat{\beta}_0$ and $\hat{\beta}_1$correspond to the Y-intercept and the slope of that fitted-line, respectively.

3.1.2.2 Multiple Linear Regression (MLR)

The multiple linear regression, or univariate multiple regression, is the generalization of the simple linear regression model. The model in multiple linear regression includes more than one input variable. So if it is believed that the dependent variable Y is effected by *n* number of independent variables $X_1, X_2, \ldots, X_n$, then the regression equation of Y can be represented as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \tag{5}$$

Where:

- Y = the dependent variable (output),
- $X_1, X_2, \ldots, X_n$ = the independent variables (inputs) with n number of observations,
- $\beta_0, \beta_{1, \ldots,} \beta_n$ = the regression coefficients or regression parameters, and
- ε = an error to account for the difference between the predicted data. The predicted value form of Eq. (5) is

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_n X_n \tag{6}$$

Where:

- $\hat{Y}$ = the predicted value of the variable Y when the independent variables are represented by the values $X_1, X_2, \ldots, X_n$.

The estimated regression coefficients $\beta_0, \beta_{1, \ldots,} \beta_n$ are evaluated similar to SLR by minimizing the sum of the e_{yi} distances of observation points from the plane expressed by the regression equation, as follows;

$$\sum_{i=1}^{n} e_{yi}^2 = \sum_{i=1}^{n} (\widehat{Y_1} - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \cdots - \hat{\beta}_n X_{ni})^2 \tag{7}$$

In this study values of $\beta_0, \beta_{1, \ldots,} \beta_n$ are determined using the Microsoft Excel 2016 spreadsheet

3.2 Machine Learning Approaches

This section shows the two most widely used artificial intelligence approaches which have been utilized by many scientists and researchers in the literature to simulate the suspended sediment concentration problem in numerous streams. It is worthy to note that these two machine learning techniques have been used in this study.

3.2.1 Artificial Neural Networks

Artificial Neural networks (ANN) are kind of statistical modelling approaches designed basically to mimic the human brain in its ability to arbitrate inputs and eventually reach to conclusion(s). These networks are designed to learn from the provided data “training sets” and then estimate the parameters of some populations. Applications of neural networks can be found in data modeling, system optimization and statistical analysis. Fields like econometrics, engineering, psychology and physics use neural networks as the statistical tools (Lewis-Beck *et al.*, 2004). ANN is one of the most significant strengths of the Artificial Intelligence (AI) techniques that hydrologists used in the last couple decades, which helped researchers to handle all data types, and predict different nonlinear phenomena.

ANN power lays deep in its ability to approximate arbitrary continuous functions between the input(s) and output(s) based on a set of given examples. This ability is gained during the stage of training or sometimes called learning. As this ability is obtained, they are known as truly adaptive systems, which do not require any previous knowledge about the nature of relationships between parameters (Afaghi *et al.*, 2001)

The structure of the ANN consists of simple units called neurons. In a network, neurons are connected through weighted connections. Learning process inside the network is achieved throughout adjusting those weights (Lipták, 2002). Networks are usually arranged in the form of layers (columns or rows) where the first layer represents the input and the last one corresponds to the output. Intermediate layers in between input and output layers are known as hidden layers. therefore, the very basic form of the ANN is where the input is processed to predict the output, single-layer or multi-layer network.

Analyzing data process starts with feeding inputs to the first layer neurons and then further modification is conducted by propagation of data to the neurons of the second layer. Then results are transferred to the next layer and so on until they reach the output layer. The objective of ANN is to learn by training and to eventually determine a logical connection between input and output patterns, or to analyze, or to find the structure of the input patterns. When a network is fed with data, network training achieved through the modification of the connection weights between units. This process is similar to interpreting the value of the connections between units as parameters from statistical point of view. The training process identifies the “algorithm” used to find these parameters (Lewis-Beck *et al.*, 2004)

3.2.1.1 ANN Structure

Neural network is a black box approach in which the input variables are processed to generate outputs. Generally, the neural network composed of three main components, called layers, as shown in Figure 3-1, and known as follows:

i. Input Layer

This layer is responsible for receiving information from an external source. The input layer consists of neurons equal to the number of inputs. Inputs could be measured data from an external environment or signals from sensory systems which passes through the network during the data processing phase.

ii. Hidden or Intermediate Layers

These layers do not receive data from the outside environment as it is the case for the input layer and it is only connected to the layers within the network. Neurons of these layers hiddenly perform most of the internal processing work and are responsible for acquiring patterns associated with the process or network being analyzed. Number of neurons in these layers are variable and is set during the writing algorithm phase.

iii. Output Layer

Number of neurons in this layer is the same as for those of the system output. This layer is responsible for producing and presenting the final outputs network that resulted from neurons of the previously processed layers.

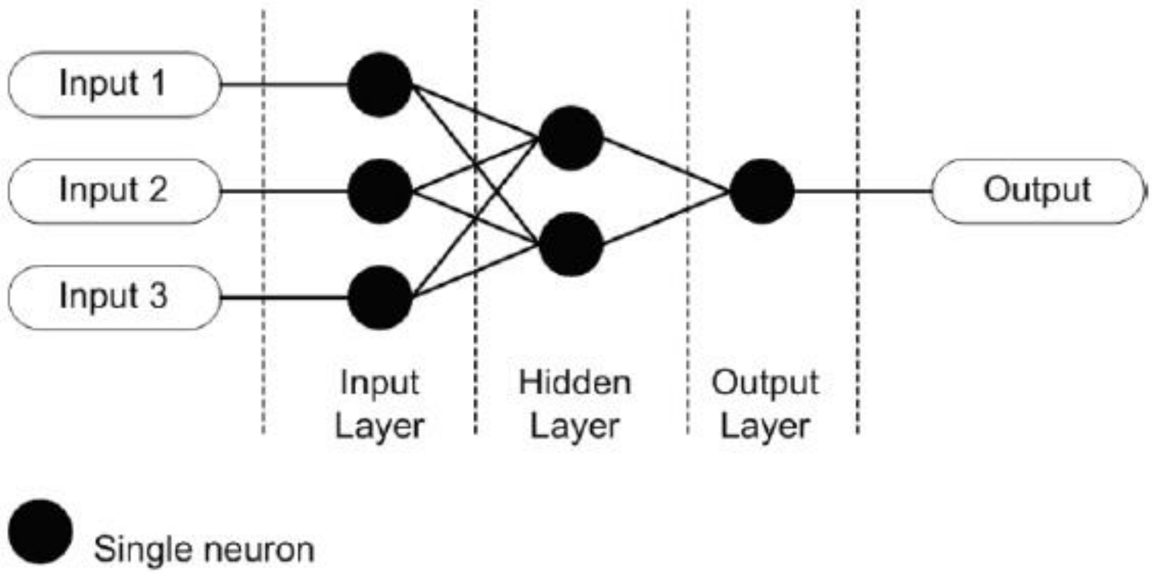


Figure 3-1 Example of simple artificial neural network (Krenker *et al*. 2011)

3.2.1.2 Artificial Neuron

Artificial neuron is the main elementary unit of any artificial neural network. It is designed to resemble the structure and the function of a biological neuron which is a biological neural network's key building block that includes; the brain, spinal cord and peripheral ganglia. Equivalence in structural composition and functionalities are shown in Figure 3-3 and Figure 3-3 where Figure 3-2 represents the biological neuron including its components (dendrites, nucleus, axon, …) while Figure 3-3 represents the artificial neuron including inputs, weights, activation (transfer) function, bias and outputs

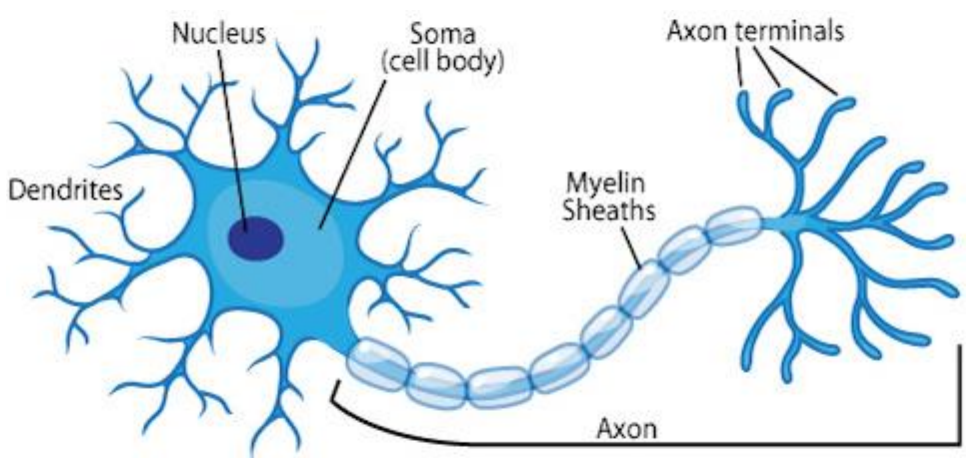


Figure 3-2 Biological Neuron Design (https://askabiologist.asu.edu/neuron-anatomy)

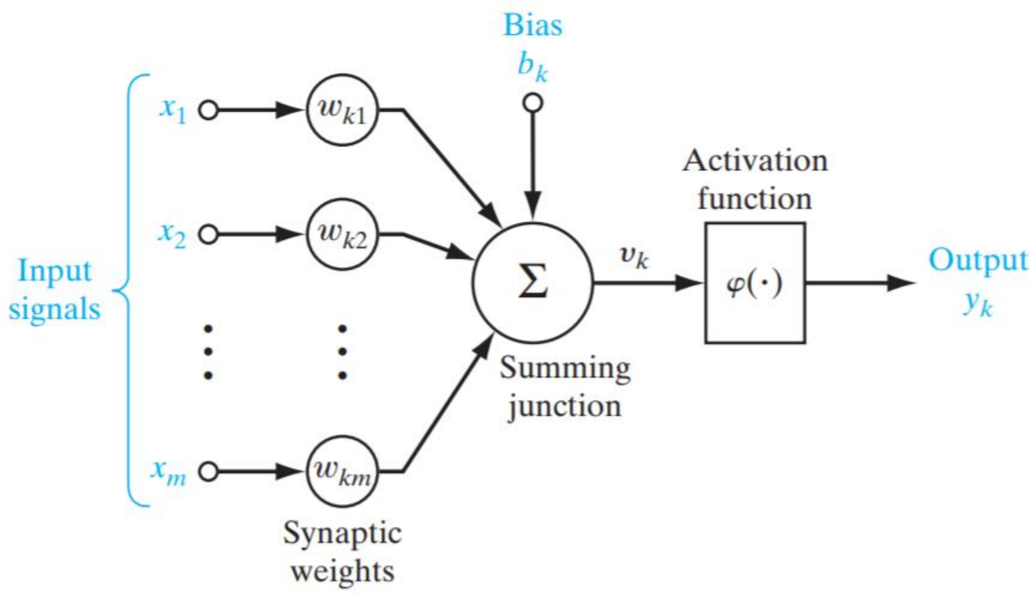


Figure 3-3 Artificial Neuron Design (Haykin, 2008)

For the biological neuron case, the information is received by neurons through dendrite, then the cell body, namely soma processes the information and transfers it on via axon. Where in the case of the artificial neuron the information is received by the artificial neuron's body by inputs that are then weighted (individually multiplied with a weight) then the body sums up the weighted inputs and bias and then processes the sum with the activation function also known as transfer function. At the final phase, the artificial neuron runs the processed information via output(s). The simple form of a mathematical representation for an artificial neuron can be denoted as follows:

$$y(k) = F\left(\sum_{i=1}^{n} w_i(k).x_i(k) + b\right) \tag{7}$$

Where:

- $x_i(k)$ = input value in discrete time k where i goes from 0 to m,
- $w_i(k)$= weight value in discrete time k where i goes from 0 to m,
- b = bias,
- F = a transfer function,
- $y_i(k)$ = output value in discrete time k.

For this equation, the main unknown variable is the transfer function. Transfer function outlines the properties of the artificial neuron. Generally, the various set of functions that the transfer function can be chosen out of are described below:

- Step function

It is a binary function where the output can have only two probable values, either zero or one for instance as illustrated in Figure 3-4. That if the input value for a specific threshold was met then the resulted output can be in one value and the resulted output will be different if the input value for that specific threshold was not met. This type of transfer function when used in artificial neuron then this neuron is called perceptron. This function can be mathematically defined with the following equation:

$$y = \begin{cases} 1 \; if \; w_i x_i \; \geq threshold \\ 0 \; if \; w_i x_i \; < threshold \end{cases} \tag{8}$$

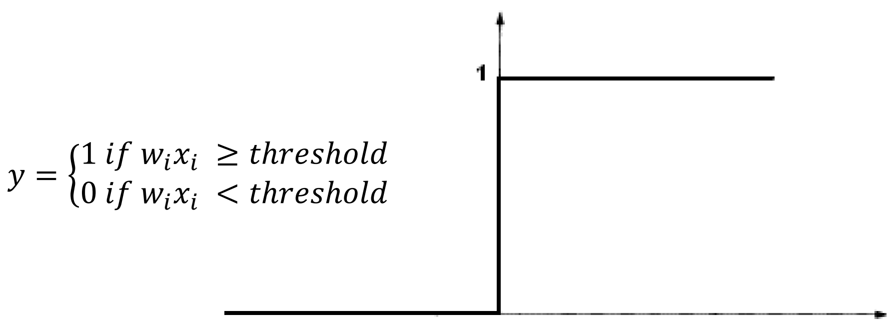
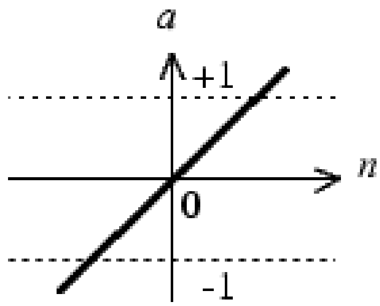


Figure 3-4 Step transfer function

- Linear function

In this case as shown in Figure 3-5, the artificial neuron is applying the simple linear transformation over the sum of weighted inputs and bias. This function is mostly used in the input layer of the artificial network.



$a = Purelin(n)$

Figure 3-5 Linear (Purelin) transfer function

The neuron's output is calculated using the simple equation taking the following form:

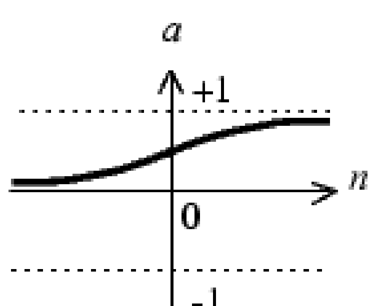
$$a(n) = \alpha x \tag{9}$$

- Non-linear function

The most well-known type of non-linear functions is the sigmoid function which produces an "S" shape function. While the perceptron outputs discrete 0 or 1, the sigmoid neuron outputs a more smooth or continuous range between 0 and 1. The most two well-known classes of the sigmoid function are:

i. Log-sigmoid Transfer Function (LOGSIG)

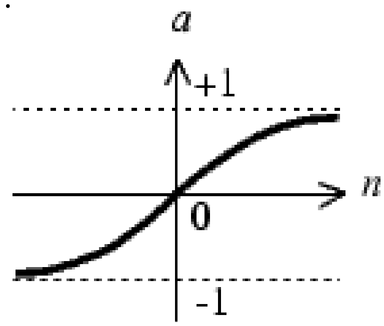
This type of transfer function (shown in Figure 3-6) takes the input (that could have any value falls in the range between plus and minus infinity) and limits the output into the range of 0 to 1. Since this function is differentiable, it is most commonly used in multilayer networks that are trained using the backpropagation algorithm (da Silva *et al.* 2017).



$$a = Logsig(n) = \frac{1}{1 + e^{-n}}$$

Figure 3-6 Log-Sigmoid transfer function *ii.* Hyperbolic Tangent Transfer Function (TANSIG)

The output of this type of a function ranges from -1 to +1, as it can be seen in the Figure 3-7.



$$a = Tansig(n) = \frac{2}{1 + e^{-2n}} - 1$$

Figure 3-7 Tan-Sigmoid transfer function

3.2.1.3 Composition of ANN

The artificial neural network is a result of combining two or more artificial neurons. Unlike the single artificial neuron which has nearly no importance in case of solving a real-life problem, the artificial neural network has the capability to solve it. Moreover, the artificial neural networks have also the ability to solve complicated real-life problems through their powerful processing ability of the information in their artificial neurons in a non-linear, distributed, parallel and local way.

The topology (architecture) of the artificial neural network is the way of how single neurons are connected to each other. This interconnection can be performed in several ways which lead to different potential topologies. Predominately, these numerous possible topologies are categorized into two basic classes, namely, feed-forward topology in which the information transforms from inputs to outputs in one direction only, and recurrent topology in which some of the information can transform to the opposite direction as well and not limited to only one direction. Figure 3-8 illustrate the difference between the main two topologies, the left side represents the feed-forward neural network type of topology while the right-hand side one shows the recurrent neural network.

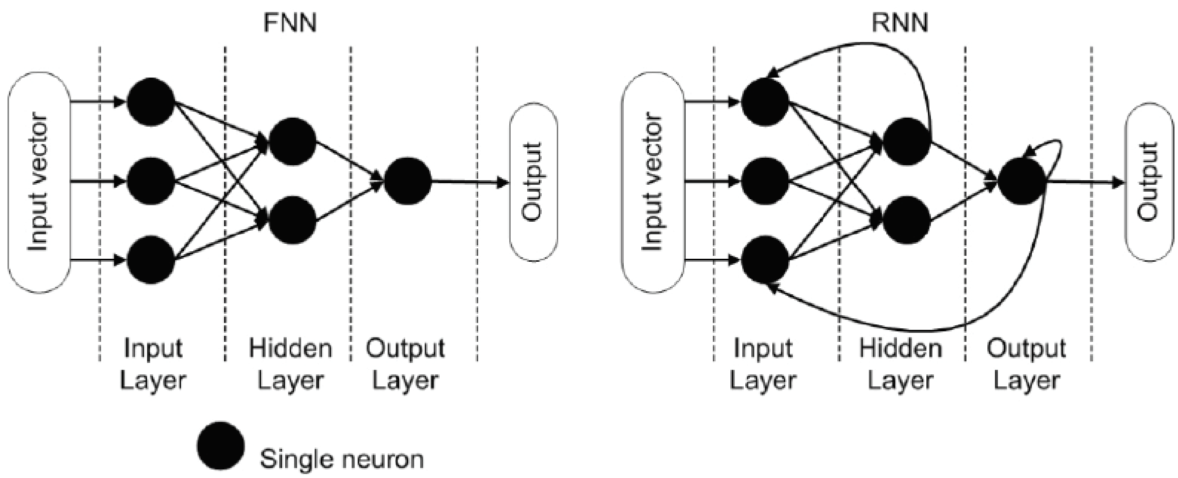


Figure 3-8 Feed-forward (FNN) and recurrent (RNN) topologies of ANN (Krenker *et al*. 2011)

Choosing and building topology of an artificial neural network is only half of the task towards using it in solving any given problem. Similar to the biological neural networks and how they need to process given inputs to give a fit response, the artificial neural networks go through the similar procedure. To achieve such a task, the modeler has to learn the artificial neural network using one of the following learning algorithms; supervised, un-supervised or reinforcement learning.

3.2.1.4 Feed-forward Artificial Neural Networks

In this type of topology, the information always flows in a single direction (unidirectional) from input to output and has no back-loops. No restrictions on number of layers or the transfer function type used in single artificial neural or the number of connections among the individual artificial neurons. For the purpose of analytical description, the simple multi-layer feed-forward artificial neural network is shown in the Figure 3-9 and is represented by the following equations:

$$\begin{aligned} n_1 &= F_1(w_1 x_1 + b_1) \\ n_2 &= F_2(w_2 x_2 + b_2) \\ n_3 &= F_3(w_3 x_3 + b_3) \\ n_4 &= F_4(w_4 x_4 + b_4) \end{aligned} \tag{10}$$

$$m_1 = F_4(q_1 n_1 + q_2 n_2 + b_4)$$
$$m_2 = F_5(q_3 n_3 + q_4 n_4 + b_5) \tag{11}$$
$$y = F_6(r_1 m_1 + r_2 m_2 + b_6)$$

$$y = \begin{pmatrix} r_1(F_4[q_1 F_1(w_1 x_1 + b_1) + q_2 F_2(w_2 x_2 + b_2)] + b_4) + \ \dots \\ \dots + \ r_2(F_5[q_3 F_2(w_2 x_2 + b_2) + q_4 F_3(w_3 x_3 + b_3)] + b_5) + b_6 \end{pmatrix} \tag{12}$$

As seen in Figure 3-9 and the corresponding analytical representation with the set of equations (10), (11), and (12) the simple feed-forward artificial neural network would lead to a long mathematical equation making its solving by hand is impractical. Despite that for any given complex artificial neural network, the mathematical representation can be used, however; computers and technical software (MATLAB r2016b was used in this study) are used to assist in building and optimizing any type of artificial neural network.

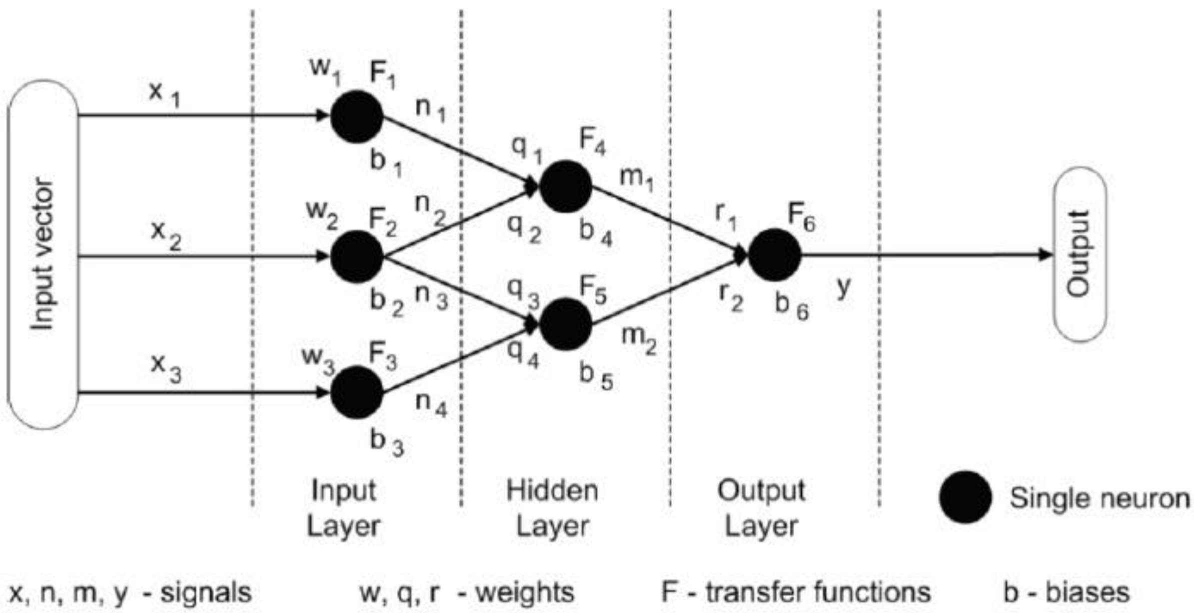


Figure 3-9 Multi-layer feed-forward artificial neural network (Krenker *et al.* 2011)

3.2.1.5 Recurrent Artificial Neural Networks

An artificial neural network with the recurrent topology has the same concept of the feed-forward type topology with no restrictions in regards with the back-looping procedure. For this the input information is no longer transferred only in a unidirectional way but it is also transmitted backwards. In other words, the outputs of the neurons are used as a feedback inputs for other neurons, this gives this type of networks the dynamic feature, meaning that they can be applied on time-series systems. The most basic topology of the recurrent artificial neural network is the fully recurrent artificial network shown below in Figure 3-10, however; there are other special cases of the recurrent artificial networks such as Hopfield, Elman, Jordan and bi-directional.

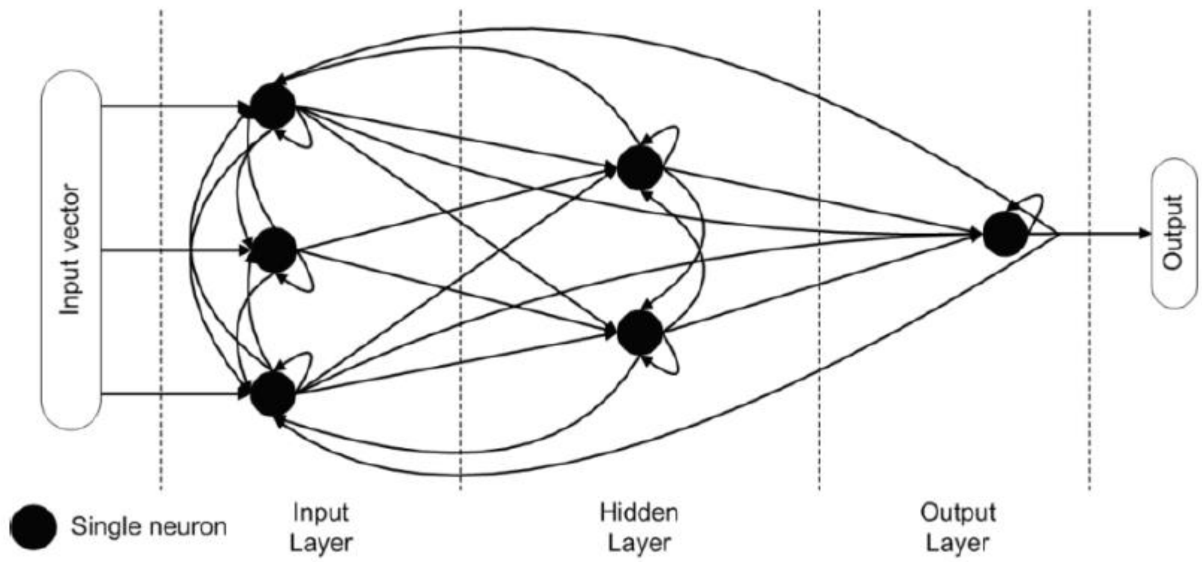


Figure 3-10 Fully recurrent artificial neural network (Krenker *et al.* 2011)

3.2.1.6 Learning Models

As discussed in the section 3.2.2.3 of this chapter that there are three main learning patterns, namely; supervised learning, unsupervised learning and reinforcement learning. Every single learning model has many training algorithms, and all can be applied by any given type of artificial neural network.

i. Supervised Learning

It is documented that Donald Hebb proposed the first supervised leaning strategy inspired by neurophysiological observations (Hebb, 1949)

The desired outputs for a given set of input variable(s) must be available for this type of leaning approach; in other words, each training sample is composed of the input variables and their corresponding outputs. Therefore, the supervised learning strategy necessitates a table of input and output data. The neural structure therefore will be learning from these given set of data. This is then limits the application of this learning method to be depended only on the availability of that input/output table, and the network will therefore respond to what it learned to be correct from the fed set of "teaching" data.

The weights and bias (thresholds) of the network are in a continuous adjustment by applying the comparative actions, accomplished by the learning algorithmic rules, thus the difference between calculated and measured data is always under supervision of the modeler. The network is given the title "trained" when the difference between calculated and measured data is within a satisfactory range. It is noteworthy to mention, that this is the type of learning approach was considered in this study.

ii. Unsupervised Learning

Contrary to the supervised learning approach, the application of an algorithm based on unsupervised learning technique does not need any prior information in regards with the targeted outputs. The application of unsupervised learning is commonly seen in the problems of assessment, such as statistical modelling, compression, filtering, blind source separation and clustering. In this type of learning the concern is to find how data is organized. Only unlabeled examples are given to the artificial neural network with unsupervised learning type and this is the key difference between it and the supervised learning and reinforcement learning. One of the most common forms of unsupervised learning is clustering that aims to categorize data in different clusters based on their similarity.

iii. Reinforcement Learning

Reinforcement learning is a type of a machine learning approach where parameters of an artificial neural network are set. This case is useful where data is not given so it will be generated by interactions with the environment. Reinforcement learning is often utilized as a part of the artificial neural network's overall learning algorithm.

Models built utilizing the reinforcement learning are considered to be similar to a certain degree to those using supervised learning approach, because both continuously examine the difference between the network's response and the desired output (da Silva *et al.* 2017). The internal neural elements are modified by the learning algorithm used on the reinforcement learning depending on whatever quantitative or qualitative data derived by the interaction with the environment being assessed. The performance evaluation can therefore be determined using this acquired information.

3.2.1.7 Back Propagation Training Algorithm

The back-propagation of errors technique is one of the most popular artificial neural network's algorithms. To train a network using this algorithm, generally two phases must take place; first phase, namely, "forward phase", happens when the inputs are presented and propagated forward through the network to calculate the output for every processing element. However, in the second stage, also known as "backward phase", the recurrent difference calculation (of the forward phase) is performed in a backward direction. The algorithm will stop training when the error value of the error function become adequately insignificant.

Rojas (2005) claimed that the back-propagation algorithm could be broken down into four main steps. Next after selecting the weights of the network randomly, the needed corrections are then calculated using the back-propagation algorithm. The four main steps forming the algorithm are as follows:

i. Feed-forward computation;

ii. Backpropagation to the output layer;

iii. Backpropagation to the hidden layer;

iv. Weight updates.

This technique is basically a gradient descent method to process the total squared error of the output calculated by the net and make it negligible (minimum). Back-propagation is a systematic method for training multiple artificial systems. Back-error propagation is the most broadly used of the neural network models and has been applied successfully in applications in a wide range of areas. Back-propagation network consists of layers, each layer entirely linked to the layers underneath and above as shown in Figure 3-11.

In back-propagation algorithm, same as every other type of algorithms, middle and output layers use activation functions. Typically, sigmoid activation functions are among the most used; where the output of the network will be between 0 and 1. However; Gaussian distribution can be used as an activation function as well. After this process computed output value compared with the expected output value and the distance between them are taken as an error to back propagate, and that is why it is called back-propagation. The predetermined error function is:

$$E = \sum_{n=1}^{N} (t_n - z_n)^2 \tag{10}$$

Where:

- E = the total error, and
- t and z = the calculated and measured outputs for the input n, respectively.

The back-propagation system takes the following form:

The derivative of the error with respect to the weight connecting i to j is;

$$\frac{\partial E}{\partial W_{ij}} = \delta_j y_i \tag{11}$$

To change weight (W) from unit i to unit j by;

$$\Delta W_{ij} = -\eta \delta_j y_i \tag{12}$$

Where:

- η = the learning rate (η > 0),
- δ_j = the error for unit j, and
- y_i = the input for unit i.

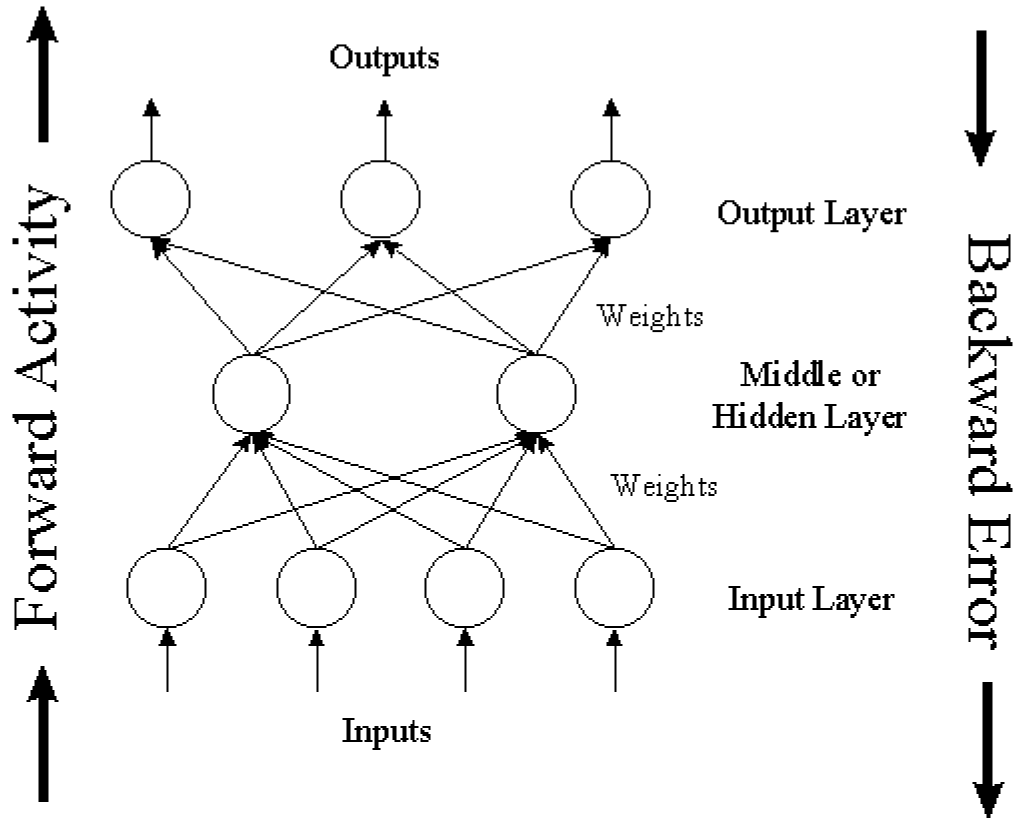


Figure 3-11 Architecture of MLP feed-forward ANN (Nastos *et al*. 2011)

Every middle layer node employs an activation function. In this back-propagation process, a sigmoid (LOG type) function can will be used because sigmoid function can easily be calculated and differentiated.

$$y = f(a) = \frac{1}{1 + e^{-a}} \tag{13}$$

Derivative of Eq.(14);

$$f'(a) = f(a)\big(1 - f(a)\big) \tag{14}$$

Each input variable is computed in a weighted form;

$$y(x) = w^T f(x) \tag{15}$$

It is necessary to calculate the error term for both output and middle units as follows;

For input units

$$\delta_k = (y_k - y_{target}) \tag{16}$$

For hidden units

$$\delta_j = y_j(1 - y_j)\sum_k W_{ij}\delta_k \tag{17}$$

Then to minimize the error, new weight values are driven in the opposite direction. The learning rate determines the amount of update in the specified direction. It is noteworthy that this method is the one employed in this study.

3.2.2 Adaptive Neuro-Fuzzy Inference System

A neuro-fuzzy technique called Adaptive network based fuzzy inference system or semantically equivalent to adaptive neuro fuzzy inference system (ANFIS) (J.-S. R. Jang & Sun, 1995; J. R. Jang *et al.* 1997) has been used as a modeling tool in the present study. ANFIS may also refer to Adaptive network based fuzzy inference system, and it is a neuro fuzzy technique where the fusion is made between the neural network and the fuzzy inference system. In ANFIS the parameters can be estimated in such a way that both the Sugeno and Tsukamoto fuzzy models are represented by the ANFIS architecture (J. R. Jang *et al*., 1997). ANFIS technique comprises of a hybrid system of fuzzy logic and neural network technique in order to have better results for systems possessing nonlinear behavior. The fuzzy logic takes into account the imprecision and uncertainty of the system that is being modeled while the neural network gives it a sense of adaptability. Using this hybrid method, at first an initial fuzzy model along with its input variables are derived with the help of the rules extracted from the input output data of the system that is being modeled. Next the neural network is used to fine tune the rules of the initial fuzzy model to produce the final ANFIS model of the system.

3.2.2.1 Fuzzy Inference system

The adaptive neuro-fuzzy system can be defined as a fuzzy inference system (FIS) equipped with a training algorithm (Bentaher & Elmazoghi, 2013). Suparta and Alhasa (2016) stated that a FIS was built on three core units; (i) IF-THEN fuzzy rule base, (ii) membership functions to be used in the fuzzy rules and (iii) reasoning fuzzy inference techniques from basic rules to get the output. Comprehensive architecture of the FIS is illustrated in Figure 3-12. The inputs will be fuzzified from actual values using the fuzzification process based on each one's membership function, where the fuzzy value ranges between 0 and 1. Basic rules (rule-base) and databases are usually known as knowledge base, and that both of them are main components in the decision-making process.

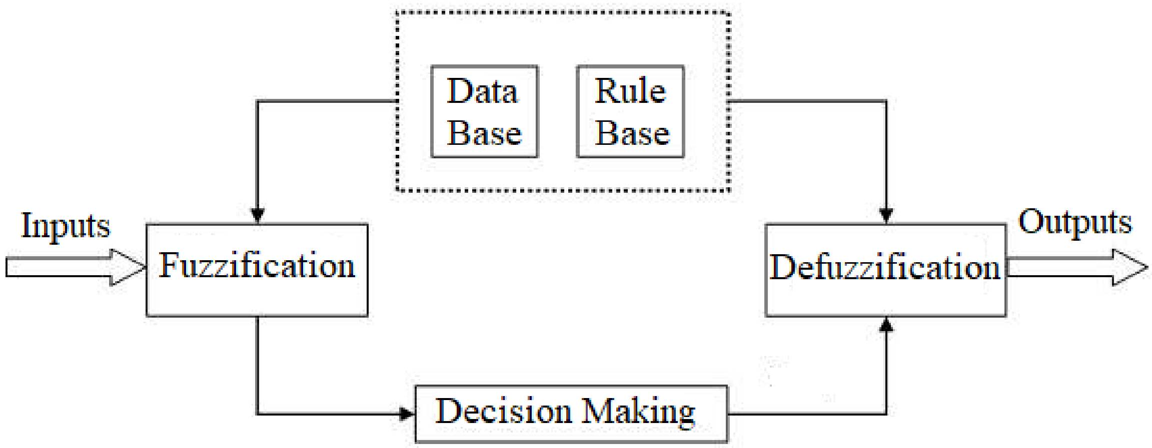


Figure 3-12 Architecture components of a fuzzy inference system

Database unit typically contains information on fuzzy sets parameter along with a function that outlines each existing linguistic variable. While developing the database unit, number of the linguistic values to represent each corresponding linguistic term is determined, the membership function is established in this phase too. Rule base includes fuzzy logic operators and IF-THEN conditional statements, it is generated either by the modeler (human) or automatically from the environment. There are several types of FIS, including Takagi-Sugeno, Mamdani and Tsukamoto. However; Takagi-Sugeno model is found to be the one widely used in the application of ANFIS technique (Suparta & Alhasa, 2016).

3.2.2.2 Adaptive Network

Adaptive network is a feedforward neural network with multiple layers as shown in Figure 3-13. It uses the supervised learning algorithm for the learning process. Moreover, the adaptive network consists of various adaptive nodes connected directly without any weight value in-between them. The output is then dependable on the signals and parameters of each node, and those nodes have different purposes and task.

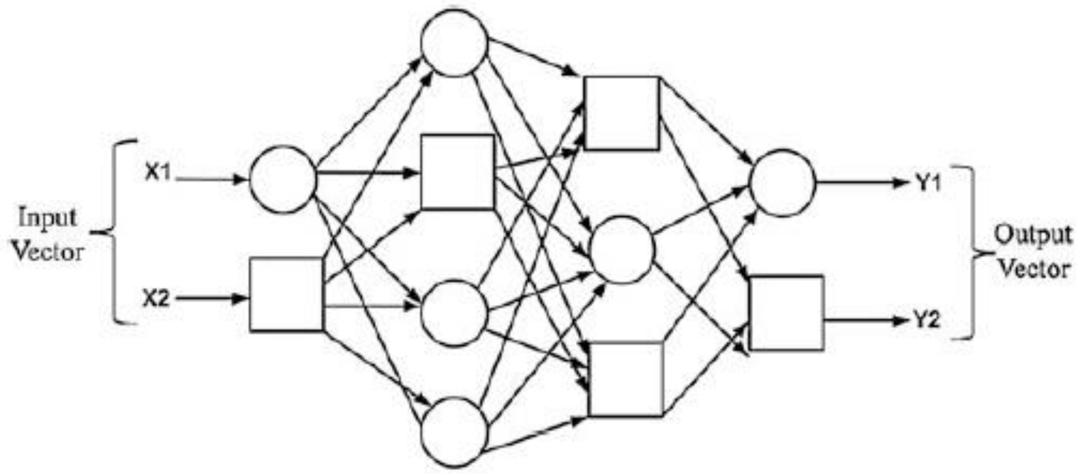


Figure 3-13 Architecture of an adaptive network

3.2.2.3 ANFIS Structure

The structure of the ANFIS could be simply described as an adaptive network that uses type supervised learning as a learning algorithm and has a function similar to the model of Takagi-Sugeno fuzzy inference system. In other words, ANFIS is a hybrid system integrating the learning capabilities of ANN and knowledge representation and inference abilities of fuzzy logic that could self modify their membership function to achieve a desired performance. The structure of fuzzy reasoning mechanism for Takagi-Sugeno model is shown in the Figure 3-14, and the corresponding ANFIS scheme is illustrated in Figure 3-15. It is assumed that the system has two inputs x and y, and one output f. IF-THEN rules are two rules for the Takagi-Sugeno model, as follows:

Rule 1 = If x is A_1 and y is B_1 Then $f_1=p_1x+q_1x+r_1$

Rule 2 = If x is A_2 and y is B_2 Then $f_2=p_2x+q_2x+r_2$

Where:

- A_1, A_2 and B_1, B_2 = the membership functions of inputs x and y, and
- p_1, q_1, r_1 and p_2, q_2, r_2 = linear parameters of Takagi-Sugeno fuzzy inference model.

The corresponding ANFIS architecture shown in Figure 3-15 appears to have five layers. Layer no.1 and layer no. 4 contain adaptive nodes (rectangular), whereas the rest contain fixed nodes (circular). J. R. Jang *et al.* (1997) provided a brief explanation of each layer as follows:

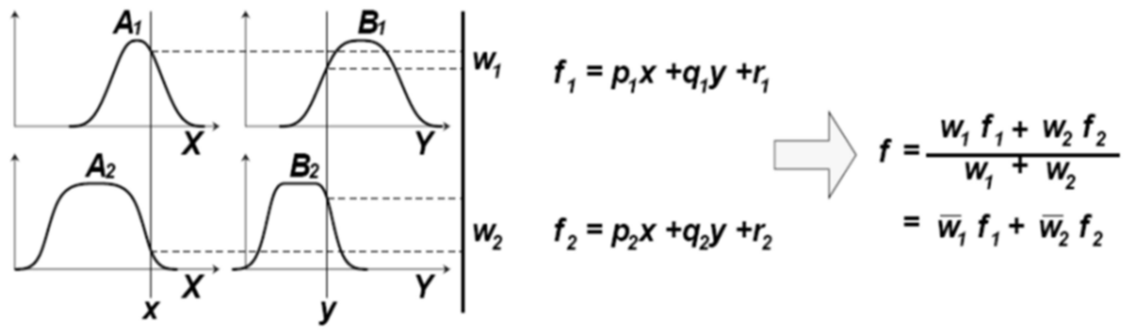


Figure 3-14 Reasoning from a two-input first-order Sugeno fuzzy model with two rule (Foroozesh *et al.* 2013)

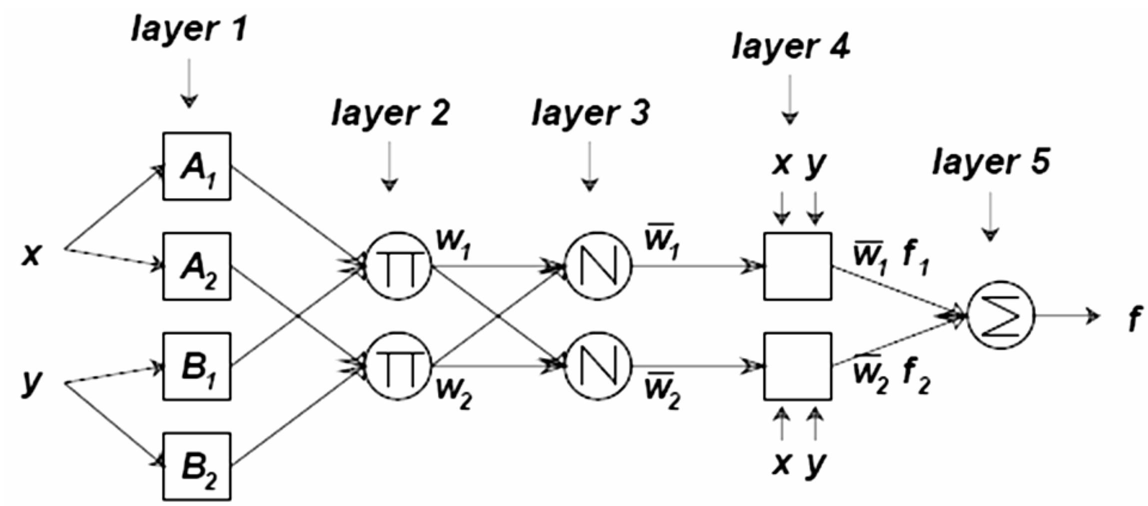


Figure 3-15 ANFIS architecture corresponding to Figure 3-13 (Foroozesh *et al.*, 2013)

Layer 1: Every node i in this layer is an adaptive node with a node function. The degree of membership value which is given by the input of the membership functions can be derived from the output from each node. For instance, the membership function can be Gaussian or generalized bell membership function (Eq. 19) and (Eq. 20), respectively. Note that there are several other types of the membership function, including but not limited to;

triangular membership function, trapezoidal membership function, Pi-shaped curve membership function and sigmoid curve membership function.

$$\mu_{Ai}(x) = exp\left[-\left(\frac{x - c_i}{2a_i}\right)^2\right] \tag{18}$$

$$\mu_{Ai}(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b}} \tag{19}$$

$$O_{1,i} = \mu_{Ai}(x), \qquad i = 1, 2 \tag{20}$$

$$O_{1,i} = \mu_{Bi-2}(y), \quad i = 3, 4 \tag{21}$$

Where:

- μ_{Ai} and μ_{Bi-2} = the degree of membership functions for the fuzzy sets A_i and Bi, respectively, and
- (a_i, b_i, c_i) are the parameters of a membership function that can change the shape of the membership function. The parameters in this layer are typically referred to as the premise parameters.

Layer 2: Nodes in this layer are fixed (non-adaptive), and the circle node is marked as Π. The output node is the result of multiplying of signal coming into the node and delivered to the next node. Each node in this layer represents the firing strength for each rule. In the second layer, the T-norm operator with general performance, such as the AND, is applied to obtain the output.

$$O_{2,i} = w_i = \mu_{Ai}(x) * \mu_{Bi}(y), \quad i = 1, 2 \tag{22}$$

Where:

- w_i= the output that represents the firing strength of each rule.

Layer 3: Nodes in this layer are fixed (non-adaptive), and the circle node is marked as N. Each node is a calculation of the ratio between the *i*-th rules firing strength and the sum of all rules' firing strengths. This result is known as the normalized firing strength.

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{23}$$

Layer 4: Every node in this layer is an adaptive node to an output, with a node function defined as

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \tag{24}$$

Where:

- $\overline{w}_i$ = the normalized firing strength from the previous layer (third layer), and
- $(p_i x + q_i y + r_i)$ = a parameter in the node. The parameters in this layer are referred to as consequent parameters.

Layer 5: The single node in this layer is a fixed or non-adaptive node that computes the overall output as the summation of all incoming signals from the previous node. In this layer, a circle node is labeled as Σ.

$$O_{5,i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{25}$$

3.2.2.4 ANFIS Hybrid Learning Algorithm

Hybrid learning algorithm is briefly discussed in this subsection. The learning algorithm uses two-passes learning cycle; forward pass and backward pass. During the forward pass (forward path), the parameters of the premises in the first layer should be fixed and the recursive least square estimator (RLSE) method is applied to modify the consequent parameter of the fourth layer. Hence the consequent parameters are linear, therefore the RSLE scheme could be utilized to fast process the convergence rate in hybrid learning technique. Subsequently, after obtaining the consequent parameters, passing back process of the input data to the adaptive network input unit takes place to generate a new output which will be compared with the measured ones.

Where in the backward pass step, the consequent parameters are set to a steady state. Backpropagation of the error that resulted from the comparison between the actual and computed outputs is generated to the first layer. At the same time, premises parameters in the first layer are updated using gradient descent or back-propagation learning systems. One level of hybrid learning is named epochs. It is noteworthy to mention that this is the learning method used to train various FIS models. A summary of the ANFIS hybrid learning procedure is described in Table 3-1.

Table 3-1 ANFIS hybrid learning process

	Forward Pass	Backward Pass
Premise parameters	Fixed	Gradient descent
Consequent parameters	recursive least square estimator	Fixed
Signals	Node output	Error rate

3.2.3 Study Area and Datasets

The study area, shown in Figure 3-16, comprises the watershed of River-Bend that is one of the 28 Upper Thames River basin's watersheds. Its total area is 5830 ha, approximately 58 km², and is located at the most downstream of the Upper Thames River basin that covers an area of 3362 km², so all water coming from the upstream watersheds passes through this area. The watershed is located within three municipalities; London (50%, 29 sq. km), Middlesex Centre (31%, 18 sq. km), Strathroy-Caradoc (19%, 11 sq. km) and it includes nine significant natural sites; four provincially significant wetlands and five significant natural areas. It also contains four other watercourses besides the Thames, namely; Van Hecke, GM Ireland, Kelly, and Stanton.

3.2.3.1 Soil Type and Land Use

Agricultural activities form 44% of the total land use activities of the area, while urban activities compose 27%, others of 25% natural, 3% water and less than 1% aggregates shape the entire land use of the watershed. The watershed soil type can be broken down into the following categories:

i. 25% not mapped (urban);
ii. 21% silty loam;
iii. 15% bottomland;
iv. 13% coarse sand;

v. 9% loamy fine sand;

vi. 6% silty clay loam;

vii. 6% clay loam;

viii. 5% sandy loam.

Total of 6% of the entire watershed land is considered to be highly erodible, that means it could potentially contribute more than 7 tonnes/ha/year of soil to the watercourse (UTRCA, 2012).

Total area of the vegetation cover is approximately of about 1483 ha, forming about the one fourth of the total area of the watershed. Almost half of which is of a deciduous type, 27% is mixed and 4% is coniferous.

3.2.3.2 Streamflow and Water Quality

The annual 5-year (2006-2010) mean annual flow reported to be 46.1 cubic meter per second, while on a scale of 15-year the mean annual flow was 41.8 m^3/sec when measured near Byron area, denoted as “Water Quality Monitoring Site” on Figure 3-16. The River Bend watershed has a total length of 76 km of watercourses, 81% of which is natural the rest is either buried or channelized. The flow type in the watershed is 66% permanent, about one fifth of the total flow is intermittent and nearly 13% is of the buried type.

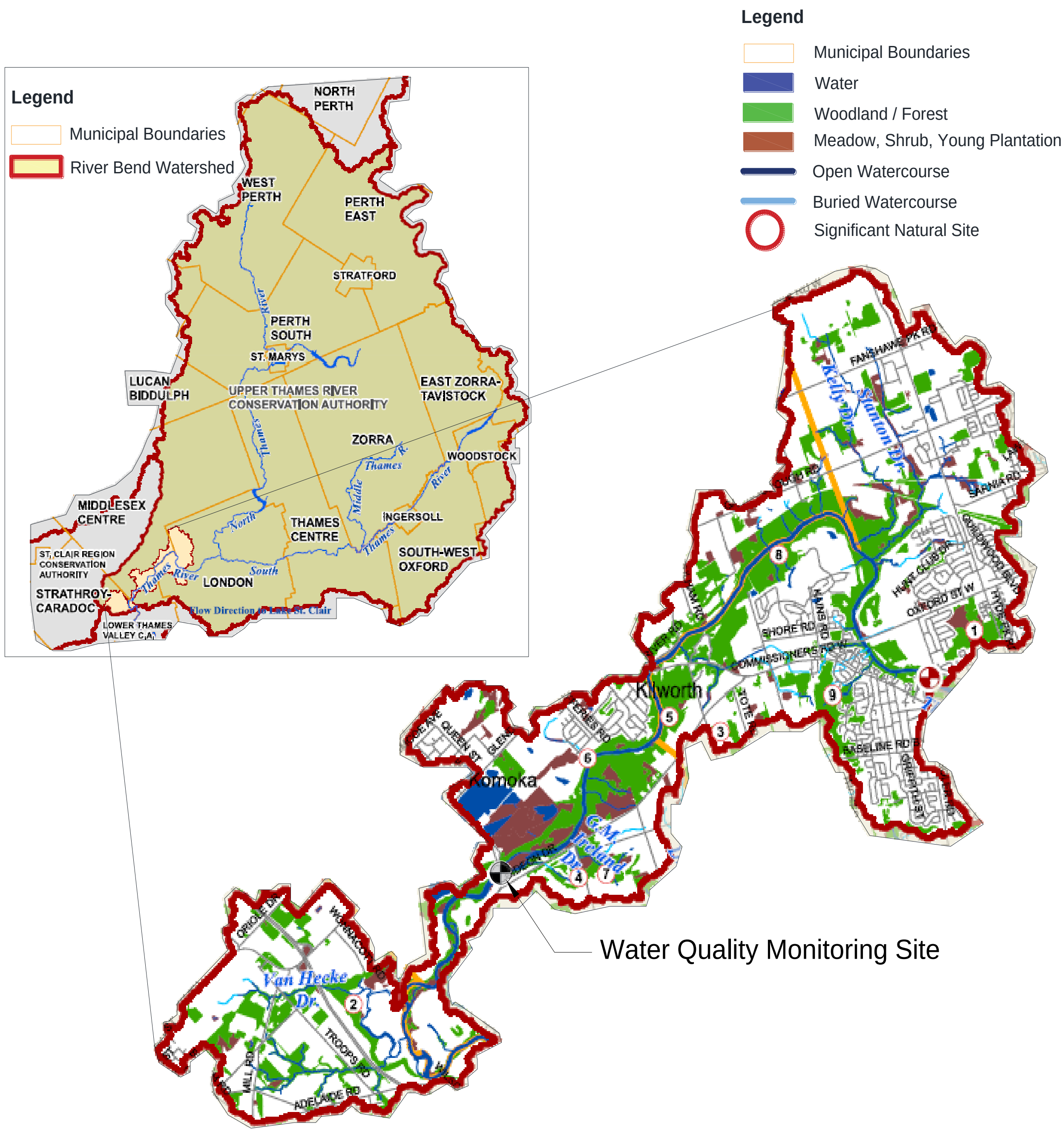


Figure 3-16 Upper Thames River Basin (on the left-had side) and River Bend Subwatershed, the research area of study (UTRCA, 2012)

River Bend watershed is the most downstream section of the Upper Thames River and its water quality is influenced by land uses and activities in this watershed and throughout the Upper Thames watershed. Four sewage treatment plants discharge treated effluent to the Thames in this watershed including London's Oxford Pollution Control Plant, Kilworth

Heights Wastewater Treatment Plant (WWTP), Komoka WWTP, and Mount Brydges WWTP. This watershed also has Springbank Dam, which marks the upstream end of this watershed as well as 24 privately-owned barriers. However; samples taken at Byron water quality station shows that the river's water quality in this watershed is in improving, for instance, phosphorus levels at the outlet have improved since 1990 and have remained steady since 2005 (UTRCA, 2012).

3.2.3.3 Fisheries and Great Lakes Connection

The River Bend watershed is one of the Thames River watersheds, which is a part of the Lake Erie watershed. It takes 4 to 10 days for the water to reach Lake St. Clair and another two weeks approximately to arrive at Lake Erie. Lake Erie is a drinking water source for millions of people from Canada and the US.

River Bend watershed is recorded to be a habitat of different 54 fish species and more than 20 mussel species. More than a few species out of those are categorized as species at risk, so this study is crucial in conserving the suitable habitat for these fish types to survive, water quality concerns are a big part too (UTRCA, 2012).

3.2.3.4 Dataset for River Bend Watershed

Dataset used in this study, for tainting and testing purposes, were obtained from the water quality monitoring site at Byron (Latitude: 42°57'46.9" and Longitude: 81°19'54.9"W). Almost weekly time series data of river discharge (Q), water temperature (T), electrical conductivity (C) and suspended sediment concentration (SSC) from 1993 to 2016 were employed to develop various training and testing models. These data were downloaded from the web server of the City of London (https://www.london.ca/residents/Environment/Rivers-Creeks/Pages/Water-Quality.aspx). Simultaneous datasets of Q, T, C and SSC were recorded during that period. Appendix A shows the raw data used in this study. Separate scatter plots of time versus raw datasets of river discharge in (m^3/s), suspended sediment concentration (mg/l) and electric conductivity (μS/cm) are illustrated in the Figure 3-17, Figure 3-18 and Figure 3-19, respectively.

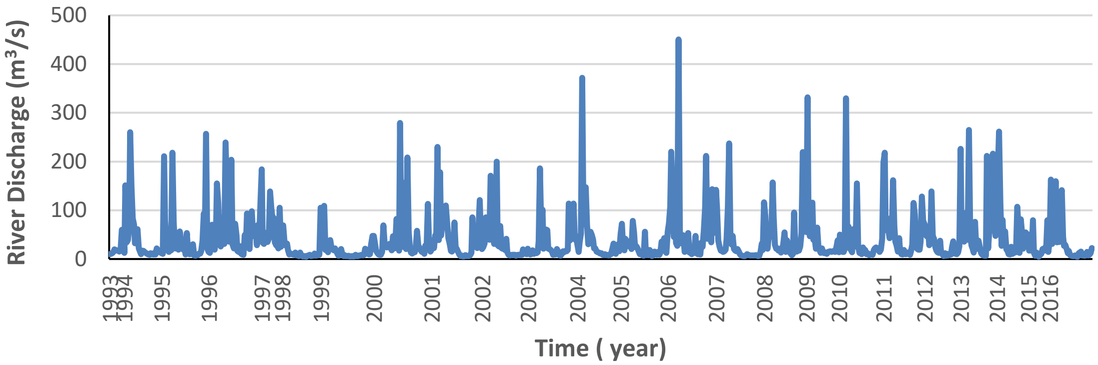


Figure 3-17 Scatter plot of the raw river discharge (m^3/s) data over the period between 1993-2016

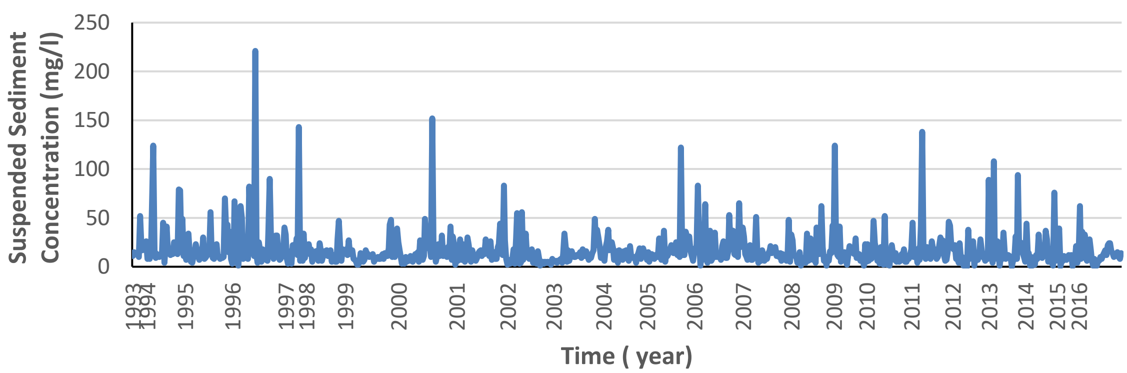


Figure 3-18 Scatter plot of the raw suspended sediment concentration (mg/l) data over the period between 1993-2016

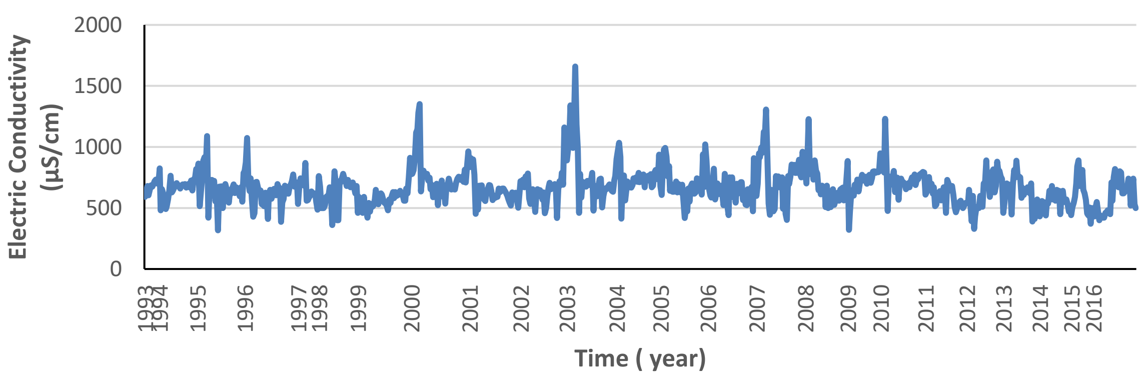


Figure 3-19 Scatter plot of the raw electric conductivity (µS/cm) data over the period between 1993-2016

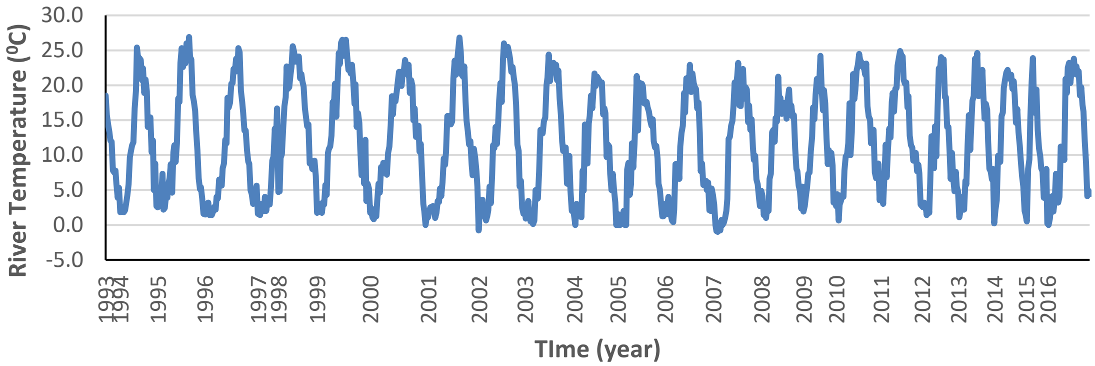


Figure 3-20 Scatter plot of the raw river temperature (^{0}C) data over the period between 1993-2016

CHAPTER 4
MODELS DEVELOPMENT AND RESULTS

4.1 Data Preprocessing

To model the SSC using SRC and SLR approaches, the only independent variable to be considered is the streamflow (m^3/s). Presence of some identical values of Q yielding different SSC values creates added difficulties for prediction ability of such models. A sample of these data is presented in Table 4-1. Only one value of Q which believed to be the most representative was considered, taking into consideration the corresponding trend of the observed SSC values. Thus, a fair comparison between the various models using conventional and machine learning approaches is guaranteed.

Table 4-1 Samples of Q data duplications

Date	**Time**	**T (ºC)**	**Q (m^3/s)**	**C (µS/cm)**	**SSC (mg/l)**
11-07-05	8:15 AM	18.8	6.7	556	18
26-08-13	7:45 AM	22.2	6.7	670	6
17-08-15	8:45 AM	23.9	6.7	890	11
26-09-16	7:55 AM	17.2	6.7	740	10
23-08-00	10:25 AM	21.5	13.0	741	5
21-07-03	7:50 AM	22.2	13.0	479	15
27-05-13	7:55 AM	14.8	13.0	887	8
14-09-15	7:50 AM	16.1	13.0	680	9

470 datasets for each of the input and output variables was considered in this study, to make better estimations for scenarios where Q is the only input. Outliers were also considered and the predicting procedure was performed using the Grubbs' test (Grubbs, 1969) with a significance level α of 95%. A percentage of 0.088% of the total raw data was found to be outliers and therefore they were removed. An example below shows how Grubbs' test is conducted:

A set of data of the discharge values were recorded as 80, 67, 76, 78, 66, 120 (m^3/s), Grubbs' test was conducted for the data outliers as follows:

i. Make a normal probability plot as below

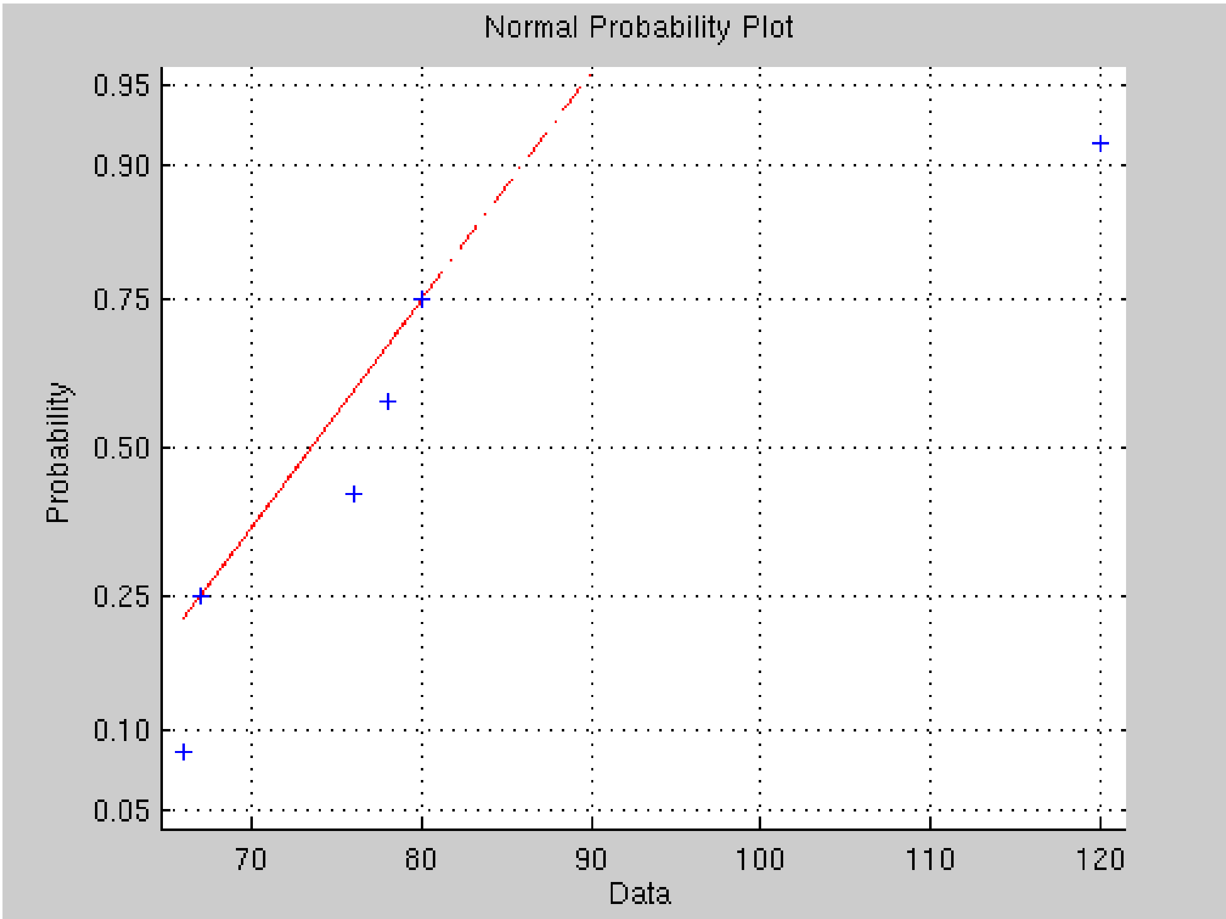


Figure 4-1 Example data, normal probability plot

ii. This plot looks nearly linear except for the point at 120. Therefore, this is probably an outlier.

iii. Use the Grubbs' test:

$$G = \frac{Y_{max} - \bar{Y}}{s} = \frac{120 - 81.2}{19.9} = 1.95 \tag{26}$$

$$G_{crit} = \frac{n-1}{\sqrt{n}} * \sqrt{\frac{t^2_{\frac{\alpha}{n},n-2}}{n-2+t^2_{\frac{\alpha}{n},n-2}}} = \frac{6-1}{\sqrt{6}} * \sqrt{\frac{3.964^2}{6-2+3.964^2}} = 1.82 \tag{27}$$

Where:

- Y_{max} = the outlier that is needed to be checked,
- $\bar{Y}$ = the mean of data,
- s = the data standard deviation,
- n = the number of data points, and
- $t_{\frac{\alpha}{n},n-2}$ = the t value for probability of $\frac{\alpha}{n}$ and dF of $n-2$.

iv. Because $G > G_{crit}$ the null hypothesis can be rejected, and the point is an outlier.

v. This point can be removed.

MATLAB code shown in Appendix B was used to remove some of the significant outliers which may affect the forecasting process of the various models to be developed, and the processed data after removing the duplications and the outliers are organized and shown in the Appendix C.

4.1.1 Training Dataset

Inputs were selected based on previous studies conducted to simulate the phenomena discussed broadly in the literature review section of this thesis. A study by Tyrrell (2015) investigated the effect of water temperature on sediment concentration and the study results indicate that sediment concentration is sensitive to changes in water temperature. The trend is that the sediment concentration decreases as water temperature warms. Mkpenie, Ebong, and Abasiekong (2007) also studied the effect of temperature on sedimentation (i.e. settling out of solid particles (sediments) in a liquid by gravity) and they concluded that the rate of sedimentation typically doubles for a 20˚C raise in temperature for some soluble substances. For such reasons temperature records were taking into consideration as an effective input for this study.. Dai *et al.*, (2009) studied the relationship between SSC and electric conductivity, the results show that good linear relationship exists between the SSC and the electrical conductivity. Since pure water does not conduct electricity very well because it contains very few ions, and that suspended sediment may contain dissolvable solids which can make the water more conductive, simultaneous data observed of electric conductivity was considered as an input to study its effect on the SSC in this study. The selecting procedure of the training dataset was performed randomly and more than 85% of the total processed data was chosen to be used for the training purposes of the various models of this study. Figures Figure 4-2, Figure 4-3, Figure 4-4 and Figure 4-5 show the distribution of the temperature, river flow, electric conductivity and suspended sediment concentration data over the training period, respectively.

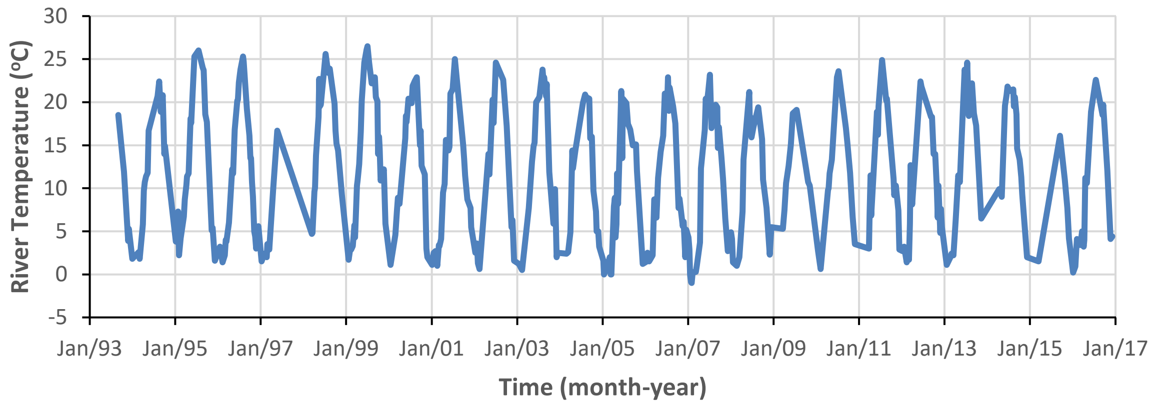


Figure 4-2 Temperature data (ºC) used for training various models

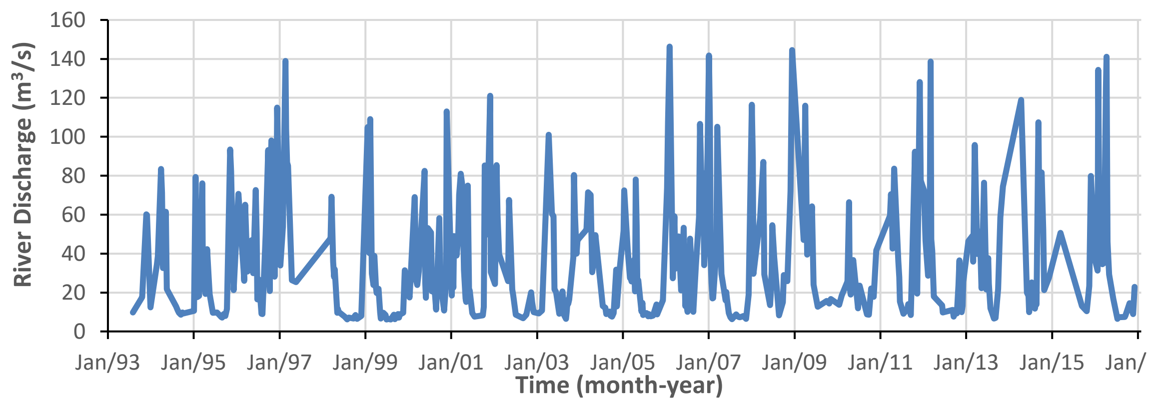


Figure 4-3 Discharge data (m^3/s) used for training various models

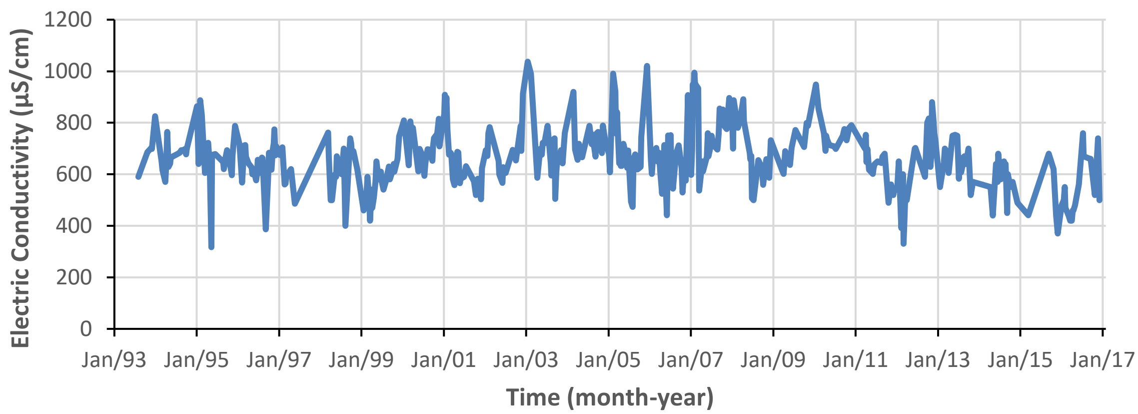


Figure 4-4 Electric Conductivity data (µS/cm) used for training various models

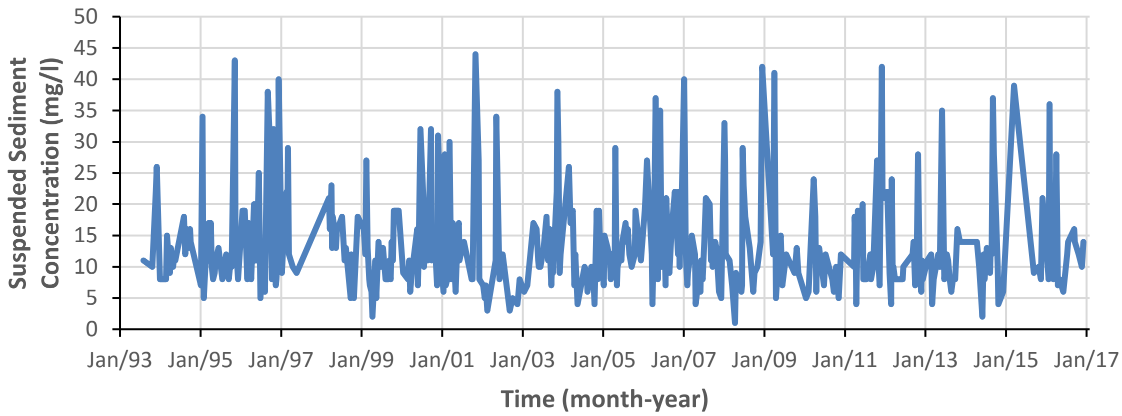


Figure 4-5 Suspended Sediment Concentration data (mg/l) used for training various models

4.1.2 Testing Dataset

Testing dataset selection was performed taking into consideration the data distribution for the dependent variable. Values from peaks, troughs and from the medium ranges, to assure a complete coverage of the entire dataset was considered. Nearly 15% of the total processed data was chosen to be used in the testing phase of the various models of this study in order to determine the best model. Figures Figure 4-6Figure 4-7Figure 4-8 andFigure 4-9 show the distribution of the temperature, river flow, electric conductivity and suspended sediment concentration data over the data testing period, respectively.

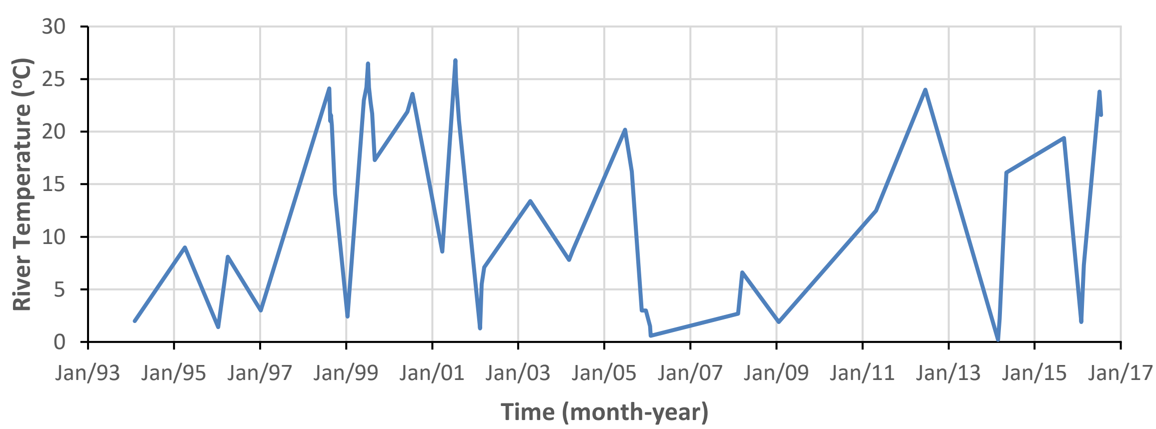


Figure 4-6 River temperature data (°C) used for testing various models

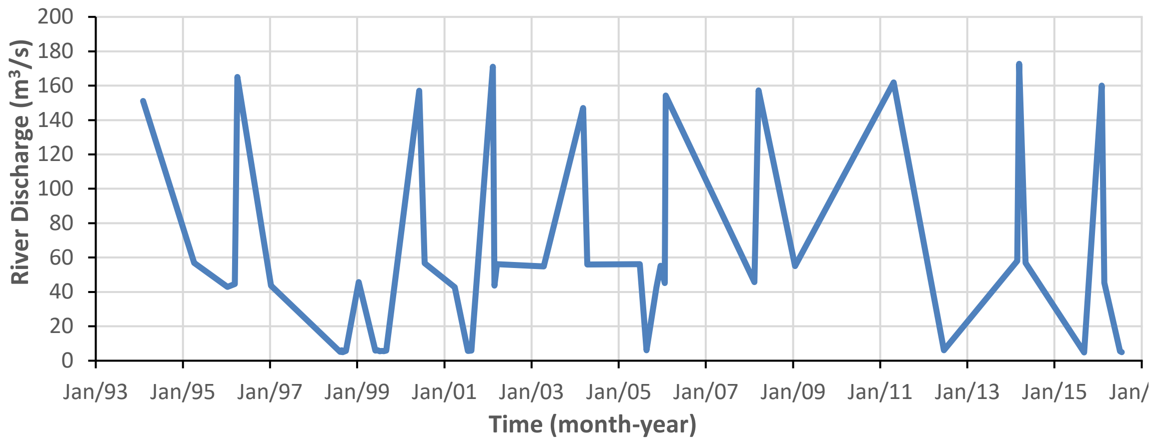


Figure 4-7 Discharge data (m^3/s) used for testing various models

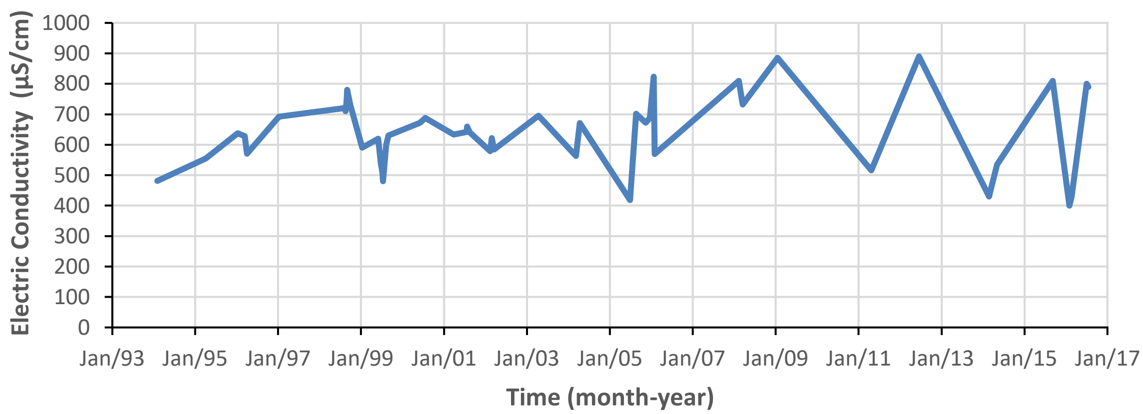


Figure 4-8 Electric conductivity data (µS/cm) used for testing various models

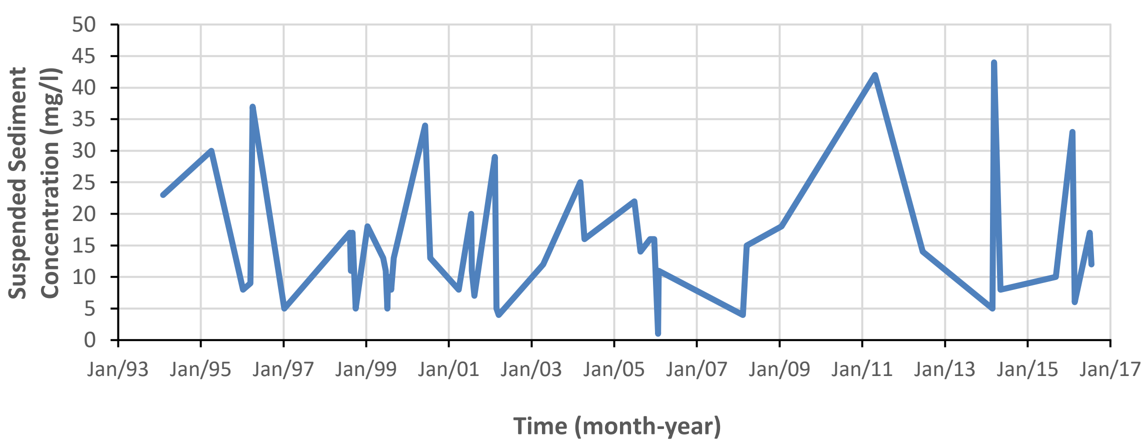


Figure 4-9 Suspended sediment concentration data (mg/l) used for testing various models

4.1.3 Models Performance Evaluation

To evaluate and examine the performance of the various models employed in this study, three different statistical measures were used to compute the goodness of each simulated model's results compared with the measured data. These statistical performance measures are; mean absolute error, root mean square error and Nash-Sutcliffe efficiency. Uncertainty analysis for the various models developed in this study is also conducted.

4.1.3.1 Mean Absolute Error (MAE)

The MAE measures the average magnitude of the errors without considering their direction. The MAE is the average over the verification sample of the absolute values of the differences between the calculated and the corresponding observed data. MAE values range from 0 to infinity and the smaller the value, the better the model. It is represented by the following form of equation:

$$MAE = \frac{\sum_{i=1}^{n}|x_i - y_i|}{n} \tag{28}$$

Where:

- x_i = The predicted data.
- y_i = The observed data.
- n = The number of observed data.

4.1.3.2 Root Mean Square Error (RMSE)

The RMSE is a quadratic scoring rule which measures the average magnitude of the error. The RMSE is the difference between computed and corresponding observed values are each squared and then averaged over the sample. After that, the square root of the average is taken. Because of squaring the errors takes place before they are averaged, the RMSE gives a high weight to large errors. It can range from 0 to infinity, the smaller the RMSE the better the forecasting model. RMSE is represented by the following form of equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{29}$$

4.1.3.3 Nash-Sutcliffe Efficiency (NSE)

The efficiency factor (NSE) proposed by Nash and Sutcliffe (1970) is expressed mathematically in the following form:

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right] \quad (30)$$

Where:

- $\bar{y}$ = The mean value of the observed data.

The NSE value can range from minus infinity to one, with an efficiency NSE of a value of 1.0 means perfect match between modeled results and observed records. An efficiency (NSE) of a value of 0 indicates that the model predictions are as accurate as the mean value of the observed time series data, and an efficiency of a value lower than 0 indicates that the mean value of the observed time series would have been a better predictor than the model.

4.1.3.4 Uncertainty Analysis

According to Wahl (2004) as following method can be used to analyze the uncertainty:

1. Compute individual prediction errors in terms of the number of log cycles separating the predicted and observed value.

$$e_i = \log_{10}(\hat{x}_i) - \log_{10}(x_i) = \log_{10}(\hat{x}_i/x_i) \quad (31)$$

Where:

- e_i = the prediction error,
- $\hat{x}_i$= the predicted value, and
- x_i= is the observed value.

2. Apply the outlier-exclusion algorithm to the series of prediction errors computed in Step 1. The algorithm is described by Rousseeuw (1998) as follows:

    - Determine the estimator of location, T = median(e_i).
    - Compute the deviations from the median, and determine the median of these absolute deviations, $MAD = median|T - e_i|$.
    - Compute an estimator of scale, $S_{MAD} = 1.483 * (MAD)$. The 1.483 factor makes S_{MAD} comparable to the standard deviation, which is the usual scale parameter of a normal distribution.
    - Compute a Z score for each observation, $Z_i = (e_i - T)/S_{MAD}$, Then reject any observations for which $|Z| > 2.5$. If the samples are from a perfect normal distribution, this method rejects at the 98.7% probability level.

3. Compute the mean, $\bar{e}$, and the standard deviation, S_e, of the remaining prediction errors. If the mean value is negative, it indicates that the prediction equation underestimated the observed values, and if positive the equation overestimated the observed values.

4. Using the values of $\bar{e}$ and S_e, one can express a confidence band around the predicted value of a parameter as, $\{\hat{x} \times 10^{-\bar{e}-2S_e}, \hat{x} \times 10^{-\bar{e}+2S_e}\}$. The use of $\pm 2S_e$, approximately yields a 95% confidence band.

4.2 Input Variable Scenarios

Four different scenarios were proposed using different combinations of the various inputs (measured simultaneously) affecting the output which is suspended sediment concentration (SSC). Table 4-2 shows the different scenarios and number of data used for training and testing purposes.

Table 4-2 Different scenarios proposed for this study

Scenario no.	Input	No. of training dataset for each variable	No. of testing dataset for each variable
S1	Streamflow (Q)	420	50
S2	Streamflow + Temperature (Q & T)	420	50
S3	Stream flow + Conductivity (Q & C)	420	50
S4	Streamflow + Temperature + Conductivity (Q, T, & C)	420	50

Since each of the sediment rating curve (SRC) and the simple linear regression (SLR) is one input and one output type of model, they were developed using scenario S1.

4.2.1 Training Various Models for S1

Four different models were developed using various modeling techniques, namely, SRC, SLR, ANFIS and ANN. For these models. streamflow (Q) was the only input taken into consideration to model the targeted output, SSC.

4.2.1.1 SRC Model

Training dataset was used to train the SRC model and the plot of log(Q) against log(SSC) is shown in Figure 4-10, and as explained in section 3.1.1 of this study, the slope of the trendline represents the b value, while log a is the y-intercept.

From the equation illustrated in Figure 4-10, a and b values of equation (2) of this study are calculated and found to be as follows:

$$\log a = 0.6898 \rightarrow a = 10^{0.6898} = 4.896$$

$$\mathrm{b} = 0.271$$

Therefore, the SSC using the SRC approach of the training dataset can be written as follows:

$$SSC = 4.986Q^{0.271} \tag{32}$$

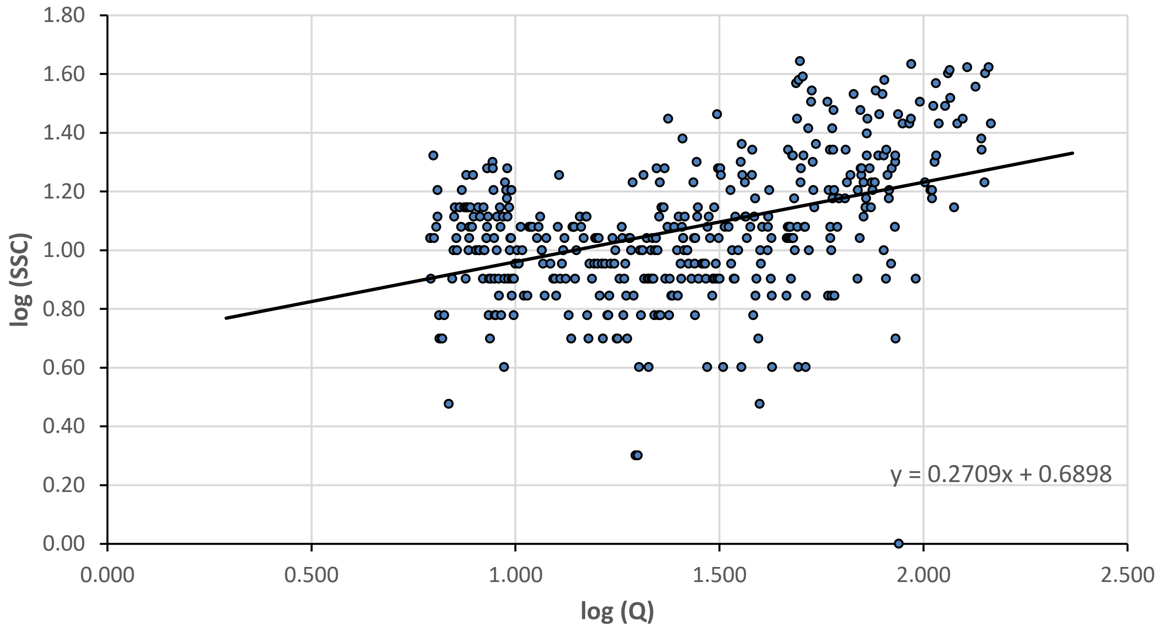


Figure 4-10 Applied SRC for the training dataset

Application of equation (32) to the training and testing dataset and calculations for MAE, RMSE and NSE were performed using Excel spreadsheet. It is worthy to note that the testing dataset was not a part of the data used in the training phase and is considered as an independent dataset to evaluate the effectiveness of the model. Thus, it has no influence on the derivation of the trendline used to determine the values of a and b. The results of the various statistical measures are shown in Table 4-3. To save space and to avoid repetition, only best model calculations for training and testing purposes are presented in Appendix D and Appendix E, respectively.

Table 4-3 Statistical measures for the training and testing phases of SRC model (S1)

	Training phase	Testing phase
MAE	4.824	6.936
RMSE	6.925	8.709
NSE	0.225	0.233

Figures Figure 4-11 and Figure 4-12 displays observed and estimated SSC (mg/l) in the training (validation) period and the extent of match between the measured and predicted

SSC (mg/l) in terms of a scatter diagram type of comparison with respect to the testing data by the SRC approach (S1), respectively.

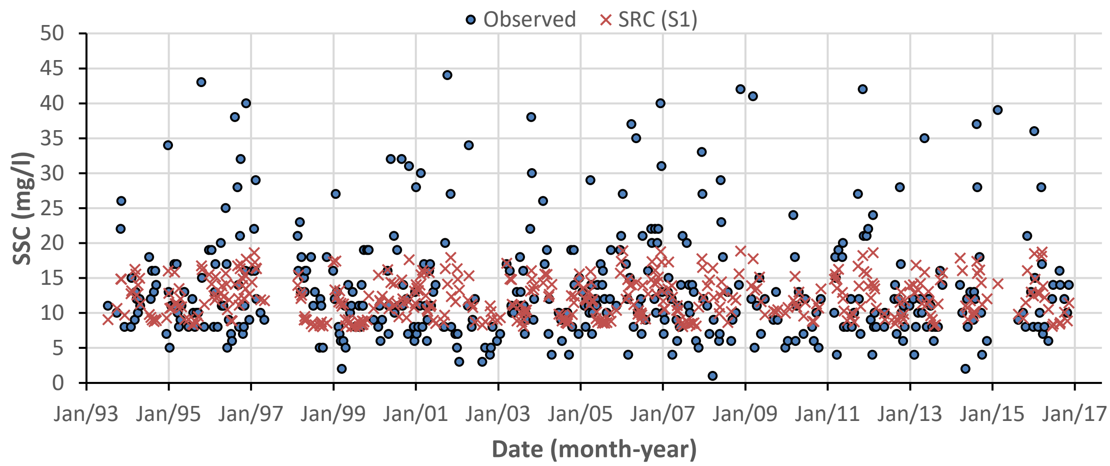


Figure 4-11 Observed and calculated SSC (mg/l), the training period using SRC (S1)

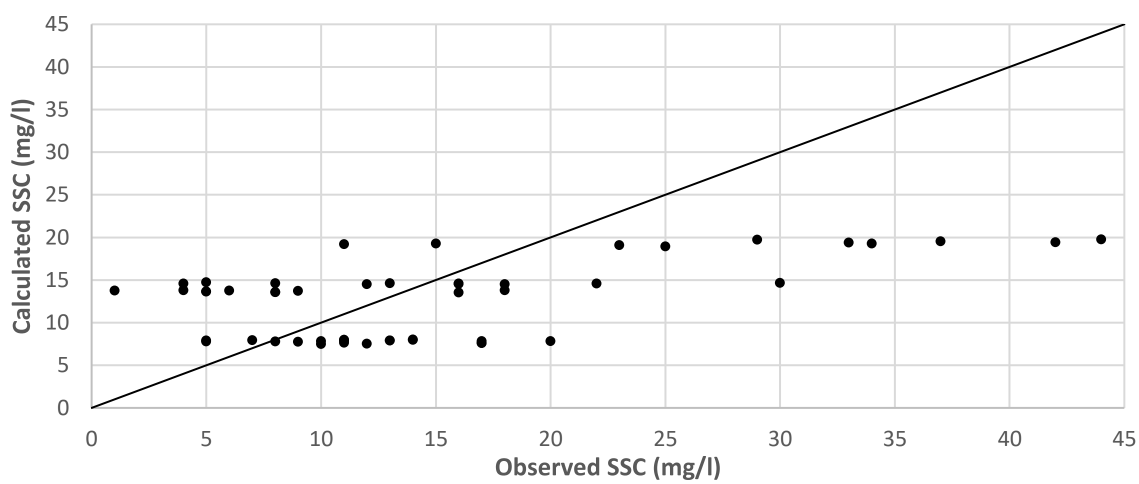


Figure 4-12 Scatter plot comparing predicted and observed SSC (mg/l) using SRC (S1), testing data

4.2.1.2 SLR Model

Training dataset was used to train the SLR model and the regression add-in tool in the Excel spreadsheet was used to train the model. With a confidence level of 95%, the regression significance F and Adjusted R Square were 2.47436E-43 and 0.365, respectively. Summary of various regression figures is presented in Table 4-4.

Table 4-4 Outputs of the SLR performed using S1's training dataset

	Coefficients	Standard Error	t Stat	P-value
Intercept	8.023657725	0.470740567	17.04475521	7.67331E-50
Q Variable 1	0.154018476	0.009907371	15.54584748	2.47436E-43

Form Table 4-4, the representative SSC equation can be written as follows:

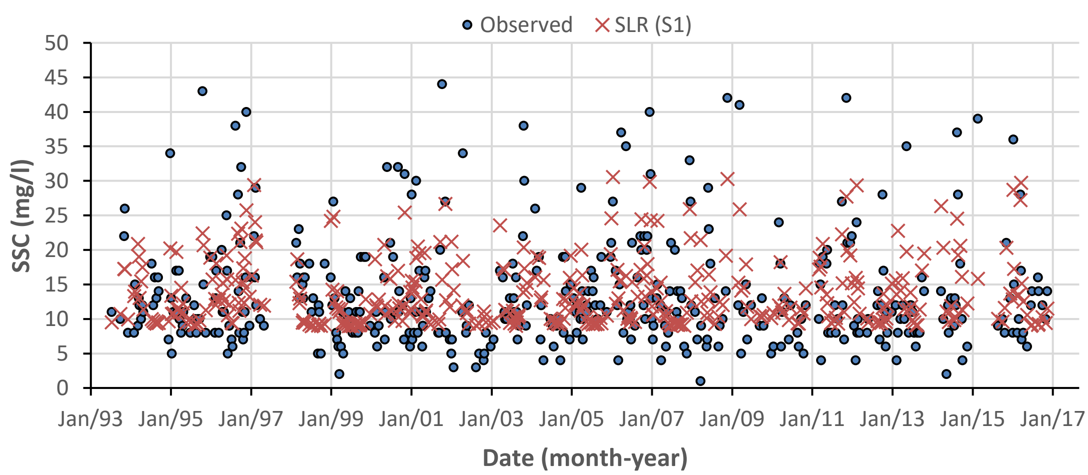
$$SSC_{SLR} = 0.154Q + 8.024 \tag{33}$$

Equation (33) was applied (exact coefficient of β_0 and $\beta_1$were used) to the training and testing datasets to determine the various statistical measures. Excel spreadsheets was used to perform the required calculations, and Table 4-5 shows the value of each performance indictor for training and testing phases separately.

Table 4-5 Statistical measures for the training and testing phases, SLR model (S1)

	Training Phase	Testing Phase
MAE	4.626	5.997
RMSE	6.262	7.563
NSE	0.366	0.421

Figures Figure 4-13



and Figure 4-14 display observed and calculated SSC (mg/l) over the training phase period and the extent of match between the measured and predicted SSC (mg/l) in terms of a scatter diagram type of comparison with respect to the testing data by the SLR approach (S1), respectively.

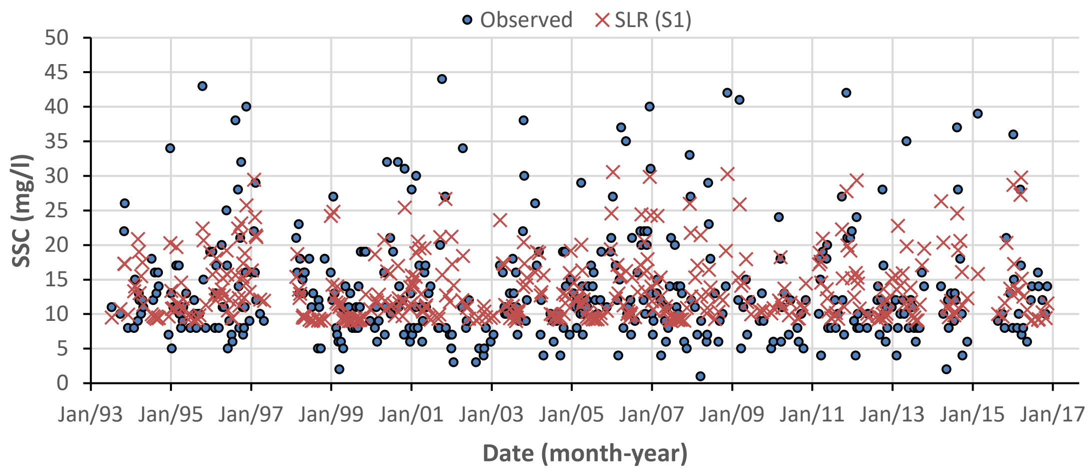


Figure 4-13 Observed and calculated SSC (mg/l), the training period using SLR (S1)

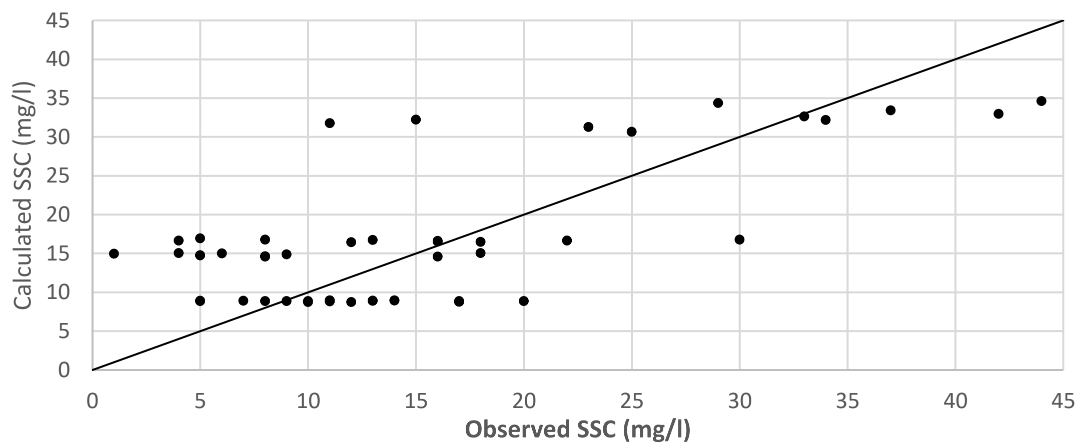


Figure 4-14 Scatter plot comparing predicted and observed SSC (mg/l) using SLR (S1), testing data

4.2.1.3 ANFIS Model

Several models were trained using the ANFISEDIT toolbox in MATLAB R2016b, using different numbers and structures of membership functions (MF). Hybrid optimization learning method used to train the various FIS. MF type constant was selected for the output. Only the best model's results will be presented here. In this case (S1) the GBELL MF type using 9 MFs was the one that best modeled the phenomena. Figure 4-15 and Figure 4-16 present the model structure and designing toolbox, the selected membership functions and the rules of the ANFIS model of S1 during training phase.

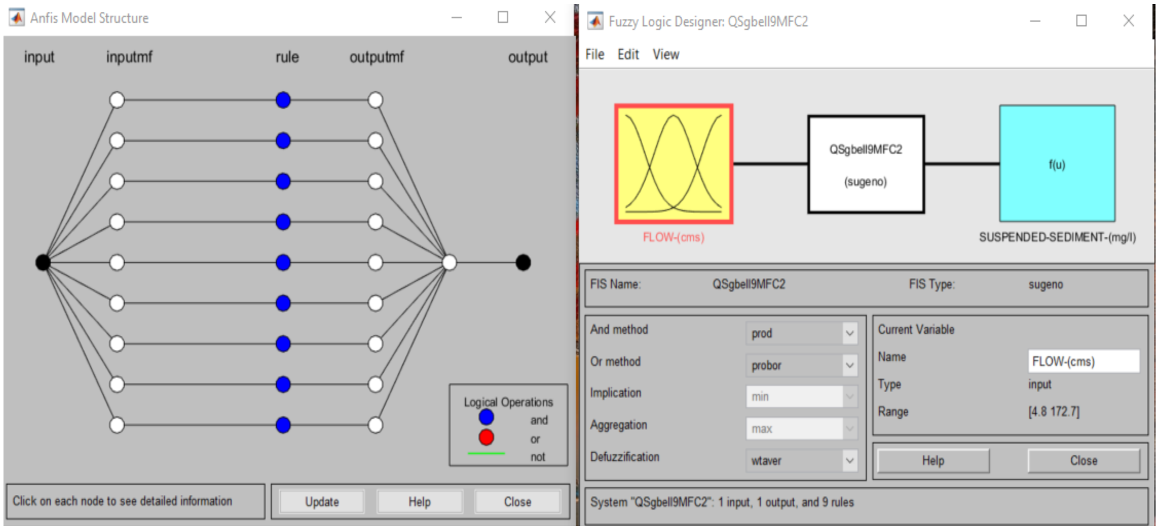


Figure 4-15 ANFIS structure and Fuzzy logic designer toolbox for (S1) training phase

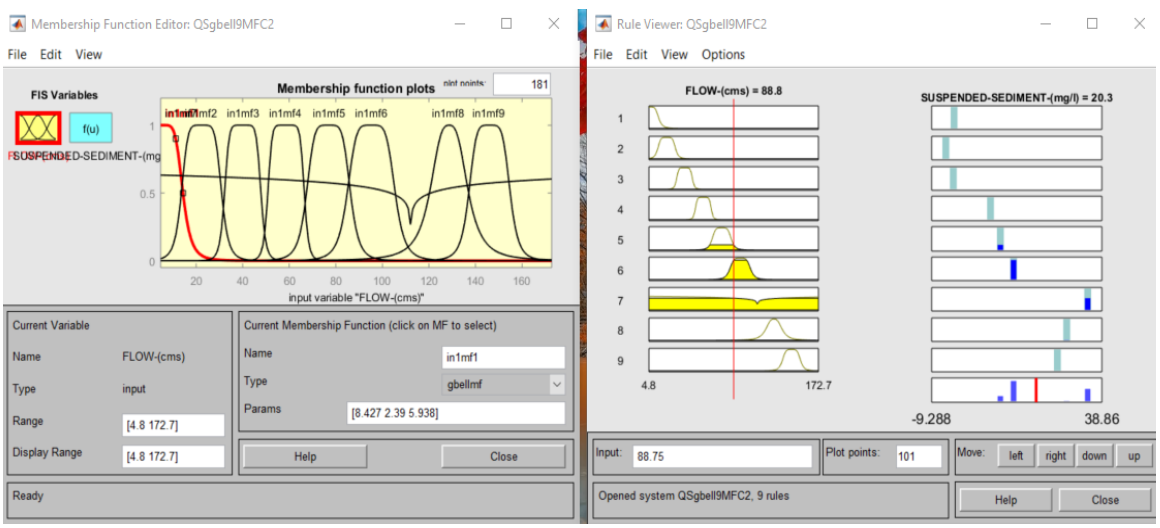


Figure 4-16 MF editor and rules for the ANFIS model (S1) training phase

The training and testing phases model performance measures are listed in Table 4-6, and Figure 4-17 displays the observed and calculated SSC (mg/l) over the training phase period for the ANFIS approach (S1). Figure 4-18 displays the extent of match between the measured and predicted SSC (mg/l) by the ANFIS model (S1) in terms of a scatter diagram type of comparison with respect to the testing data.

Table 4-6 Statistical measures for the training and testing phases, ANFIS model (S1)

	Training Phase	**Testing Phase**
MAE	4.277	5.194
RMSE	5.901	6.738
NSE	0.437	0.541

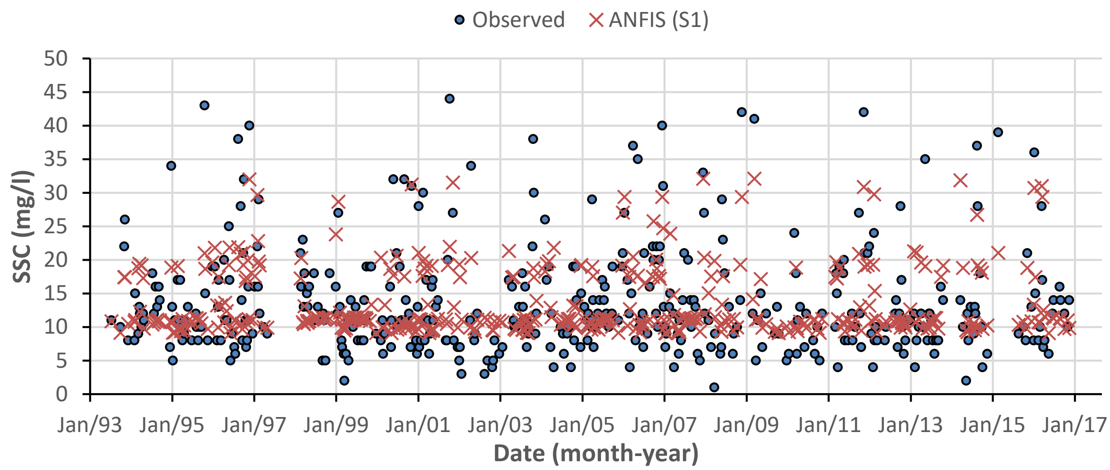


Figure 4-17 Observed and calculated SSC (mg/l), the training period using ANFIS (S1)

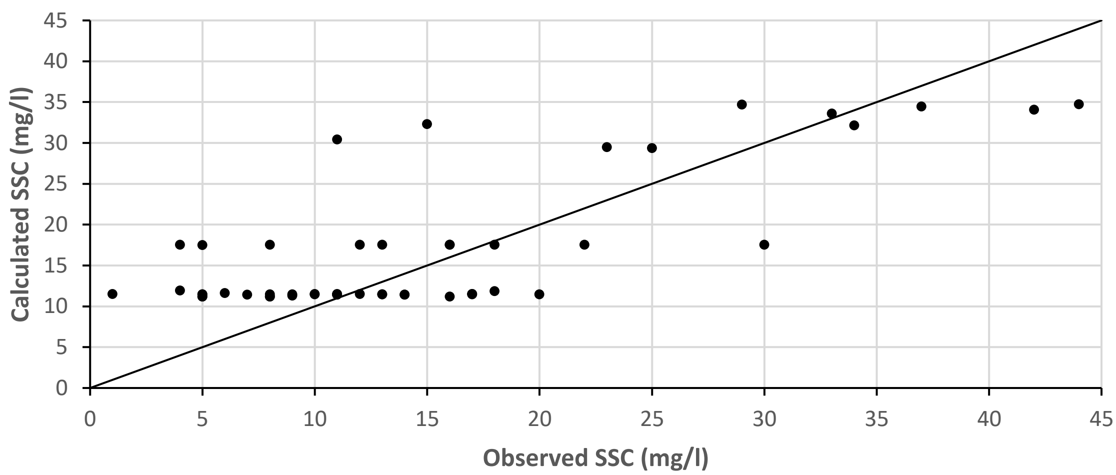


Figure 4-18 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S1), testing data

4.2.1.4 ANN Model

Several models were trained using the NNTOOL toolbox in MATLAB R2016b. Feed-forward backpropagation network type, Levenberg-Marquardt type training function and two hidden layers were chosen to train various models. Different trials were performed using different number of neurons and different types of transfer functions. Only the best model's results will be shown here. In this case (S1) types TANSIG and PURELIN transfer functions were selected as transfer functions for hidden layer 1 and 2, respectively, 20 neurons were used in the hidden layer 1. Figure 4-19 presents the best network structure for S1 using all dataset.

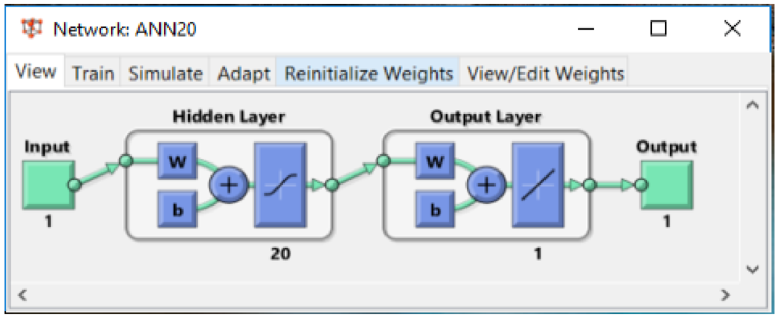


Figure 4-19 ANN best structure using all dataset (S1)

The model was trained using 85% of data in training purposes while the 15% of the entire set of the processed data was chosen for the testing determinations. After successfully achieving the best model, the calculated output is then extracted and used to calculate various performance indicators taking into consideration the same dataset chosen for training and testing purposes for the previous approaches to ensure fair comparison. Figure 4-20 displays the best model's outputs, after several epochs (trails) for the S1. Note that the model's built-in performance measure is R, which represents the square root of the coefficient of determination $\sqrt{R^2}$.

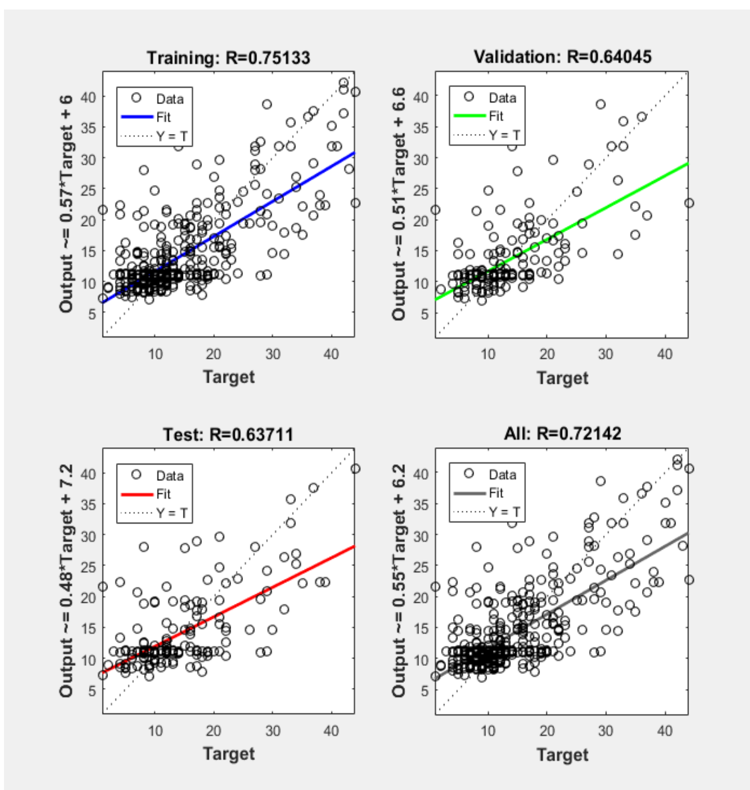


Figure 4-20 Best ANN model outputs (S1)

The training and testing phases' model performance measures are listed in Table 4-7, and Figure 4-21 displays the observed and calculated SSC (mg/l) over the training phase period for the ANN approach (S1). Figure 4-22 displays the extent of match between the measured and predicted SSC (mg/l) by the ANN model (S1) in terms of a scatter diagram type of comparison with respect to the testing data.

Table 4-7 Statistical measures for the training and testing phases, ANN model (S1)

	Training Phase	Testing Phase
MAE	4.013	4.250
RMSE	5.641	5.579
NSE	0.486	0.685

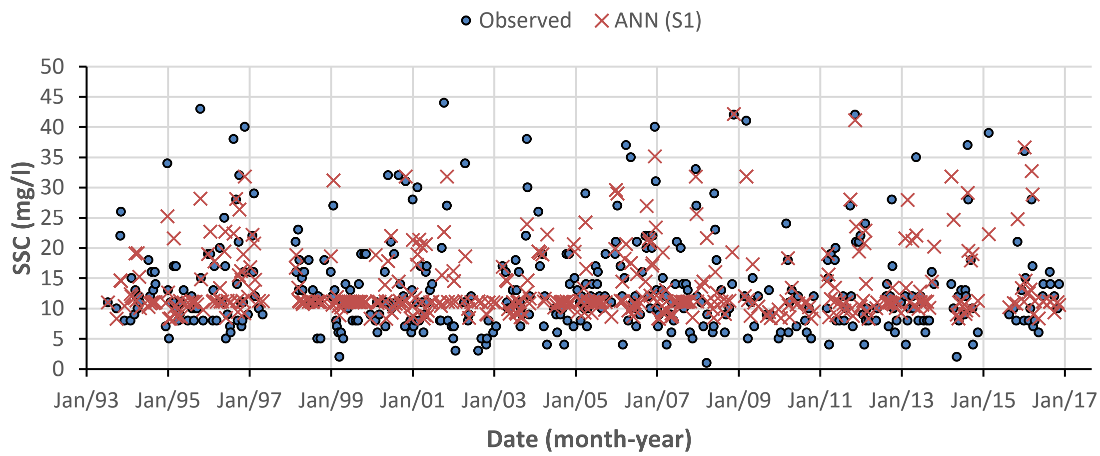


Figure 4-21 Observed and calculated SSC (mg/l), the training period using ANN (S1)

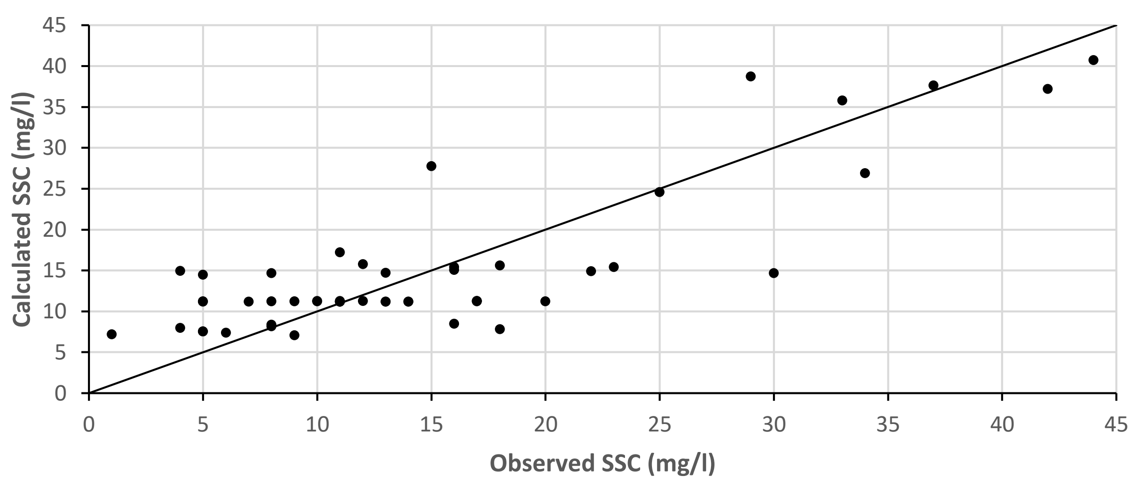
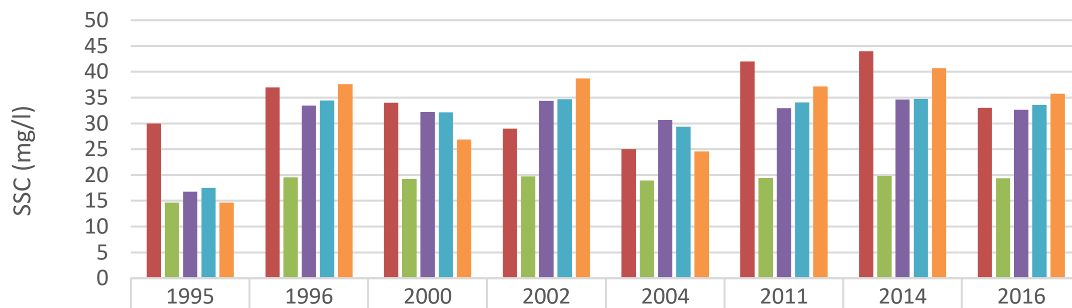


Figure 4-22 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S1), testing data

To sum up all the models for the S1 scenario, Table 4-8 shows the different performance indicators used in each model, and Figure 4-23 shows the observed SSC along with the calculated SSC (mg/l) of some peaks from the testing phase using SRC, SLR, ANFIS and ANN modeling approaches for the S1 case. An excellence performance of the machine learning approaches (ANFIS and ANN) over the conventional approaches (SRC and SLR) is recognized. A superiority of the ANN approach is observed.

Table 4-8 Summary of different performance indicators for all models of S1

	SRC		**SLR**		**ANFIS**		**ANN**	
	Training Phase	Testing Phase	Training Phase	Testing Phase	Training Phase	Testing Phase	Training Phase	Testing Phase
MAE	4.824	6.936	4.626	5.997	4.277	5.194	**4.013**	**4.250**
RMSE	6.925	8.709	6.262	7.563	5.901	6.738	**5.641**	**5.579**
NSE	0.225	0.233	0.366	0.421	0.437	0.541	**0.486**	**0.685**



	1995	1996	2000	2002	2004	2011	2014	2016
Observed	30.000	37.000	34.000	29.000	25.000	42.000	44.000	33.000
SRC	14.638	19.533	19.272	19.723	18.931	19.433	19.776	19.371
SLR	16.787	33.437	32.205	34.361	30.664	32.959	34.623	32.667
ANFIS	17.536	34.455	32.173	34.696	29.378	34.072	34.724	33.596
ANN	14.653	37.626	26.869	38.690	24.565	37.179	40.691	35.766

Figure 4-23 Selected peaks of observed SSC (mg/l) from the testing phase period for S1 and the calculated SSC (mg/l) using proposed approaches

4.2.1.5 Uncertainty Analysis of S1 Results

Uncertainty analysis for all the four models developed for S1 has been performed according to the procedure detailed in the section 4.1.3.4 in this study. In order to save space only the best model's calculation procedure for the uncertainty analysis various measures have been included in the Appendix F. Results for uncertainty estimates for S1 are given in Table 4-9 and it is observable that the four approaches (models) had an absolute mean prediction error of less than one-fifteenth except the ANN model, which had absolute mean prediction error of less than one-one hundred eightieth order of magnitude. Meaning that its prediction is of a better performance among all. The uncertainty bands were similar for all approaches (±0.3 to ±0.4 log cycles), and again with the best results for the ANN model.

Table 4-9 Uncertainty estimates for S1 various models

Approach	Mean prediction error log cycles	Width of uncertainty band, $\pm 2S_e$ log cycles	Prediction interval around hypothetical prediction value of $\hat{x}$ =1.0
SRC	-0.067	± 0.442	0.310 - 2.375
SLR	0.023	± 0.427	0.395 – 2.822
ANFIS	0.064	± 0.369	0.496 – 2.710
ANN	0.005	± 0.334	0.470 – 2.183

4.2.2 Training Various Models for S2

Three different models were developed using various modeling techniques, namely, MLR, ANFIS and ANN. For these models, the temperature (T) and the streamflow (Q) were used as inputs in order to model the targeted output (SSC).

4.2.2.1 MLR Model

Training dataset was used to train the MLR model and the regression add-in tool in the Excel spreadsheet was used to train the data. With a confidence level of 95%, the regression significance F and Adjusted R Square were 1.76626E-47 and 0.4, respectively, showing an improvement than what it was obtained in the SLR model in the S1 (Section 4.2.1.2). Meaning that the new input influences the output. The results summary of various regression figures is presented in Table 4-10.

Table 4-10 Outputs of the MLR performed using S2's training dataset

	Coefficients	Standard Error	t Stat	P-value
Intercept	4.305846954	0.862959382	4.989628765	8.89926E-07
T Variable 1	0.233848364	0.046029283	5.080425937	5.6928E-07
Q Variable 2	0.181922697	0.011082591	16.41517737	4.54393E-47

From Table 4-10, the SSC equation can be written as follows:

$$SSC_{MLR} = 0.2338T + 0.1819Q + 4.306 \tag{34}$$

Applying the Equation (34) (exact coefficients of $\beta_0,\ \beta_1 and\ \beta_2$ were used) to the training and testing datasets in order to find the various statistical measures representing the goodness of fit between the predicted and observed data using MLR approach (S2). Excel spreadsheets was used to perform different needed calculations, and the Table 4-11 shows the values of each measure. Moreover, the Figure 4-24 presents the observed and calculated SSC (mg/l) over the training phase period, and Figure 4-25 displays the extent of match between the measured and predicted SSC (mg/l) by the MLR model (S2) in terms of a scatter diagram type of comparison with respect to the testing data.

Table 4-11 Statistical measures for the training and testing phases, MLR model (S2)

	Training Phase	**Testing Phase**
MAE	4.422	5.641
RMSE	6.077	7.266
NSE	0.403	0.466

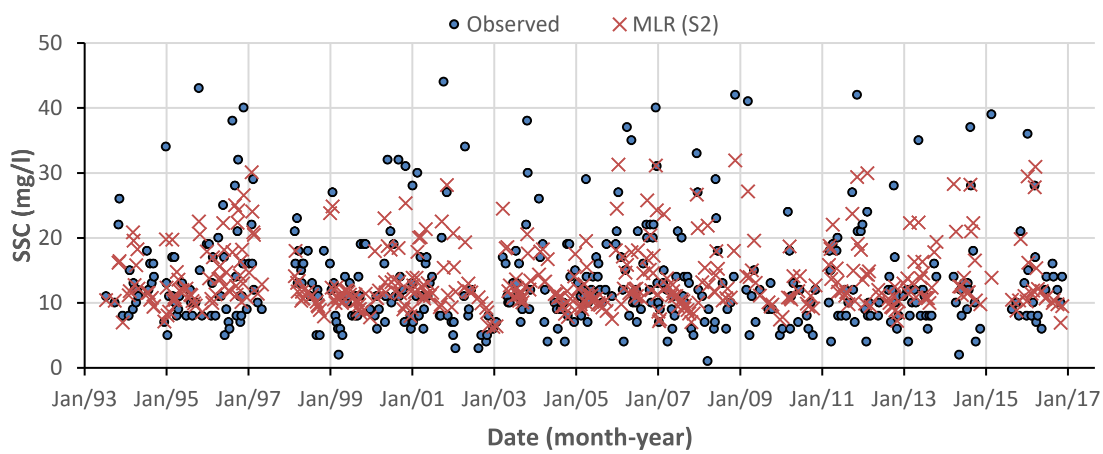


Figure 4-24 Observed and calculated SSC (mg/l), the training period using MLR (S2)

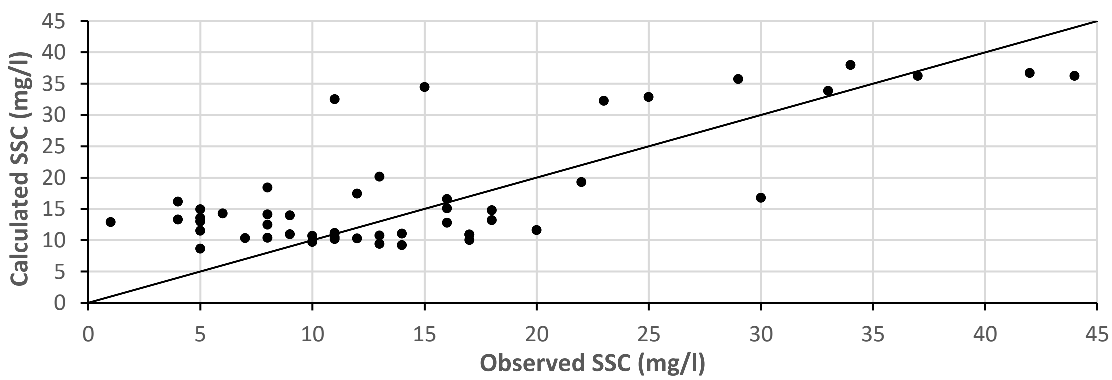


Figure 4-25 Scatter plot comparing predicted and observed SSC (mg/l) using MLR (S2), testing data

4.2.2.2 ANFIS Model

Several models were trained using the ANFISEDIT toolbox in MATLAB R2016b, using different numbers and structures of membership functions (MF). Hybrid optimization learning method used to train the various FIS. MF type constant was selected for the output. Only the best model's results will be presented here. In this case (S2) the GAUSS2 MF type using 5 and 2 MFs for Temperature and Streamflow inputs, respectively, was the one that best modeled the phenomena. Figure 4-26, Figure 4-27 and Figure 4-28 present the model structure and designing toolbox, the membership functions for the two inputs and the rules of the ANFIS model for S2 training phase.

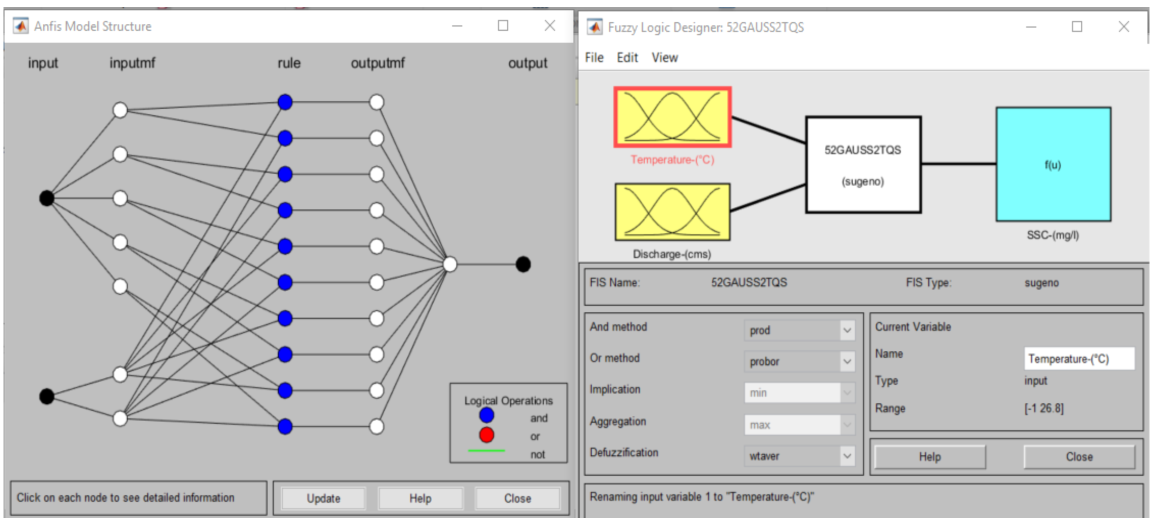


Figure 4-26 ANFIS model structure and Fuzzy logic designer toolbox for S2 training phase

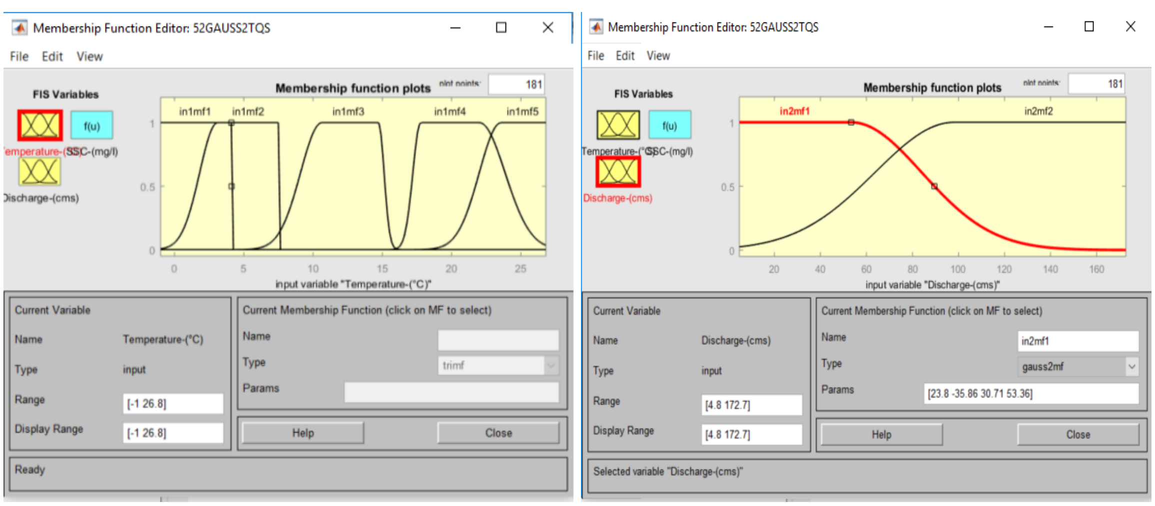


Figure 4-27 MFs editor for the two main inputs T (on the left-hand side) and Q for the ANFIS model, training phase (S2)

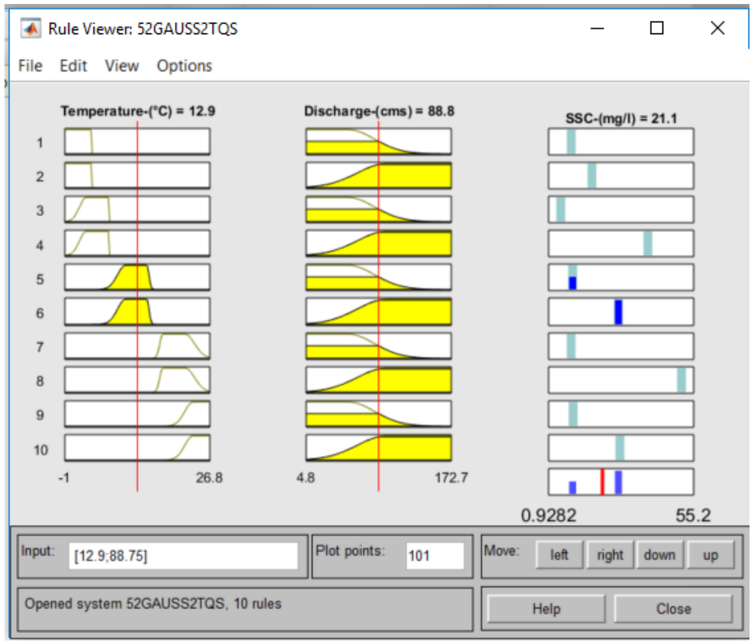


Figure 4-28 Rule viewer for the ANFIS model, training phase (S2)

The model performance measures for training and testing phases, representing the goodness of fit between the predicted and observed data for the ANFIS model (2), are listed in Table 4-12. Figure 4-29 displays the observed and calculated SSC (mg/l) over the training phase period for the ANFIS approach (S2). Figure 4-30 displays the extent of match between the measured and predicted SSC (mg/l) by the ANFIS model (S2) in terms of a scatter diagram type of comparison with respect to the testing data.

Table 4-12 Statistical measures for the training and testing phases, ANFIS model (S2)

	Training Phase	**Testing Phase**
MAE	4.641	5.420
RMSE	5.776	6.813
NSE	0.461	0.531

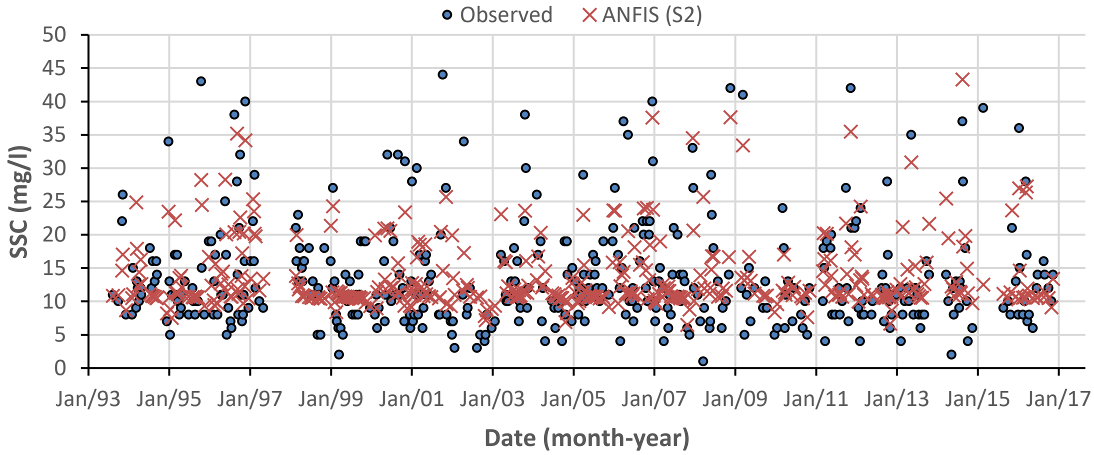


Figure 4-29 Observed and calculated SSC (mg/l), the training period using ANFIS (S2)

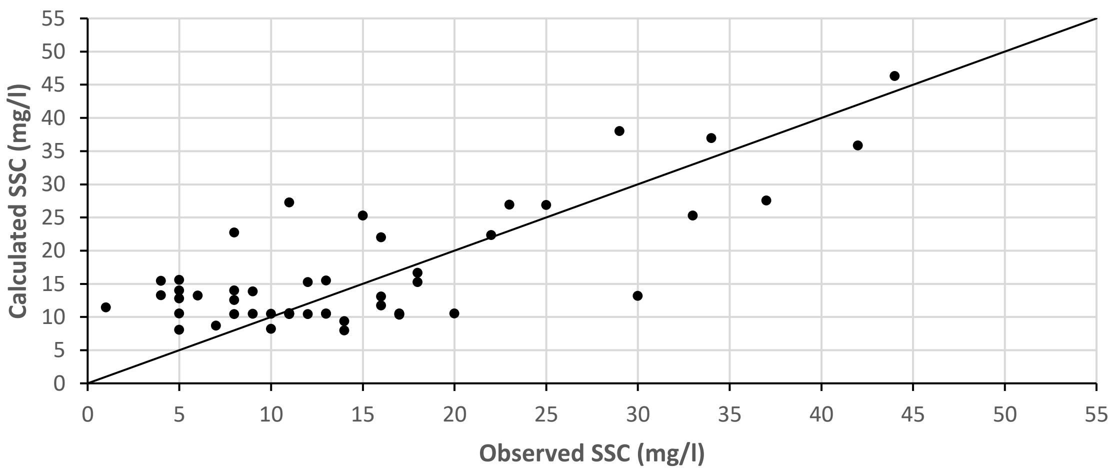


Figure 4-30 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S2), testing data

4.2.2.3 ANN Model

As discussed in section 4.2.1.4, different trails were performed using different number of neurons and different types of transfer functions. Only the best model's results will be shown here. In this case (S2) types TANSIG and PURELIN transfer functions were selected as transfer functions for hidden layer 1 and 2, respectively, 20 neurons were used in the hidden layer 1. Figure 4-31 presents the best network structure for S2 using all dataset.

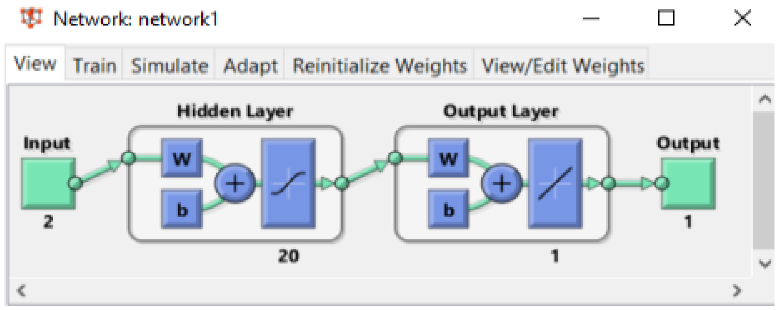


Figure 4-31 ANN best structure using all dataset (S2)

The model was trained using 85% of data in training purposes while the 15% of the entire set of the processed data was chosen for the testing determinations. After successfully achieving the best model, the calculated output is then extracted and used to calculate various performance indicators taking into consideration the same dataset chosen for training and testing purposes for the previous approaches to ensure fair comparison. Figure 4-32 displays the best model's outputs, after several epochs (trails) for the S2. Note that the model's built-in performance measure is R, which represents $\sqrt{R^2}$.

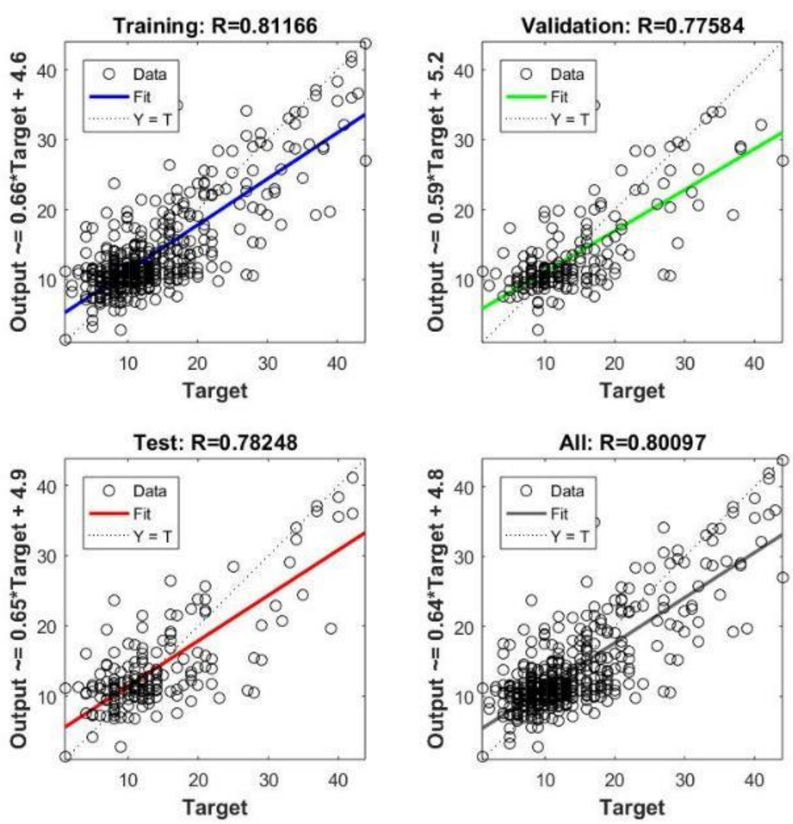


Figure 4-32 Best ANN model outputs (S2)

The training phase's model performance measures along with the testing phase's measures, in order to determine the goodness of fit between measured and calculated SSC using the ANN approach (S2), are listed in Table 4-13, and Figure 4-21 displays the observed and calculated SSC (mg/l) over the training phase period for the ANN approach (S2). Figure 4-34 displays the extent of match between the measured and predicted SSC (mg/l) by the ANN model (S2) in terms of a scatter diagram type of comparison with respect to the testing data.

Table 4-13 Statistical measures for the training and testing phases, ANN model (S2)

	Training Phase	Testing Phase
MAE	3.533	3.590
RMSE	4.865	4.869
NSE	0.617	0.760

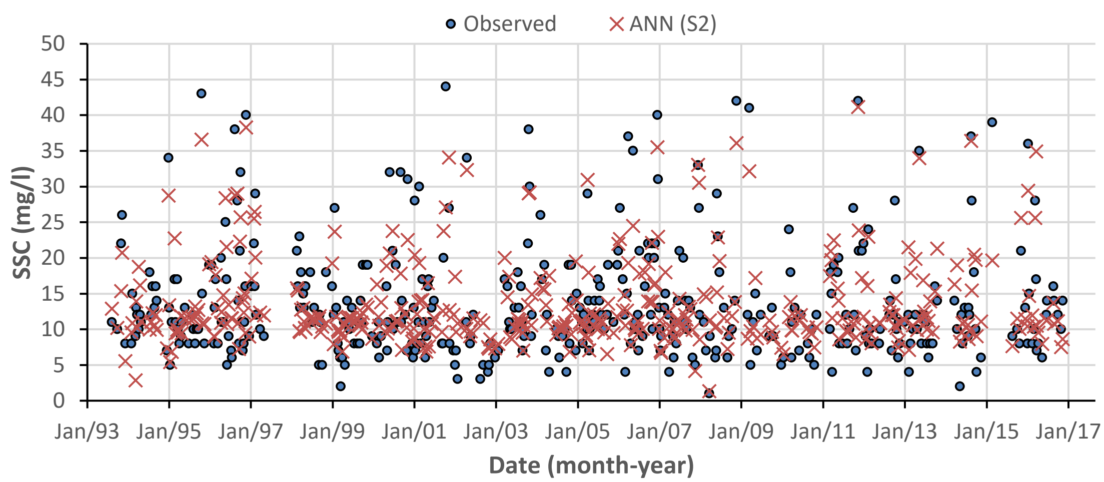


Figure 4-33 Observed and calculated SSC (mg/l), the training period using ANN (S2)

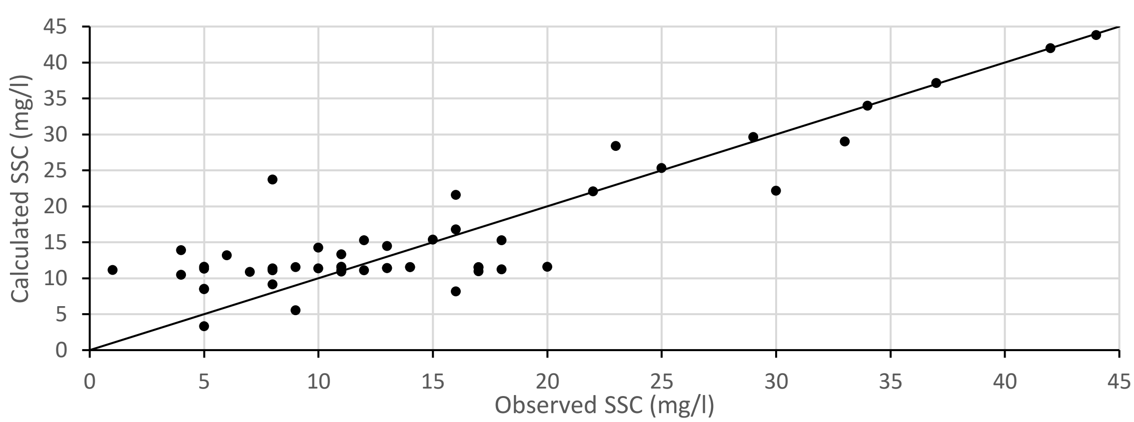
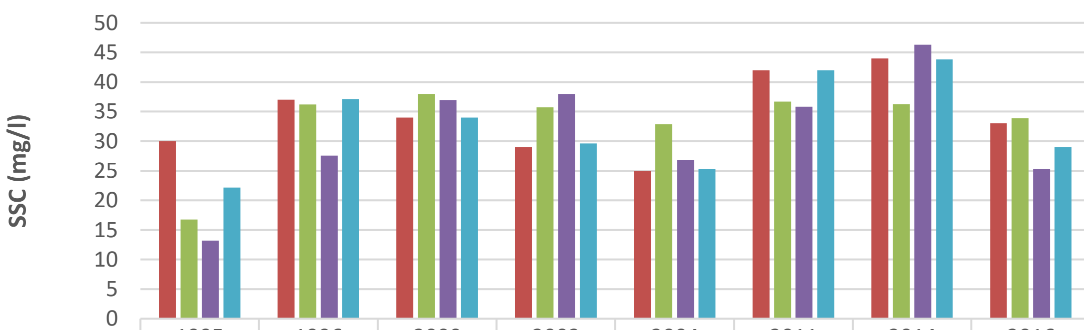


Figure 4-34 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S2), testing data

To sum up all the models for the S2 case, the Table 4-14 shows the different performance indicators used in each case, and Figure 4-35 shows the observed SSC along with the calculated SSC (mg/l) of selected peaks from the testing phase using MLR, ANFIS and ANN modeling approaches for the S2 case. An excellence performance of the machine learning approaches, i.e. ANFIS and ANN, over the conventional method, i.e. MLR, is recognized. A superiority of the ANN approach is observed.

Table 4-14 Summary of different performance indicators for all models used using the S2 inputs

	MLR		**ANFIS**		**ANN**	
	Training Phase	Testing Phase	Training Phase	Testing Phase	Training Phase	Testing Phase
MAE	4.422	5.641	4.641	5.420	**3.533**	**3.590**
RMSE	6.077	7.266	5.776	6.813	**4.865**	**4.869**
NSE	0.403	0.466	0.461	0.531	**0.617**	**0.760**



	1995	1996	2000	2002	2004	2011	2014	2016
Observed	30.000	37.000	34.000	29.000	25.000	42.000	44.000	33.000
MLR	16.762	36.217	37.989	35.719	32.873	36.682	36.238	33.858
ANFIS	13.204	27.569	36.949	38.008	26.878	35.841	46.328	25.320
ANN	22.163	37.122	33.993	29.610	25.297	41.979	43.814	29.006

Figure 4-35 Selected peaks of observed SSC (mg/l) from the testing phase period for S2 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon

4.2.2.4 Uncertainty Analysis of S2 Results

Uncertainty analysis for all the four models developed for S2 has been performed according to the procedure detailed in the section 4.1.3.4 in this study. Results are given in Table 4-15 and it is observable that the three approaches (models) had an absolute mean prediction error of less than one-thirteenth with a superiority of the ANN model, which had an absolute mean prediction error of less than one-sixtieth order of magnitude. Meaning that its prediction is of a better performance among all. The uncertainty bands were similar for MLR and ANFIS (±0.4 log cycles) while for ANN it was calculated as ±0.3 log cycles.

Table 4-15 Uncertainty estimates for S2 various models

Approach	Mean prediction error log cycles	Width of uncertainty band, $\pm 2S_e$ log cycles	Prediction interval around hypothetical prediction value of $\hat{x}$ =1.0
MLR	0.073	± 0.420	0.450 – 3.116
ANFIS	0.047	± 0.422	0.422 – 2.943
ANN	0.015	± 0.313	0.503 – 2.126

4.2.3 Training Various Models for S3

Three different models were developed using various modeling techniques, namely, MLR, ANFIS and ANN. For these models, the electrical conductivity (C) and the streamflow (Q) were used as inputs in order to model the targeted output (SSC).

4.2.3.1 MLR Model

Training dataset was used to train the MLR model and the regression add-in tool in the excel spreadsheet was used to train the data. With a confidence level of 95%, the regression significance F and Adjusted R Square were 1.51717E-42 and 0.370, respectively, with a decrease in the F significance value and an improvement in the R adjusted value from what it was in the SLR case in the S1 (Section 4.2.1.2). The results summary of various regression figures is presented in Table 4-16

Table 4-16 Outputs of the MLR performed using S3's training dataset

	Coefficients	Standard Error	t Stat	P-value
Intercept	11.01506423	2.01608806	5.46358289	8.0402E-08
Q Variable 1	0.149594538	0.010307846	14.51268705	6.39071E-39
C Variable 2	-0.004275448	0.002802089	-1.525807605	0.127815854

From Table 4-16, the SSC equation can be written as follows:

$$SSC_{MLR} = 0.1496Q - 0.0043C + 11.0151 \tag{35}$$

Applying the Equation (35) (exact coefficient of $\beta_0,\ \beta_1 and\ \beta_2$ were used) to the training and the testing dataset in order to determine the various statistical measures representing the goodness of fit between the observed and the calculated SSC (mg/l) using the MLR model (S3). Excel spreadsheets was used to perform different needed calculations, and the Table 4-17 shows the values of each measure.

Table 4-17 Statistical measures for the training and testing phases, MLR model (S3)

	Training Phase	**Testing Phase**
MAE	4.597	6.038
RMSE	6.244	7.555
NSE	0.370	0.423

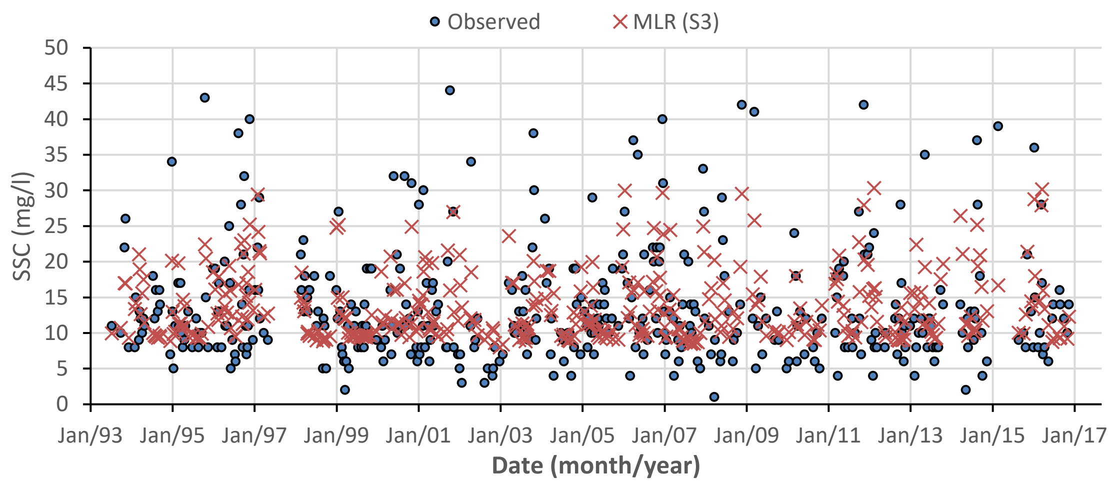


Figure 4-36 Observed and calculated SSC (mg/l), the training period using MLR (S3)

Moreover, the Figure 4-36 presents the observed and calculated SSC (mg/l) over the training phase period. and Figure 4-37 displays the extent of match between the measured and predicted SSC (mg/l) by the MLR model (S3) in terms of a scatter diagram type of comparison with respect to the testing data.

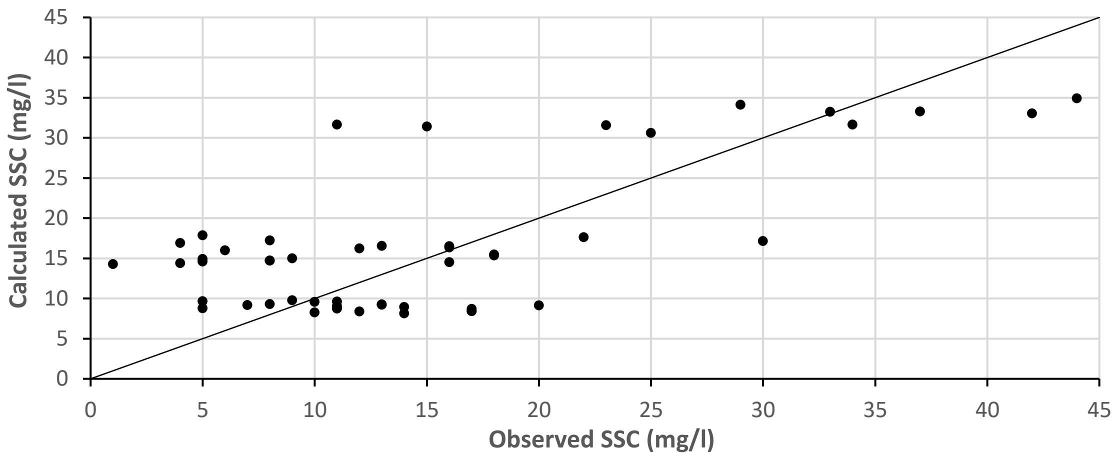


Figure 4-37 Scatter plot comparing predicted and observed SSC (mg/l) using MLR (S3). testing data

4.2.3.2 ANFIS Model

Several models were trained using the ANFISEDIT toolbox in MATLAB R2016b, using different numbers and structures of membership functions (MF). Hybrid optimization learning method used to train the various FIS. MF type constant was selected for the output. Only the best model's results will be presented here. In this case (S3) the GBELL MF type using 5 MFs for both inputs, was the one that best modeled the phenomena. Figure 4-38, Figure 4-39 and Figure 4-40 presents the model structure and designing toolbox, the membership functions for the two inputs and the rules of the ANFIS model for S3 training phase.

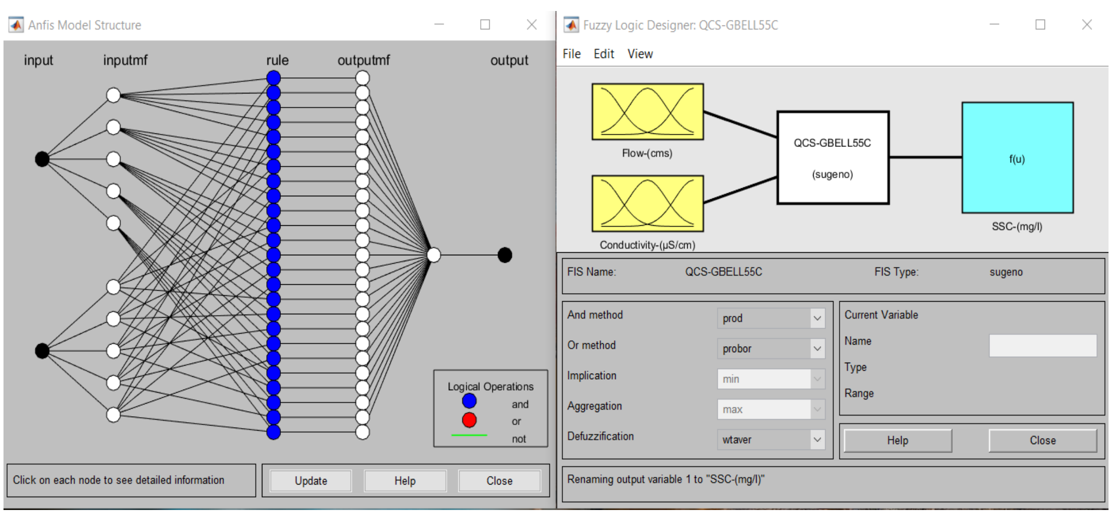


Figure 4-38 ANFIS model structure and Fuzzy logic designer toolbox for S3 training phase

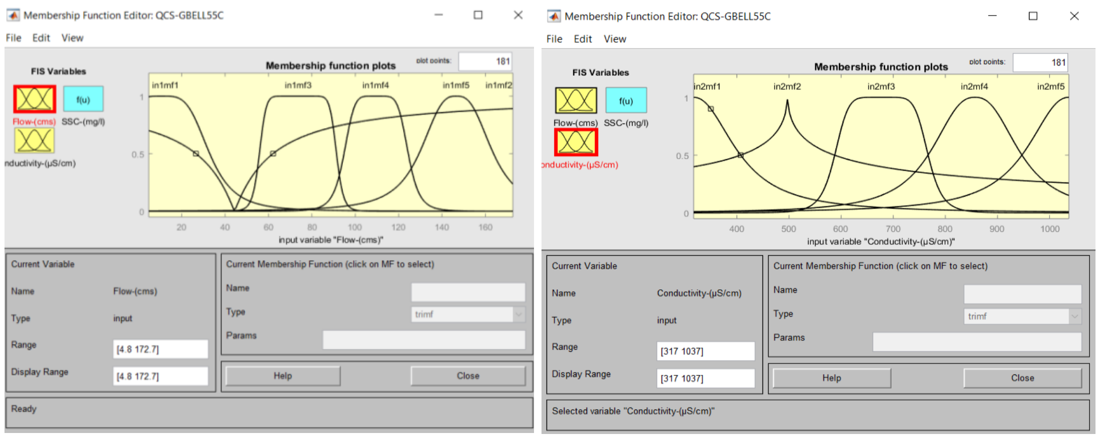


Figure 4-39 MFs editor for the two main inputs Q (on the left-hand side) and C for the ANFIS model, training phase (S3)

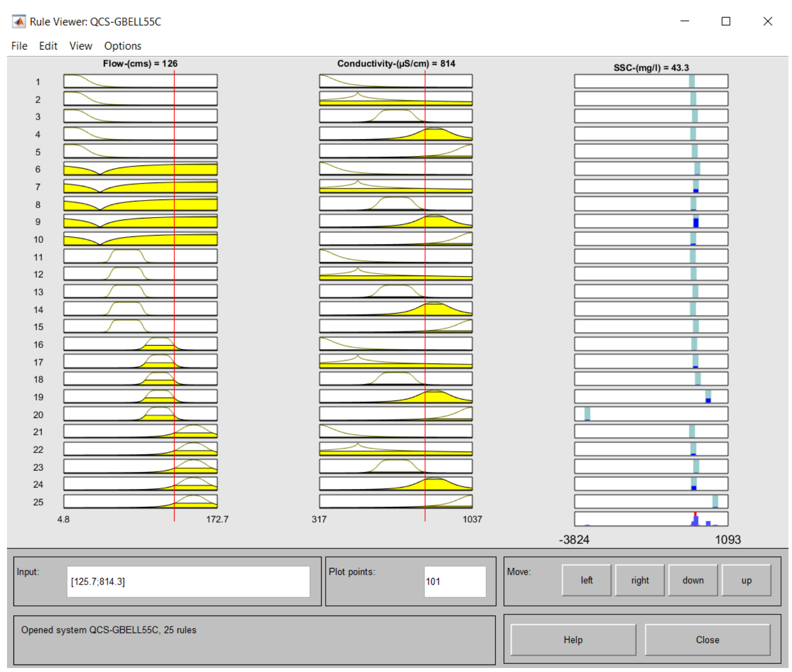


Figure 4-40 Rule viewer for the ANFIS model, training phase (S3)

Table 4-18 illustrates the goodness of fit between the predicted and observed data using this ANFIS method (S3) for the training and testing phases

Table 4-18 Statistical measures for the training and testing phases, ANFIS model (S3)

	Training Phase	**Testing Phase**
MAE	3.833	4.906
RMSE	5.311	6.770
NSE	0.544	0.536

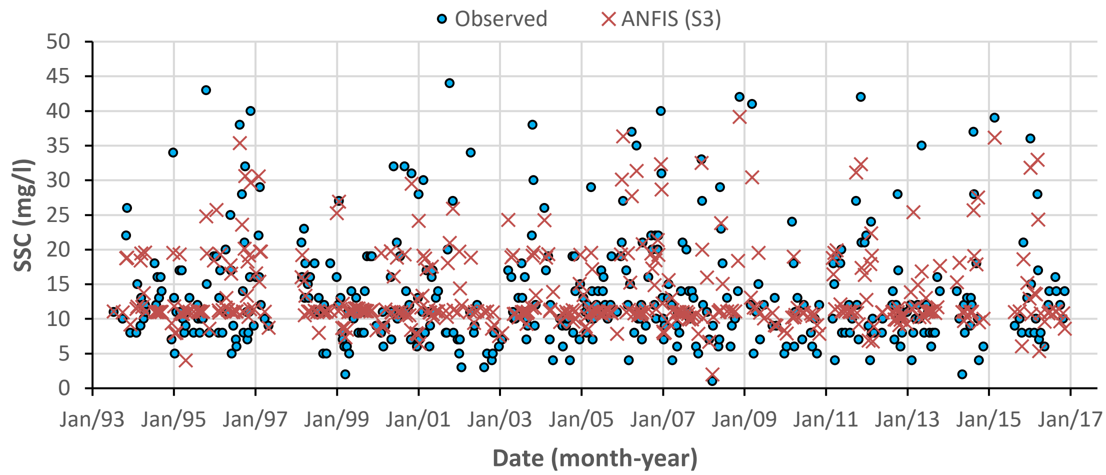


Figure 4-41 Observed and calculated SSC (mg/l), the training period using ANFIS (S3)

Figure 4-41 displays the observed and calculated SSC (mg/l) over the training phase period for the ANFIS approach (S3), and Figure 4-42 displays the extent of match between the measured and predicted SSC (mg/l) by the ANFIS model (S3) in terms of a scatter diagram type of comparison with respect to the testing data.

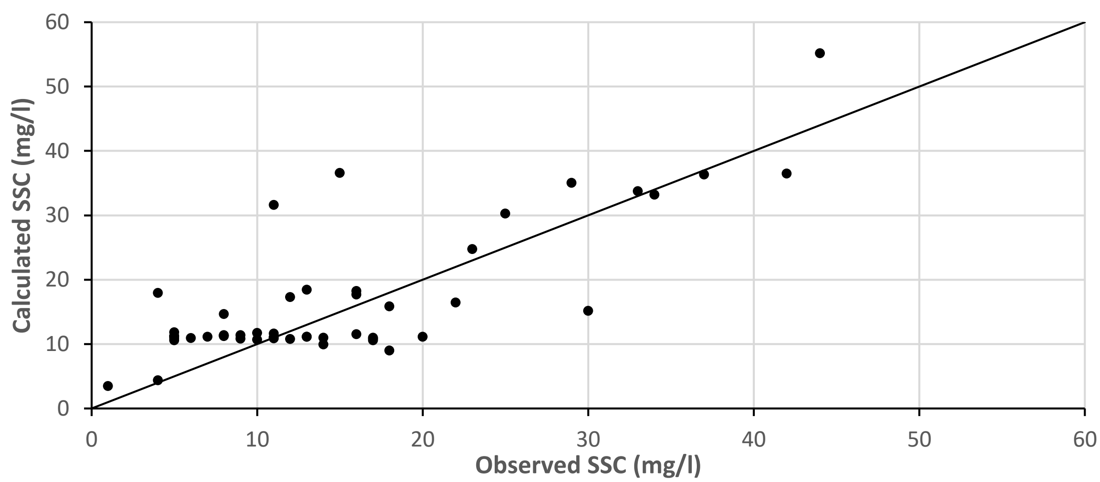


Figure 4-42 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S3). testing data

4.2.3.3 ANN Model

Different trails were performed using different number of neurons and different types of transfer functions. Only the best model's results will be shown here. In this case (S3) types TANSIG and PURELIN transfer functions were selected as transfer functions for hidden layer 1 and 2, respectively, 20 neurons were used in the hidden layer 1. Figure 4-43 presents the best network structure for S3 using all dataset.

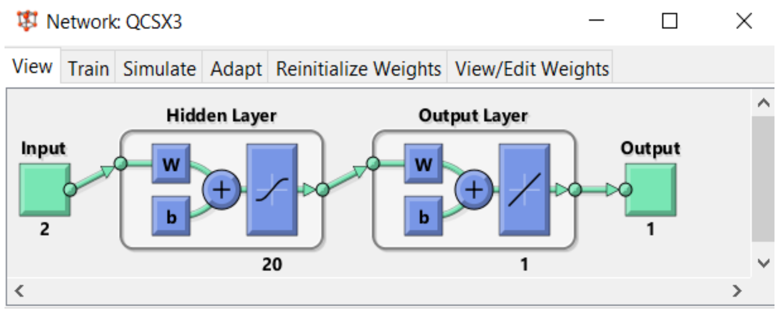


Figure 4-43 ANN best structure using all dataset (S3)

The model was trained using 85% of data in training purposes while the 15% of the entire set of the processed data was chosen for the testing determinations. After successfully achieving the best model, the calculated output is then extracted and used to calculate various performance indicators taking into consideration the same dataset chosen for training and testing purposes for the previous approaches to ensure fair comparison. Figure 4-32 displays the best model's outputs, after several epochs (trails) for the S3. Note that the model's built-in performance measure is R, which represents $\sqrt{R^2}$.

The training and testing phases' model performance measures are listed in Table 4-19, and Figures Figure 4-45 and Figure 4-46 display the observed and calculated SSC (mg/l) over the training phase period. and the extent of match between the measured and predicted SSC (mg/l) by the ANN model (S3) in terms of a scatter diagram type of comparison with respect to the testing data, respectively.

Table 4-19 Statistical measures for the training and testing phases, ANN model (S3)

	Training Phase	**Testing Phase**
MAE	3.586	3.473
RMSE	5.134	4.861
NSE	0.574	0.761

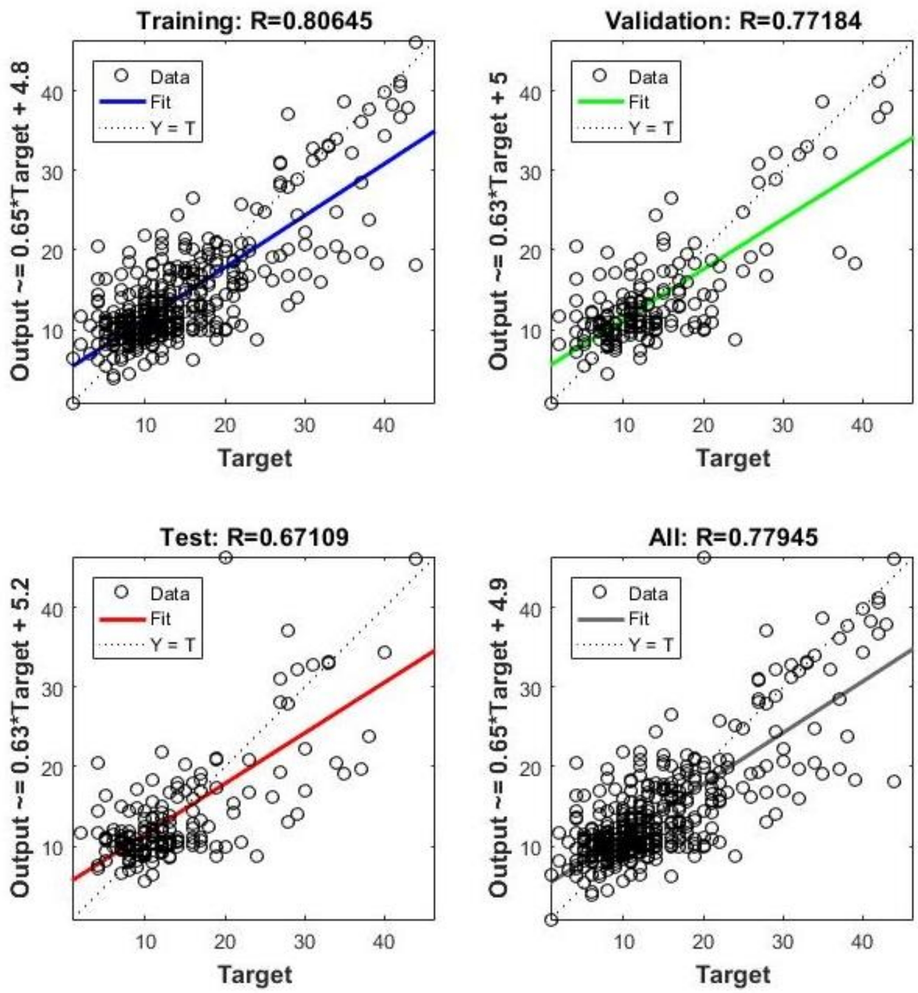


Figure 4-44 Best ANN model outputs (S3)

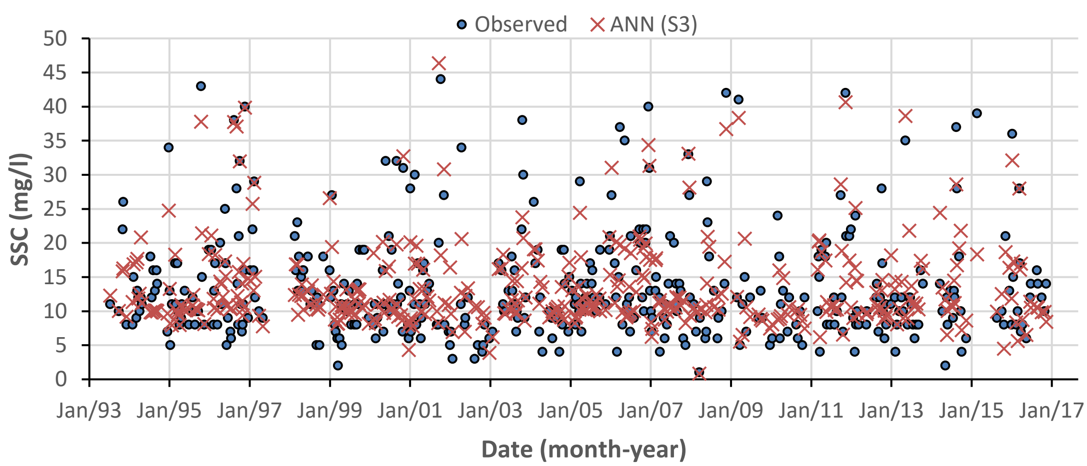


Figure 4-45 Observed and calculated SSC (mg/l), the training period using ANN (S3)

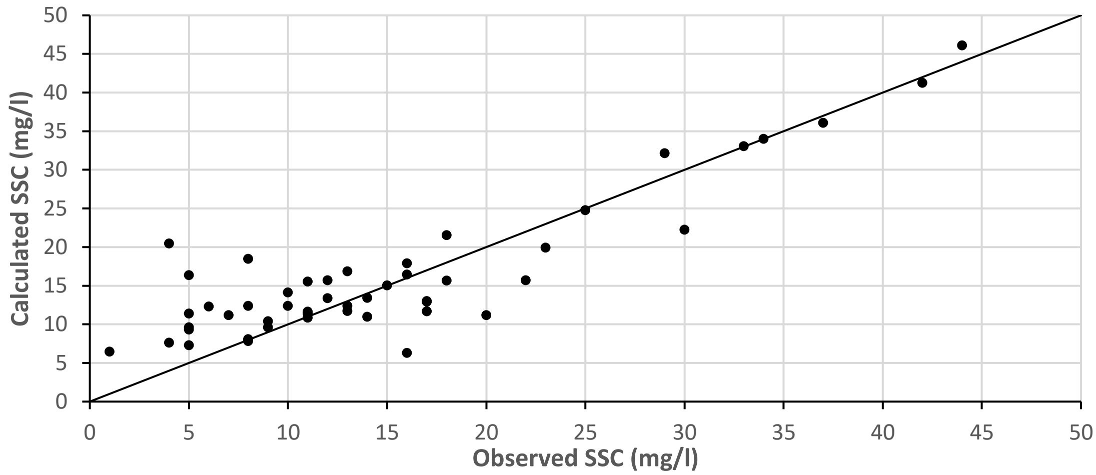
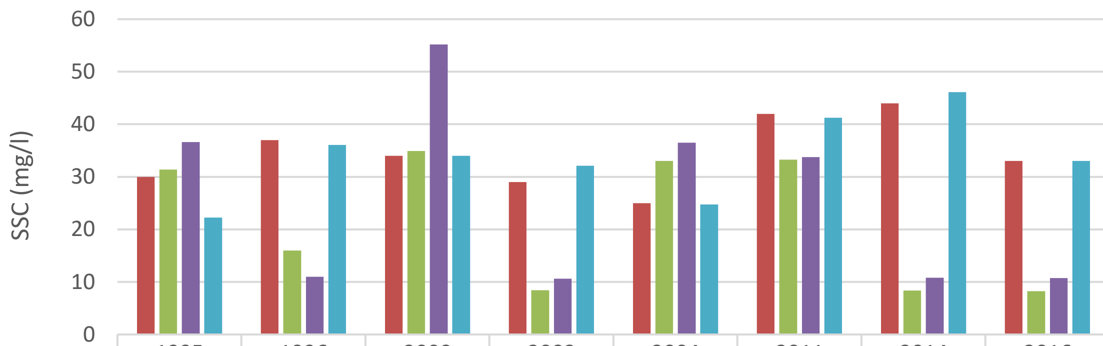


Figure 4-46 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S3), testing data

To sum up all the models for the S3 case, Table 4-20 shows the different performance indicators used in each case, and Figure 4-47 shows the observed SSC along with the calculated SSC (mg/l) of selected peaks from the testing phase using MLR, ANFIS and ANN modeling approaches for the S3 case. An excellence performance of the machine learning approaches, i.e. ANFIS and ANN, over the conventional method, i.e. MLR, is recognized. A superiority of the ANN approach is observed.

Table 4-20 Summary of different performance indicators for all models used using the S3 inputs

	MLR		**ANFIS**		**ANN**	
	Training Phase	Testing Phase	Training Phase	Testing Phase	Training Phase	Testing Phase
MAE	4.597	6.038	3.833	4.906	**3.586**	**3.473**
RMSE	6.244	7.555	5.311	6.770	**5.134**	**4.861**
NSE	0.370	0.423	0.544	0.536	**0.574**	**0.761**



	1995	1996	2000	2002	2004	2011	2014	2016
Observed	30.000	37.000	34.000	29.000	25.000	42.000	44.000	33.000
MLR	31.402	15.968	34.918	8.432	33.028	33.240	8.370	8.270
ANFIS	36.617	10.963	55.164	10.619	36.503	33.778	10.807	10.712
ANN	22.243	36.061	33.999	32.129	24.773	41.264	46.108	33.049

Figure 4-47 Selected peaks of observed SSC (mg/l) from the testing phase period for S3 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon

4.2.3.4 Uncertainty Analysis of S3 Results

Uncertainty analysis for all the four models developed for S3 has been performed according to the procedure detailed in the section 4.1.3.4 in this study. Results are given in Table 4-21 and it is observable that the three approaches (models) had an absolute mean prediction error of less than one-seventeenth with a superiority of the ANN model, which had an absolute mean prediction error of less than one-thirtieth order of magnitude. Meaning that its prediction is of a better performance among all. The uncertainty bands were similar for all approaches (±0.5 to ±0.6 log cycles) except for ANN, which had an uncertainty of ±0.2 log cycles.

Table 4-21 Uncertainty estimates for S3 various models

Approach	Mean prediction error log cycles	Width of uncertainty band, $\pm 2S_e$ log cycles	Prediction interval around hypothetical prediction value of $\hat{x}$ =1.0
MLR	0.058	± 0.571	0.306 – 4.258
ANFIS	0.056	± 0.595	0.290 – 4.476
ANN	0.033	± 0.249	0.608 – 1.911

4.2.4 Training Various Models for S4

Three different models were developed using various modeling techniques, namely, MLR, ANFIS and ANN. For these models, the temperature (T), the electrical conductivity (C) and the streamflow (Q) were used as inputs in order to model the targeted output (SSC).

4.2.4.1 MLR Model

Training dataset was used to train the MLR model and the regression add-in tool in the excel spreadsheet was used to train the data. With a confidence level of 95%, the regression significance F and Adjusted R Square were 2.21568E-46 and 0.3992, respectively, with an increase in the F significance value and an improving in the R adjusted value from what it was in the SLR case in the S1 (Section 4.2.1.2). The results summary of various regression measures is presented in Table 4-22.

Table 4-22 Outputs of the MLR performed using S4's training dataset

	Coefficients	Standard	t Stat	P-value
Intercept	3.423115474	2.5130741	1.362122778	0.17389622
T Variable 1	0.240861238	0.049745174	4.84190163	1.81651E-06
Q Variable 2	0.183900087	0.012289191	14.96437665	8.19838E-41
C Variable 3	0.001102287	0.002946908	0.374048723	0.708558639

From Table 4-22, the SSC equation can be written as follows:

$$SSC_{MLR} = 0.2409 * T + 0.1839 * Q + 0.0011 * C + 3.4231 \quad (36)$$

Applying the Equation (36) (exact coefficient of β_0, β_1, β_2 and β_3 were used) to the training and testing datasets in order to determine the various statistical measures representing the goodness of fit between the observed and the calculated SSC (mg/l) . Excel spreadsheets was used to perform these calculations, and the Table 4-23 shows the values of each measure using MLR model (S4),

Table 4-23 Statistical measures for the training and testing phases, MLR model (S4)

	Training Phase	Testing Phase
MAE	4.413	5.614
RMSE	6.076	7.269
NSE	0.403	0.466

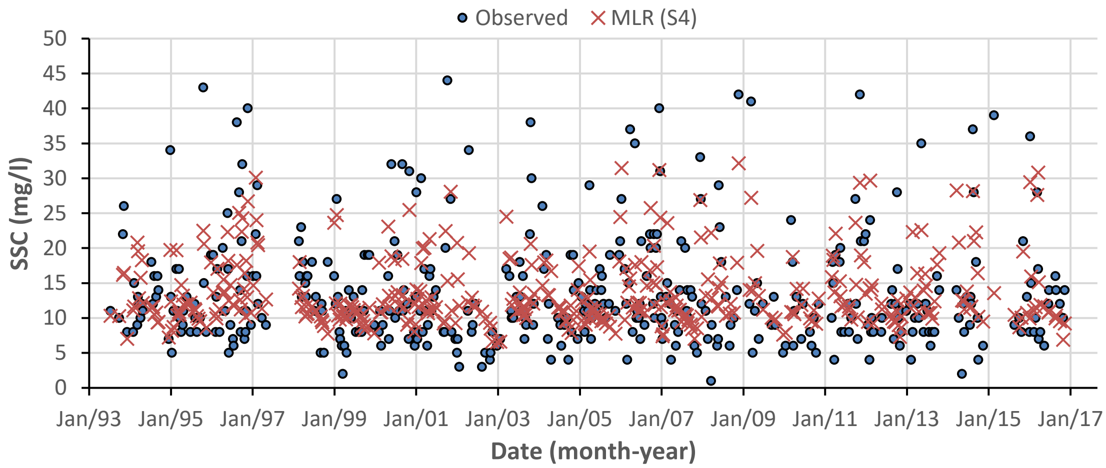


Figure 4-48 Observed and calculated SSC (mg/l), the training period using MLR (S4)

Figure 4-48 presents the observed and calculated SSC (mg/l) over the training phase period, and Figure 4-49 displays the extent of match between the measured and predicted SSC (mg/l) by the MLR model (S4) in terms of a scatter diagram type of comparison with respect to the testing data.

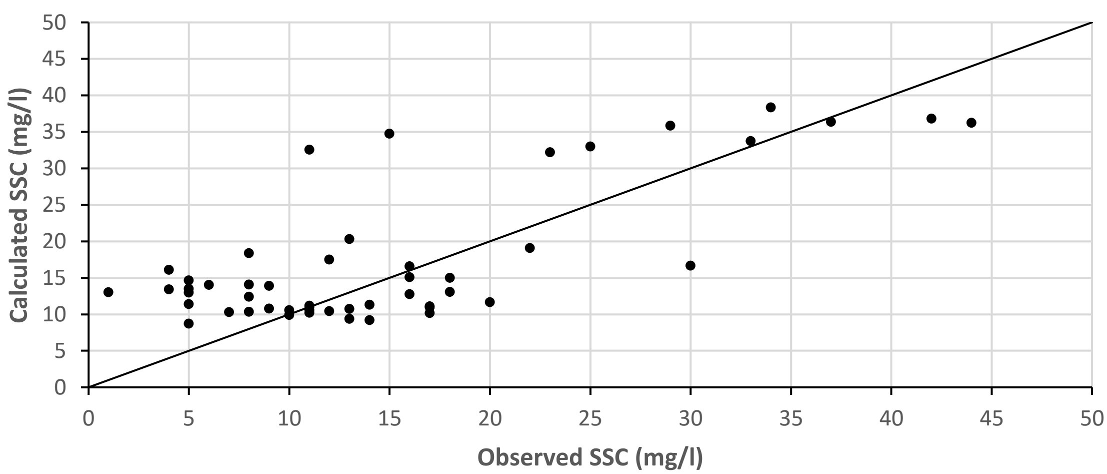


Figure 4-49 Scatter plot comparing predicted and observed SSC (mg/l) using MLR S(4), testing data

4.2.4.2 ANFIS Model

Several models were trained using the ANFISEDIT toolbox in MATLAB R2016b, using different numbers and structures of membership functions (MF). Hybrid optimization learning method used to train the various FIS. MF type constant was selected for the output. Only the best model's results will be presented here. In this case (S4) the GAUSS MF type using 3 MFs for T input, 2 MFs for Q input and 3 MFs for the C inputs, was the one that best modeled the phenomena. Figure 4-50, Figure 4-51 and Figure 4-52 present the model structure and designing toolbox, the membership functions for the three inputs and the rules of the ANFIS model for S4 training phase, respectively.

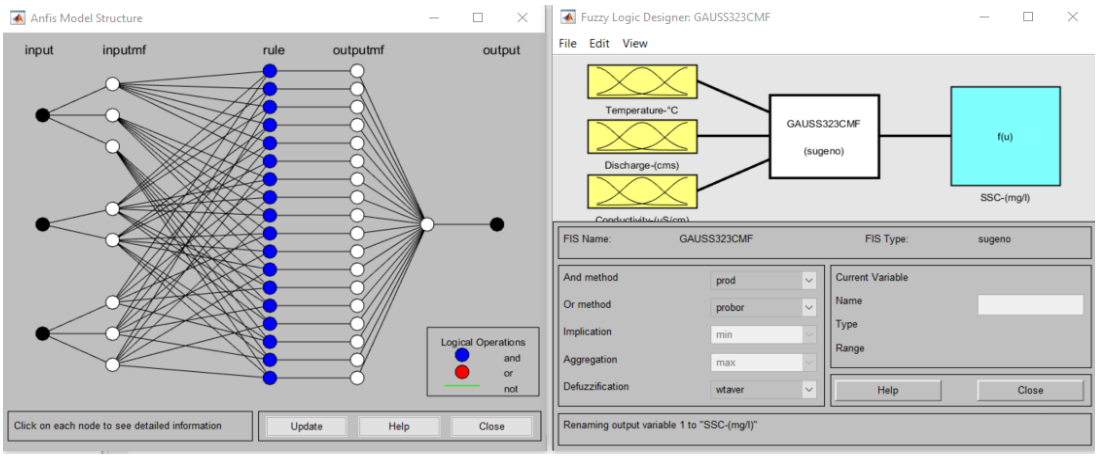


Figure 4-50 ANFIS model structure and Fuzzy logic designer toolbox for training phase (S4)

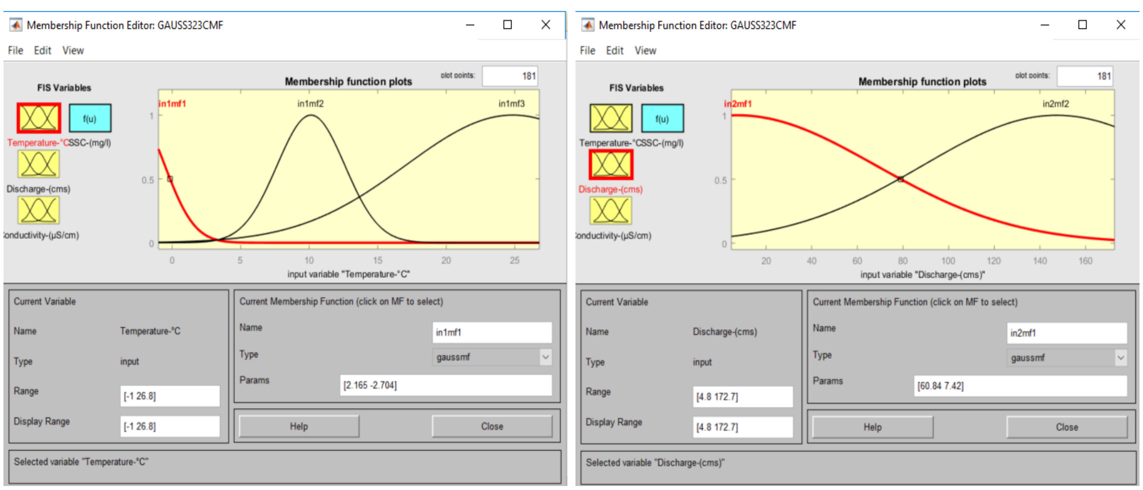


Figure 4-51 MFs editor for two of the three main inputs T (on the left-hand side) and Q for the ANFIS model, training phase (S4)

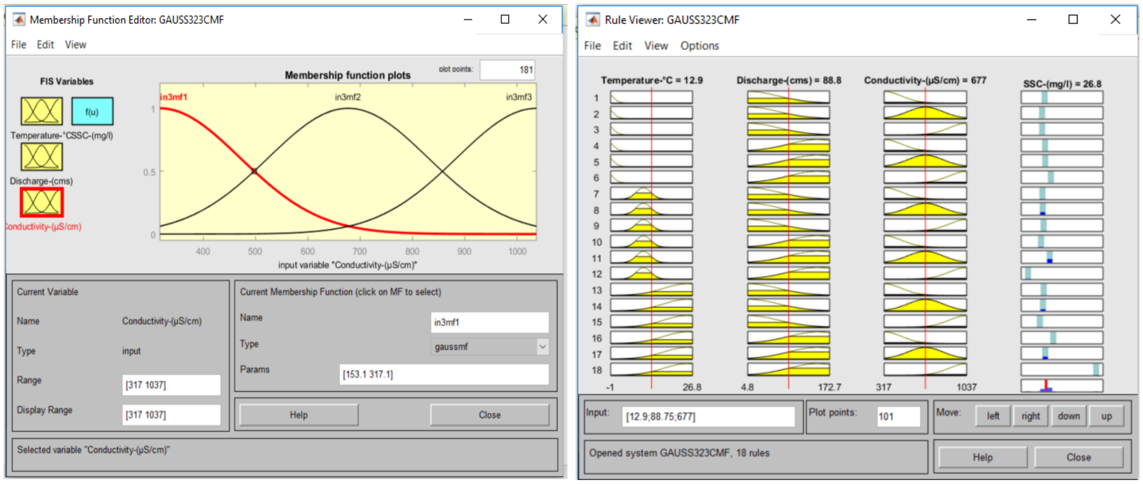


Figure 4-52 MFs editor for the input C (on the left-hand side) and Rule viewer for the ANFIS model, training phase (S4)

The training and testing phases' model performance measures are listed in Table 4-24, for the purposes of determining the goodness of fit between the observed and calculated SSC (mg/l) for ANFIS model (S4).

Table 4-24 Statistical measures for the training phase of the ANFIS model (S4)

	Training Phase	Testing Phase
MAE	4.061	5.752
RMSE	5.666	7.082
NSE	0.481	0.493

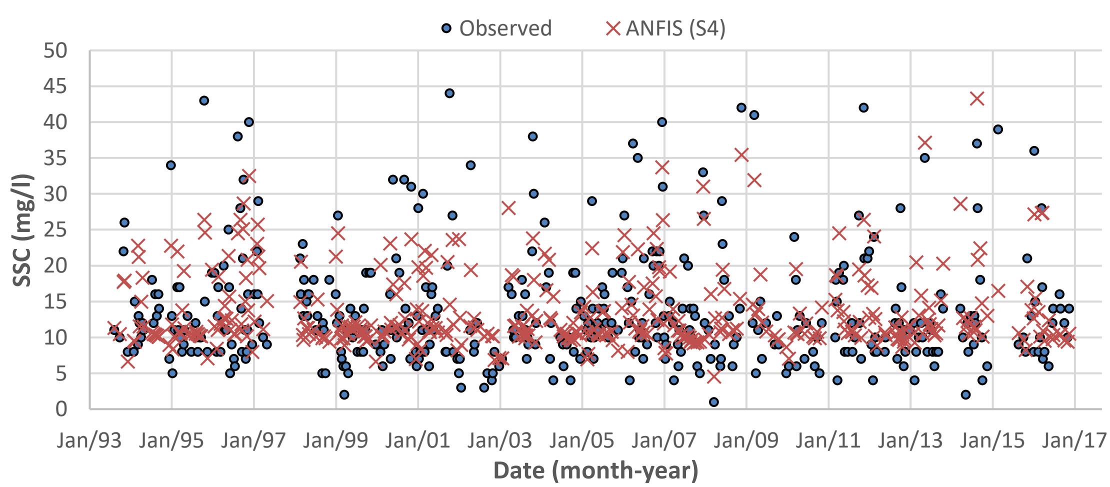


Figure 4-53 Observed and calculated SSC (mg/l), the training period using ANFIS (S4)

Figure 4-53 displays the observed and calculated SSC (mg/l) over the training phase period for the ANFIS approach (S4),and Figure 4-42 displays the extent of match between the measured and predicted SSC (mg/l) by the ANFIS model (S4) in terms of a scatter diagram type of comparison with respect to the testing data.

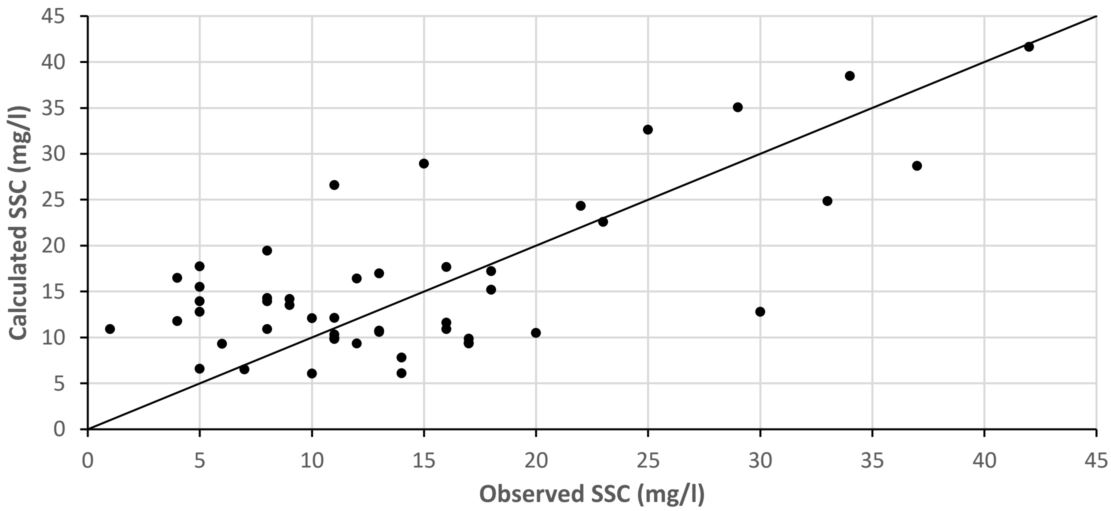


Figure 4-54 Scatter plot comparing predicted and observed SSC (mg/l) using ANFIS (S4), testing data

4.2.4.3 ANN Model

Different trails were performed using different number of neurons and different types of transfer functions. Only the best model's results will be shown here. In this case (S4) types TANSIG and PURELIN transfer functions were selected as transfer functions for hidden layer 1 and 2, respectively, 20 neurons were used in the hidden layer 1. Figure 4-55 presents the best network structure for S4 using all dataset.

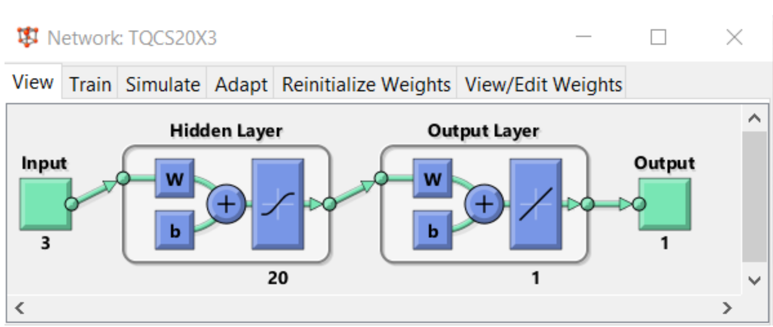


Figure 4-55 ANN best structure using all dataset (S4)

The model was trained using 85% of data in training purposes while the 15% of the entire set of the processed data was chosen for the testing determinations. After successfully achieving the best model, the calculated output is then extracted and used to calculate various performance indicators taking into consideration the same dataset chosen for training and testing purposes for the previous approaches to ensure fair comparison. Figure 4-32 displays the best model's outputs, after several epochs (trails) for the S4. Note that the model's built-in performance measure is R, which represents $\sqrt{R^2}$.

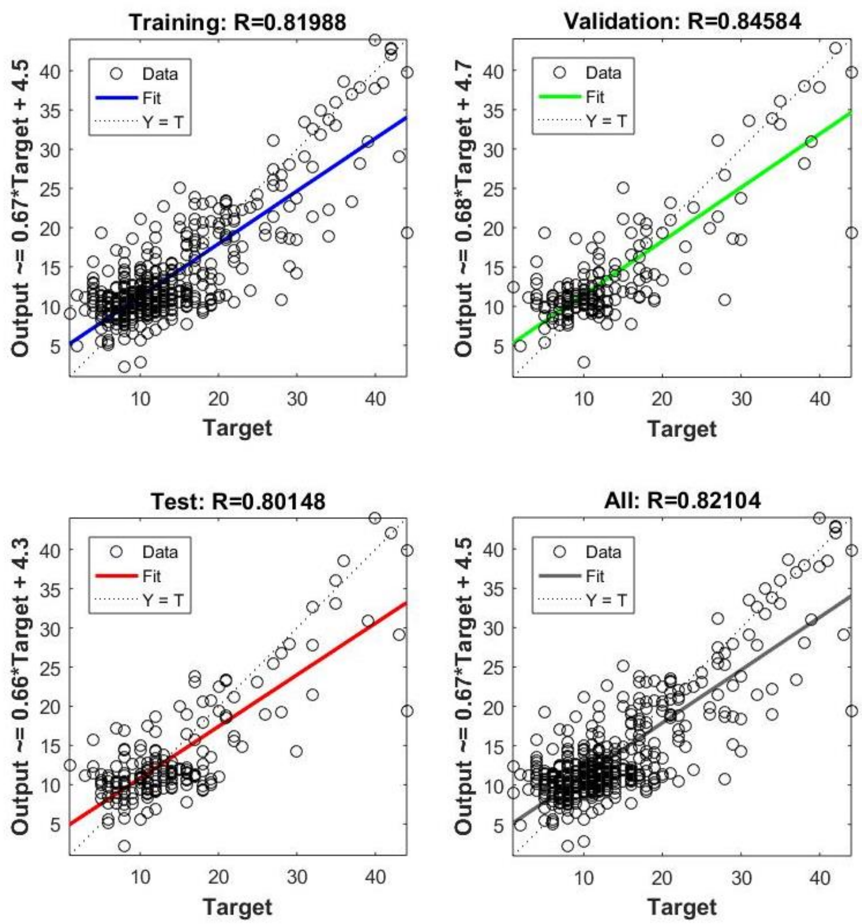


Figure 4-56 Best ANN model outputs (S4)

The training and testing phases' model performance measures are listed in Table 4-25, for the purposes of determining the goodness of fit between the observed and calculated SSC (mg/l) for ANFIS model (S4).

Table 4-25 Statistical measures for the training phase of the ANN model (S4)

	Training Phase	Testing Phase
MAE	3.456	2.823
RMSE	4.736	3.720
NSE	0.638	0.860

Figure 4-57 displays the observed and calculated SSC (mg/l) over the training phase period for the ANN approach (S4),and Figure 4-58 displays the extent of match between the measured and predicted SSC (mg/l) by the ANN model (S4) in terms of a scatter diagram type of comparison with respect to the testing data.

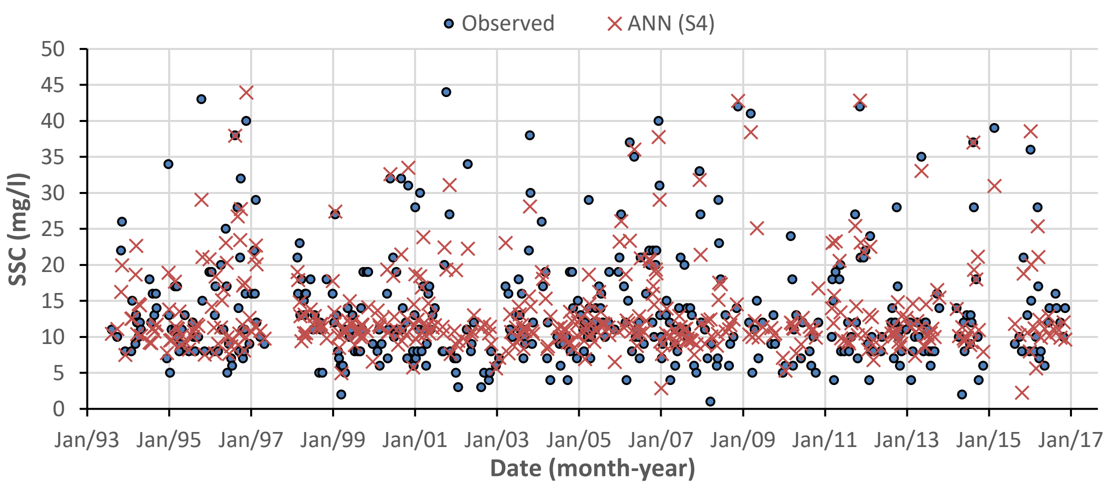


Figure 4-57 Observed and calculated SSC (mg/l), the training period using ANN (S4)

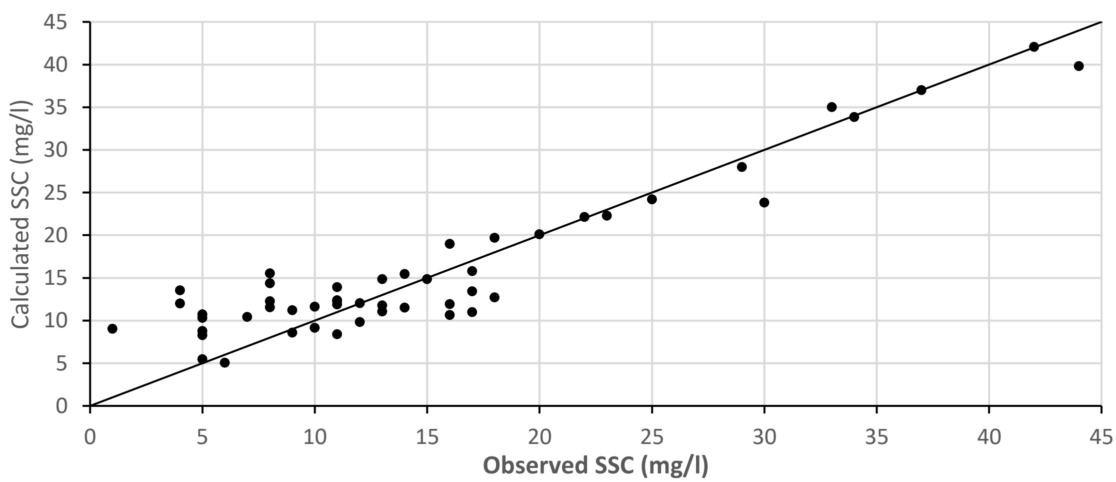
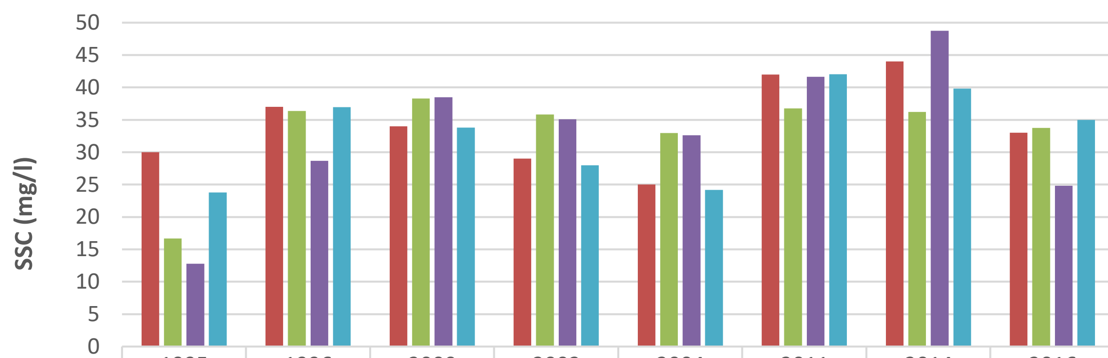


Figure 4-58 Scatter plot comparing predicted and observed SSC (mg/l) using ANN (S4), testing data

To sum up all the models for the S4 case, Table 4-26 shows the different performance indicators used in each case, and Figure 4-59 shows the observed SSC along with the calculated SSC (mg/l) of selected peaks from the testing phase using MLR, ANFIS and ANN modeling approaches for the S4 case. An excellence performance of the machine learning approaches, i.e. ANFIS and ANN, over the conventional method, i.e. MLR, is recognized. A superiority of the ANN approach is observed.

Table 4-26 Summary of different performance indicators for all models used using the (S4) inputs

	MLR		**ANFIS**		**ANN**	
	Training Phase	Testing Phase	Training Phase	Testing Phase	Training Phase	Testing Phase
MAE	4.413	5.614	4.061	5.752	**3.456**	**2.823**
RMSE	6.076	7.269	5.666	7.082	**4.736**	**3.720**
NSE	0.403	0.466	0.481	0.493	**0.638**	**0.860**



	1995	1996	2000	2002	2004	2011	2014	2016
Observed	30.000	37.000	34.000	29.000	25.000	42.000	44.000	33.000
MLR	16.667	36.346	38.311	35.821	32.956	36.776	36.211	33.746
ANFIS	12.771	28.677	38.468	35.062	32.604	41.640	48.764	24.833
ANN	23.800	36.977	33.817	27.974	24.191	42.038	39.813	34.985

Figure 4-59 Selected peaks of observed SSC (mg/l) from the testing phase period for S4 and the calculated SSC (mg/l) using various approaches to simulate the phenomenon

4.2.4.4 Uncertainty Analysis of S4 Results

Uncertainty analysis for all the four models developed for S4 has been performed according to the procedure detailed in the section 4.1.3.4 in this study. Results are given in an absolute mean prediction error of less than one-thirteenth with a superiority of the ANN model, which had an absolute mean prediction error of less than one-hundredth order of magnitude. Meaning that its prediction is of a better performance among all. The uncertainty bands were similar for MLR and ANFIS approaches (±0.4 log cycles) except for ANN, which had an uncertainty of ±0.2 log cycles.

and it is observable that the three approaches (models) an absolute mean prediction error of less than one-thirteenth with a superiority of the ANN model, which had an absolute mean prediction error of less than one-hundredth order of magnitude. Meaning that its prediction is of a better performance among all. The uncertainty bands were similar for MLR and ANFIS approaches (±0.4 log cycles) except for ANN, which had an uncertainty of ±0.2 log cycles.

Table 4-27 Uncertainty estimates for S4 various models

Approach	**Mean prediction error log cycles**	**Width of uncertainty band, ±2S_e log cycles**	**Prediction interval around hypothetical prediction value of $\hat{x}$ =1.0**
MLR	0.074	± 0.418	0.453 – 3.099
ANFIS	0.045	± 0.461	0.348 – 3.209
ANN	0.009	± 0.210	0.628 – 1.656

CHAPTER 5
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Summary

The main aim of this study was to develop the best model that estimates the Suspended Sediment Concentration (SSC) for the River Thames, London, Canada That is because the SSC is a site-specific phenomenon that ought to be modeled and estimated for every creek, stream, and river. Reliable estimation of such a phenomenon are of importance in many aspects, such as, Water resources management, hydraulic designs, environmental conservation, reservoir operation, river navigation and hydro-electric power generation. Two machine learning approaches, namely, Adaptive Neuro Fuzzy Inference System(ANFIS) and Artificial Neural Network (ANN) were developed to estimate the SSC and their performances were compared with the widely used conventional approaches, that is, Sediment Rating Curves (SRC) and Linear Regression (LR) models. In order to achieve this aim, four different scenarios were proposed in this study using different combinations of effective inputs (river discharge (Q), river temperature (T), and water electric conductivity (C)) collected over the period between 1993 and 2016. This data was used to train and test the various models for each scenario. Three main performance indicators were used to evaluate the performance of each model, namely, Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Nash-Sutcliffe Efficiency (NSE). Uncertainty analysis was also conducted to test the efficiency of each model separately. The final architecture of the best models developed for the ANFIS and ANN approaches for each scenario after several trials are illustrated in Table 5-1. Results from these best models have been compared with those found by developing the other two conventional approaches (SRC and LR).

Table 5-1 Final structure of various machine learning approaches' best model

Scenario No.	Scenario Inputs	ANFIS MFs Type	Number of ANFIS MFs	ANN Structure
S1	Q	GBELL	(9)	(1, 2, 1)
S2	T and Q	GAUSS2	(5, 2)	(2, 2, 1)
S3	Q and C	GBELL	(5, 5)	(2, 2, 1)
S4	T, Q and C	GAUSS	(3, 2, 3)	(3, 2, 1)

5.2 Conclusions

After developing various models for each scenario, a summary of the calculated performance indicators for each model in each scenario is given in Table 5-2. The following conclusions can be drawn from this study.

- A superiority of the machine learning approaches over the conventional SRC and linear regression models is observed. In scenario S1, NSE increased by more than 130% and 190% when comparing the performances of SRC with ANFIS and ANN, respectively.

- This increasing in performance was also noticeable after adding more effective inputs in scenarios S2, S3 and S4. An increase of more than 220% in the NSE indicator was achieved when comparing SRC with ANN models of S2 and S3.

- The best performance in estimating the suspended sediment concentration was accomplished by considering all inputs concerning this study (i.e. river discharge, river temperature, and water electric conductivity) with NSE of 86%.

- A quantitative analysis of the uncertainty of different conventional and machine learning approaches conducted in this study was also carried out and Table 5-3 gives the uncertainty estimates for each of the proposed scenario. Mean prediction error of the ANN model was the best among all other models, ranging from 1/30 to 1/200 order of magnitude. All in all, machine learning approaches have shown a better performance in estimating the suspended sediment concentration for the River Thames, London, Canada, than the conventional approaches. ANN model of S4 have shown the best performance considering the uncertainty analysis parameters.

- The least accuracy in predicting the suspended sediment concentration in this study was achieved when using the conventional SRC, followed by the linear regression models.

Table 5-2 Performance indicators for various models

Scenario No.		S1	S2	S3	S4
Scenario Inputs		Q	T and Q	Q and C	T, Q and C
SRC	MAE	6.936	-	-	-
	RMSE	8.709	-	-	-
	NSE	0.233	-	-	-
SLR	MAE	5.997	-	-	-
	RMSE	7.563	-	-	-
	NSE	0.421	-	-	-
MLR	MAE	-	5.641	6.038	5.614
	RMSE	-	7.266	7.555	7.269
	NSE	-	0.466	0.423	0.466
ANFIS	MAE	5.194	5.420	4.906	5.752
	RMSE	6.738	6.813	6.770	7.082
	NSE	0.541	0.531	0.536	0.493
ANN	MAE	4.250	3.590	3.473	2.823
	RMSE	5.579	4.869	4.861	3.720
	NSE	0.685	0.760	0.761	0.860

Table 5-3 Uncertainty estimates for various models

Scenario	Approach	Mean prediction error log cycles	Width of uncertainty band, $\pm 2S_e$ log cycles	Prediction interval around hypothetical prediction value of $\hat{x}$ =1.0
S1	SRC	– 0.067	± 0.442	0.310 – 2.375
	SLR	0.023	± 0.427	0.395 – 2.822
	ANFIS	0.064	± 0.369	0.496 – 2.710
	ANN	0.005	± 0.334	0.470 – 2.183
S2	MLR	0.073	± 0.420	0.450 – 3.116
	ANFIS	0.047	± 0.422	0.422 – 2.943
	ANN	0.015	± 0.313	0.503 – 2.126
S3	MLR	0.058	± 0.571	0.306 – 4.258
	ANFIS	0.056	± 0.595	0.290 – 4.476
	ANN	0.033	± 0.249	0.608 – 1.911
S4	MLR	0.074	± 0.418	0.453 – 3.099
	ANFIS	0.045	± 0.461	0.348 – 3.209
	ANN	0.009	± 0.210	0.628 – 1.656

5.3 Future Work Recommendations

The following further investigations are recommended:

- A continuous daily time-series sampling for SSC is advised to enhance the ability of various approaches to model the SSC phenomenon.

- The effect of other useful input variables (e.g. rainfall intensity) might improve the capability of various models in modeling the SSC for river Thames, London, Canada.

- Application of other techniques, such as Wavelet and Genetic Programing, that have been utilized in water resources problems especially in the areas of developing and improving optimization algorithms, could be investigated.

References

Adriaenssens, V., De Baets, B., Goethals, P. L. M., & De Pauw, N. (2004). Fuzzy rule-based models for decision support in ecosystem management. *Science of the Total Environment*, *319*(1–3), 1–12. https://doi.org/10.1016/S0048-9697(03)00433-9

Afaghi, M., Ramaswamy, H. S., & Prasher, S. O. (2001). Thermal process calculations using artificial neural network models. *Food Research International*, *34*(1), 55–65. https://doi.org/10.1016/S0963-9969(00)00132-0

Aksoy, H., & Dahamsheh, A. (2009). Artificial neural network models for forecasting monthly precipitation in Jordan. *Stochastic Environmental Research and Risk Assessment*, *23*(7), 917–931. https://doi.org/10.1007/s00477-008-0267-x

Anderson, P. G., Taylor, B. R., & Balch, G. C. (1996). Quantifying the effects of sediment release on fish and their habitats. *Canadian Manuscript Report of Fisheries and Aquatic Sciences*.

Angabini, S., Ahmadi, H., Feiznia, S., Vaziri, B. M., & Ershadi, S. (2014). Using Intelligence Models to Estimate Suspended sediment system case study : Jagin Dam. *Bulletin of Environment, Pharmacology and Life Sciences*, *3*(III), 166–172.

Aytek, A., & Kişi, Ö. (2008). A genetic programming approach to suspended sediment modelling. *Journal of Hydrology*, *351*(3–4), 288–298. https://doi.org/10.1016/j.jhydrol.2007.12.005

Banzhaf, W., Francone, F. D., Keller, R. E., & Nordin, P. (1998). *Genetic programming: an introduction on the automatic evolution of computer programs and its applications*. San Francisco, CA: MORGAN KAUFMANN PUBL Inc.

Bayram, A., Kankal, M., & Önsoy, H. (2012). Estimation of suspended sediment concentration from turbidity measurements using artificial neural networks. *Environmental Monitoring and Assessment*, *184*(7), 4355–4365. https://doi.org/10.1007/s10661-011-2269-2

Bayram, A., Kankal, M., Tayfur, G., & Önsoy, H. (2013). Prediction of suspended sediment concentration from water quality variables. *Neural Computing and Applications*, *24*(5), 1079–1087. https://doi.org/10.1007/s00521-012-1333-3

Bentaher, L. S., & Elmazoghi, H. G. (2013). ESTIMATION OF DAM BREACH WIDTHS USING A NEURO-FUZZY COMPUTING TECHNIQUE. In *Seventeenth International Water Technology Conference*. Istanbul,.

Birtwell, I. (1999). *The effects of sediment on fish and their habitat*. *Canadian Stock Assessment Secretariat Research Document 99/139*. West Vancouver, B.C.

Burke, E. K., & Kendall, G. (2005). *Search methodologies: introductory tutorials in optimization and decision support techniques*. Springer.

Campbell, F. B., & Bauder, H. A. (1940). A rating-curve method for determining silt-discharge of streams. *Eos, Transactions American Geophysical Union*, *21*(2), 603–607. https://doi.org/10.1029/TR021i002p00603

Caux, P. Y., Moore, D. R. J., & MacDonald, D. (1997). Ambient water quality guidelines (criteria) for turbidity, suspended and benthic sediments: technical appendix. Prepared for the British Columbia Ministry of Environment, Lands and Parks. Victoria, BC: Water Quality Branch, Environment and Resource Management Division.

Cigizoglu, H., & Alp, M. (2003). Suspended Sediment Forecasting by Artificial Neural Networks Using Hydro Meteorological Data. *World Water & Environmental Resources Congress 2003*, (90), 1–8. https://doi.org/doi:10.1061/40685(2003)173\r10.1061/40685(2003)173

Cigizoglu, H. K. (2003). Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Advances in Water Resources*, *27*, 185–195. https://doi.org/10.1016/j.jhydrol.2008.06.013

Cigizoglu, H. K., & Kisi, Ö. (2003). Flow prediction by three back propagation techniques using k -fold partitioning of neural network training data. *Nordic*

Hydrology, *36*(1), 49–64.

Cobaner, M., Unal, B., & Kisi, O. (2009). Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *Journal of Hydrology*, *367*(1–2), 52–61. https://doi.org/10.1016/j.jhydrol.2008.12.024

da Silva, I. N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L. H. B., & dos Reis Alves, S. F. (2017). Artificial Neural Networks, 21–29. https://doi.org/10.1007/978-3-319-43162-8

Dai, Q., Shan, H., Jia, Y., & Cui, W. (2009). Laboratory Study on the Relationships Between Suspended Sediment Concentration and Electrical Conductivity. In *ASME 2009 28th International Conference on Ocean, Offshore and Arctic Engineering : Volume 7* (pp. 179–186). Honolulu, Hawaii, USA. https://doi.org/10.1115/OMAE2009-79211

Dawson, C. W., Harpham, C., Wilby, R. L., & Chen, Y. (2002). Evaluation of artificial neural network techniques for flow forecasting in the River Yangtze, China. *Hydrology and Earth System Sciences*, *6*(4), 619–626. https://doi.org/10.5194/hess-6-619-2002

Demirci, M., & Baltaci, A. (2013). Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. *Neural Computing and Applications*, *23*(S1), 145–151. https://doi.org/10.1007/s00521-012-1280-z

European Inland Fisheries Advisory Commission (EIFAC). (1964). *Water quality criteria for European freshwater fish. Report on Finely Divided Solids and Inland Fisheries*. *EUROPEAN INLAND FISHERIES ADVISORY COMMISSION*. Rome.

Firat, M., & Güngör, M. (2010). Monthly total sediment forecasting using adaptive neuro fuzzy inference system. *Stochastic Environmental Research and Risk Assessment*, *24*(2), 259–270. https://doi.org/10.1007/s00477-009-0315-1

Firat, M., Yurdusev, M. A., & Turan, M. E. (2009). Evaluation of artificial neural network techniques for municipal water consumption modeling. *Water Resources Management*, *23*(4), 617–632. https://doi.org/10.1007/s11269-008-9291-3

Foroozesh, J., Khosravani, A., Mohsenzadeh, A., & Mesbahi, A. H. (2013). Application of Artificial Intelligence (AI) Modeling in Kinetics of Methane Hydrate Growth. *American Journal of Analytical Chemistry*, *4*(11), 616–622. https://doi.org/10.4236/ajac.2013.411073

Foster, I. D. L., & Charlesworth, S. M. (1996). Heavy metals in the hydrological cycle: trends and explanation. *Hydrological Processes*, *10*(2), 227–261. https://doi.org/10.1002/(SICI)1099-1085(199602)10:2<227::AID-HYP357>3.0.CO;2-X

Garg, V. (2011). Modeling catchment sediment yield: A genetic programming approach. *Natural Hazards*, *70*(1), 39–50. https://doi.org/10.1007/s11069-011-0014-3

Ghorbani, M. A., Hosseini, S. H., Fazelifard, M. H., & Abbasi, H. (2013). Sediment load estimation by MLR, ANN, NF and Sediment Rating Curve (SRC) in Rio Chama River. *Journal of Civil Engineering and Urbanism*, *3*(4), 136–141.

Government of Canada, E. and C. C. C. (2013). Water – The Transporter. Retrieved August 22, 2017, from https://www.ec.gc.ca/eau-water/default.asp?lang=En&n=ADB791B6-1

Grubbs, F. (1969). Procedures for Detecting Outlying Observations in Samples. *Technometrics*, *11*(1), 1–21.

Guven, A., & Kişi, Ö. (2011). Estimation of Suspended Sediment Yield in Natural Rivers Using Machine-coded Linear Genetic Programming. *Water Resources Management*, *25*(2), 691–704. https://doi.org/10.1007/s11269-010-9721-x

Haykin, S. (2008). *Neural Networks and Learning Machines*. *Pearson Prentice Hall New Jersey USA 936 pLinks* (Vol. 3). https://doi.org/978-0131471399

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological TheoryNo*

Title. New York, NY: John Wiley & Sons.

Heng, S., & Suetsugi, T. (2013). Using Artificial Neural Network to Estimate Sediment Load in Ungauged Catchments of the Tonle Sap River Basin, Cambodia. *J. Water Resour. Prot.*, *5*(2), 111–123. https://doi.org/10.4236/jwarp.2013.52013

Hollis, E. H., Boone, J. G., De Rose, C. R., & Murphy, G. J. (1964). *Literature review of the effects of turbidity and siltation on aquatic life*. Annapolis, Maryland.

Jain, S. K. (2001). Development of Integrated Sediment Rating Curves Using ANNs. *JOURNAL OF HYDRAULIC ENGINEERING*, *127*(1), 30–37.

Jang, J.-S. R. (1993). ANFIS: Adaptive-Network_Based Fuzzy Inference system. *IEEE*, *23*(3), 665–685.

Jang, J.-S. R., & Sun, C.-T. (1995). Neuro-Fuzzy Modeling and Control. *Proceedings of the Institution of Electrical Engineers IEEE*, *83*(3), 378–406.

Jang, J. R., Sun, C.-T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Upper Saddle River, NJ: Prentice Hall Inc.

Jeong, D.-I., & Kim, Y.-O. (2005). Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. *Hydrological Processes*, *19*(19), 3819–3835. https://doi.org/10.1002/hyp.5983

Joshi, R., Kumar, K., & Adhikari, V. P. S. (2015). Modelling suspended sediment concentration using artificial neural networks for Gangotri glacier. *Hydrological Processes*, *1366*(November 2015), 1354–1366. https://doi.org/10.1002/hyp.10723

Karim, M. F., & Kennedy, J. F. (1990). Menu of Coupled Velocity and Sediment-Discharge Relations for Rivers, *116*(8), 978–996.

Kisi, O. (2005). Suspended sediment estimation using neuro-fuzzy and neural network approaches. *Hydrological Sciences–Journal–des Sciences Hydrologiques*, *50*(4).

https://doi.org/10.1623/hysj.2005.50.4.683

Kisi, Ö. (2004). Multi-layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation / Prévision et estimation de la concentration en matières en suspension avec des perceptrons multi-couches et l'algorithme d. *Hydrological Sciences Journal*, *49*(6). https://doi.org/10.1623/hysj.49.6.1025.55720

Kisi, Ö. (2007). Streamflow Forecasting Using Different Artificial Neural Network Algorithms. *JOURNAL OF HYDROLOGIC ENGINEERING*, *12*(5), 532–539.

Kisi, O., Dailr, A. H., Cimen, M., & Shiri, J. (2012). Suspended sediment modeling using genetic programming and soft computing techniques. *Journal of Hydrology*, *450–451*, 48–58. https://doi.org/10.1016/j.jhydrol.2012.05.031

Kisi, O., Haktanir, T., Ardiclioglu, M., Ozturk, O., Yalcin, E., & Uludag, S. (2009). Adaptive neuro-fuzzy computing technique for suspended sediment estimation. *Advances in Engineering Software*, *40*(6), 438–444. https://doi.org/10.1016/j.advengsoft.2008.06.004

Kisi, O., Karahan, M. E., & Şen, Z. (2006). River suspended sediment modelling using a fuzzy logic approach. *Hydrological Processes*, *20*(20), 4351–4362. https://doi.org/10.1002/hyp.6166

Kisi, O., & Shiri, J. (2012). River suspended sediment estimation by climatic variables implication: Comparative study among soft computing techniques. *Computers and Geosciences*, *43*, 73–82. https://doi.org/10.1016/j.cageo.2012.02.007

Kisi, O., & Zounemat-Kermani, M. (2016). Suspended Sediment Modeling Using Neuro-Fuzzy Embedded Fuzzy c-Means Clustering Technique. *Water Resources Management*, *30*(11), 3979–3994. https://doi.org/10.1007/s11269-016-1405-8

Kizhisseri, A., Simmonds, D., Rafiq, Y., & Borthwick, M. (2005). An Evolutionary Computation Approach to Sediment Transport Modelling. *Coastal Dynamics 2005*, (1963), 1–14. https://doi.org/doi:10.1061/40855(214)81\r10.1061/40855(214)81

Kondolf, G. M. (1997). Hungry water: Effects of dams and gravel mining on river channels. *Environmental Management*, *21*(4), 533–551. https://doi.org/10.1007/s002679900048

Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection*. Cambridge, Mass: MIT press.

Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the Artificial Neural Networks. *European Journal of Gastroenterology & Hepatology*, *19*(12), 1046–1054. https://doi.org/10.1097/MEG.0b013e3282f198a0

Lewis-Beck, M. S., Bryman, A., & Liao, T. F. (2004). *The Sage encyclopedia of social science research methods*. Sage.

Lipták, B. G. (2002). *Instrument engineers' handbook: Process software and digital networks* (3rd ed.). CRC Press.

Lloyd, D. S. (1987). Turbidity as a Water Quality Standard for Salmonid Habitats in Alaska. *North American Journal of Fisheries Management*, *7*(1), 34–45. https://doi.org/10.1577/1548-8659(1987)7<34:TAAWQS>2.0.CO;2

Lohani, A. K., Goel, N. K., & Bhatia, K. K. S. (2007). Deriving stage–discharge–sediment concentration relationships using fuzzy logic. *Hydrological Sciences Journal*, *52*(4), 793–807. https://doi.org/10.1623/hysj.52.4.793

Lopes, V. L., & Ffolliott, P. F. (1993). SEDIMENT RATING CURVES FOR A CLEARCUT PONDEROSA PINE WATERSHED IN NORTHERN ARIZONA. *WATER RESOURCES BULLETIN, AMERICAN WATER RESOURCES ASSOCIATION*, *29*(3), 1:14.

Mamdani, E. H. (1974). Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers*, *121*(12), 1585–1588. https://doi.org/10.1049/piee.1974.0328

McBean, E. A., & AI-Nassr, S. (1988). Uncertainty in Suspended Sediment Transport

Curves. *Journal of Hydraulic Engineering*, *114*(1), 63–74.

Melesse, A. M., Ahmad, S., McClain, M. E., Wang, X., & Lim, Y. H. (2011). Suspended sediment load prediction of river systems: An artificial neural network approach. *Agricultural Water Management*, *98*(5), 855–866. https://doi.org/10.1016/j.agwat.2010.12.012

Mkpenie, V. N., Ebong, G., & Abasiekong, B. (2007). Studies on the effect of temperature on the sedimentation of insoluble metal carbonates 1*. *J. Appl. Sci. Environ. Manage. December*, *11*(4), 67–69. Retrieved from www.bioline.org.br/ja

Morehead, M. D., Syvitski, J. P., Hutton, E. W. H., & Peckham, S. D. (2003). Modeling the temporal variability in the flux of sediment from ungauged river basins. *Global and Planetary Change*, *39*(1–2), 95–110. https://doi.org/10.1016/S0921-8181(03)00019-5

Morgan, R. P. C. (2009). *Soil Erosion and Conservation* (3., Auflag). New York, NY: John Wiley & Sons.

Morris, G., Annandale, G., & Hotchkiss, R. (2008). “Reservoir Sedimentation,” in Sedimentation Engineering: Processes, Measurements, Modeling, and Practice. *American Society of Civil Engineers*.

Nagy, H., Watanabe, K., & Hirano, M. (2002). Prediction of Sediment Load Concentration in Rivers using Artificial Neural Network Model. *Journal of Hydraulic Engineering*, *128*(6), 588–595. https://doi.org/10.1061/(asce)0733-9429(2002)128:6(588)

Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting Through Conceptual Models Part I-a Discussion of Principles*. *Journal of Hydrology*, *10*, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Nastos, P., Moustris, K., Larissi, I., & Paliatsos, A. (2011). Air Quality and Bioclimatic Conditions within the Greater Athens Area , Greece - Development and Applications of Artificial Neural Networks. In F. Nejadkoorki (Ed.), *Advanced Air*

Pollution. InTech.

Newcombe, C. P., & Macdonald, D. D. (1991). Effects of Suspended Sediments on Aquatic Ecosystems. *North American Journal of Fisheries Management*, *11*(1).

Olyaie, E., Banejad, H., Chau, K.-W., & Melesse, A. M. (2015). A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. *Environmental Monitoring and Assessment*, *187*(4), 189. https://doi.org/10.1007/s10661-015-4381-1

Ongley, E. D., Krishnappan, B. G., Droppo, G., Rao, S. S., & Maguire, R. J. (1992). Cohesive sediment transport: emerging issues for toxic chemical management. *Hydrobiologia*, *235–236*(1), 177–187. https://doi.org/10.1007/BF00026210

Owens, P. N., Batalla, R. J., Collins, A. J., Gomez, B., Hicks, D. M., Horowitz, A. J., … Trustrum, N. A. (2005). Fine-grained sediment in river systems: environmental significance and management issues. *River Research and Applications*, *21*(7), 693–717. https://doi.org/10.1002/rra.878

Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). *Field Guide to Genetic Programming*. Morrisville, NC: Lulu Press. Retrieved from http://digitalcommons.morris.umn.edu/cs_facpubs%0Ahttp://lulu.com%0Ahttp://www.gp-field-guide.org.uk

Rajaee, T., Mirbagheri, S. A., Zounemat-Kermani, M., & Nourani, V. (2009). Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Science of the Total Environment*, *407*(17), 4916–4927. https://doi.org/10.1016/j.scitotenv.2009.05.016

Rezaei, M., & Fereydooni, M. (2015). COMPARATIVE EVALUATION OF ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS) AND ARTIFICIAL NEURAL NETWORK (ANN) IN SIMULATION OF SUSPENDED SEDIMENT LOAD (CASE STUDY : DALAKI RIVER , CHAM CHIT STATION). *Indian Journal of Fundamental and Applied Life Sciences*, *5*(S1), 3598–3606.

Rojas, R. (1996). *Neural Networks: A systematic introduction*. New York;Berlin: Springer-Verlag. https://doi.org/10.1016/0893-6080(94)90051-5

Rousseeuw, P. J. (1998). Chapter 17: Robust estimation and identifying outliers. In *Handbook of statistical methods for engineers and scientists* (2nd Ed., p. 17.1-17.15). New York, NY: McGraw–Hill.

Shamaei, E., & Kaedi, M. (2016). Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions. *Applied Soft Computing Journal*, *45*, 187–196. https://doi.org/10.1016/j.asoc.2016.03.009

Srinivasulu, S., & Jain, A. (2006). A comparative analysis of training methods for artificial neural network rainfall–runoff models. *Applied Soft Computing*, *6*(3), 295–306. https://doi.org/10.1016/j.asoc.2005.02.002

Steven Sobieszczyk, Heather M. Bragg, and M. A. U. (2015). *Water-Quality Conditions and Suspended-Sediment Transport in the Wilson and Trask Rivers , Northwestern Oregon , Water Years 2012 – 14 Scientific Investigations Report 2015 – 5109*. Reston, Virginia. https://doi.org/http://dx.doi.org/10.3133/sir20155109.

Suparta, W., & Alhasa, K. M. (2016). *Modeling of Tropospheric Delays Using ANFIS*. Bangi: Springer. https://doi.org/10.1007/978-3-319-28437-8

Tachi, S. E. (2017). *Contribution to the characterization and the modelling of sediment transport in urban hydro- systems*. University of Hassiba Benbouali, Chlef.

Tahmoures, M., Nia, A. R. M., & Naghiloo, M. (2015). Modeling of streamflow-suspended sediment load relationship by adaptive neuro-fuzzy and artificial neural network approaches (Case study : Dalaki River , Iran). *DESERT*, *2*, 177–195.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, *SMC-15*(1), 116–132. https://doi.org/10.1109/TSMC.1985.6313399

Tuan, L. T., & Shibayama, T. (2003). Application of Gis To Evaluate Long-Term Variation of Sediment Discharge To Coastal Environment. *Coastal Engineering Journal*, *45*(2), 275–293. https://doi.org/10.1142/S0578563403000774

Tyrrell, J. L. (2015). the Effect of Water Temperature on in-Stream Sediment Concentration and Transport Rate.

Uri, N. D. (1999). The environmental implications of soil erosion in the United States. *Environmental Monitoring and Assessment*, *66*(3), 293–312. https://doi.org/10.1023/A:1006333329653

UTRCA. (2012). *River Bend, Watershed Report Card*. London.

Wahl, T. (2004). Uncertainty of Predictions of Embankment Dam Breach Parameters. *Journal of Hydraulic Engineering*, *130*(5), 389–397. https://doi.org/10.1061/(ASCE)0733-9429(2004)130:5(389)

Wang, H., Yang, Z., Wang, Y., Saito, Y., & Liu, J. P. (2008). Reconstruction of sediment flux from the Changjiang (Yangtze River) to the sea since the 1860s. *Journal of Hydrology*, *349*(3–4), 318–332. https://doi.org/10.1016/j.jhydrol.2007.11.005

Waters, T. (1995). Sediment in streams. Sources, biological effects, and control. *AMERICAN FISHERIES SOCIETY, BETHESDA, MD (USA)*, 251.

Williams, G. P. (1989). Sediment concentration versus water discharge during single hydrologic events in rivers. *Journal of Hydrology*, *111*(1–4), 89–106. https://doi.org/10.1016/0022-1694(89)90254-0

Wolman, M., & Schick, A. (1967). Eects of Construction on Fluvial Sediment, Urban and Suburban Areas of Marylan. *Water Resources Research*, *3*(2), 451–464. Retrieved from https://books.google.ca/books?hl=en&lr=&id=kTSBAAAAQBAJ&oi=fnd&pg=PT140&dq=wolman+1967+sediment&ots=KhTw_7J2u-&sig=Vm1h55l7ESrr-d8xd9id1EAmkME

Wood, P. J., & Armitage, P. D. (1997). Biological Effects of Fine Sediment in the Lotic Environment. *Environmental Management*, *21*(2), 203–217.

Yang, G., Chen, Z., Yu, F., Wang, Z., Zhao, Y., & Wang, Z. (2007). Sediment rating parameters and their implications: Yangtze River, China. *Geomorphology*, *85*(3–4), 166–175. https://doi.org/10.1016/j.geomorph.2006.03.016

Zadeh, L. a. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zhu, Y. M., Lu, X. X., & Zhou, Y. (2007). Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology*, *84*(1–2), 111–125. https://doi.org/10.1016/j.geomorph.2006.07.010

Appendices

Appendix A: Raw data table used in this study

Date	Time	River Temperature (degrees Celsius)	Flow (m^3/sec)	Conductivity (µS/cm)	Suspended Solids (mg/L)
8-09-93	9:25 AM	18.5	9.73	590	11
22-09-93	11:00 AM	15.7	10.40	641	15
29-09-93	11:00 AM	14.4	11.70	681	13
6-10-93	11:00 AM	13.3	14.30	603	14
20-10-93	11:00 AM	12.1	20.10	648	13
27-10-93	11:00 AM	11.9	17.60	686	10
3-11-93	10:40 AM	7.8	18.00	674	52
11-11-93	10:55 AM	7.5	16.70	698	21
17-11-93	10:40 AM	7.8	15.40	720	17
24-11-93	12:10 PM	5.8	17.20	732	18
30-11-93	11:50 AM	3.9	60.10	702	22
7-12-93	11:45 AM	5.3	59.80	698	26
5-01-94	12:00 PM	1.8	12.40	825	8
23-02-94	11:28 AM	2.0	151.00	481	23
1-03-94	11:20 AM	2.6	34.30	653	8
9-03-94	10:30 AM	1.8	38.70	616	15
16-03-94	10:25 AM	2.1	69.50	581	76
23-03-94	11:55 AM	3.3	260.00	491	124
30-03-94	10:55 AM	4.2	151.00	515	17
6-04-94	11:40 AM	5.8	83.40	571	9
13-04-94	11:25 AM	9.7	71.30	644	13
21-04-94	11:35 AM	10.7	32.60	764	12
4-05-94	11:20 AM	11.4	48.40	629	10
17-05-94	11:15 AM	11.8	61.50	641	12
24-05-94	11:25 AM	16.7	21.70	665	11
8-06-94	11:25 AM	19.2	16.60	680	45
22-06-94	11:00 AM	25.4	10.20	705	4
6-07-94	11:45 AM	23.8	17.20	686	19
20-07-94	12:00 PM	23.9	16.30	682	41
3-08-94	11:35 AM	23.6	14.50	658	14
10-08-94	11:20 AM	20.7	12.80	683	18
24-08-94	10:55 AM	22.4	10.90	693	12
6-09-94	11:05 AM	18.9	9.50	691	16
21-09-94	11:10 AM	20.8	8.57	697	13
28-09-94	11:30 AM	19.1	12.00	668	25
5-10-94	10:45 AM	14.0	9.78	677	16
12-10-94	11:20 AM	14.9	9.17	699	14
19-10-94	11:35 AM	15.4	12.70	667	13

26-10-94	11:05 AM	10.4	9.14	724	79
9-11-94	11:20 AM	12.1	21.70	643	78
30-11-94	11:10 AM	5.0	18.20	736	18
6-12-94	11:05 AM	8.8	14.20	738	49
14-12-94	11:50 AM	2.7	16.90	819	11
4-01-95	12:05 PM	2.5	12.20	771	18
11-01-95	11:50 AM	3.8	10.50	864	7
18-01-95	11:40 AM	4.9	211.00	514	26
25-01-95	11:40 AM	4.2	79.40	640	34
1-02-95	11:30 AM	7.3	34.60	752	13
8-02-95	11:10 AM	2.2	17.70	888	5
14-02-95	11:30 AM	2.4	14.80	915	4
23-02-95	11:35 AM	4.2	18.30	832	11
1-03-95	11:50 AM	3.9	18.00	1,089	13
15-03-95	10:50 AM	6.0	218.00	420	23
22-03-95	11:35 AM	6.7	76.00	605	17
29-03-95	11:35 AM	8.7	22.10	705	11
3-04-95	11:05 AM	4.6	22.00	730	7
12-04-95	11:40 AM	10.4	22.60	718	17
19-04-95	1:35 PM	11.4	19.20	722	10
26-04-95	11:30 AM	9.0	56.90	555	30
3-05-95	11:00 AM	11.7	42.30	631	8
17-05-95	1:35 PM	18.1	27.50	317	9
24-05-95	11:20 AM	17.6	19.10	675	11
7-06-95	11:50 AM	23.2	23.10	578	16
20-06-95	1:15 PM	25.3	9.57	678	13
28-06-95	10:50 AM	22.6	53.10	498	56
5-07-95	12:05 PM	25.0	15.10	633	20
12-07-95	11:20 AM	23.1	9.72	696	9
26-07-95	11:45 AM	26.0	9.79	664	8
9-08-95	11:35 AM	23.8	12.20	586	13
16-08-95	11:35 AM	26.9	30.40	544	23
30-08-95	11:45 AM	24.0	7.37	648	12
6-09-95	11:40 AM	23.7	7.05	620	10
20-09-95	11:40 AM	18.6	8.89	664	8
4-10-95	11:00 AM	17.7	8.00	693	10
11-10-95	11:40 AM	16.3	11.50	623	9
18-10-95	11:40 AM	13.5	11.60	670	10
1-11-95	11:15 AM	10.7	16.30	659	70
8-11-95	11:15 AM	6.7	37.70	639	36
15-11-95	11:00 AM	5.1	93.40	596	43
22-11-95	11:10 AM	5.4	82.30	671	15
29-11-95	11:10 AM	3.8	257.00	552	36
13-12-95	10:45 AM	1.6	21.30	788	8
3-01-96	11:15 AM	1.5	14.40	798	3

10-01-96	11:00 AM	1.5	12.40	873	5
17-01-96	11:15 AM	3.0	34.60	1,074	67
24-01-96	11:05 AM	3.2	70.60	716	19
31-01-96	11:20 AM	1.4	43.00	638	8
7-02-96	11:10 AM	1.8	18.90	742	1
14-02-96	11:30 AM	1.4	50.10	568	19
22-02-96	1:35 PM	1.9	155.00	429	62
28-02-96	1:45 PM	2.6	116.00	457	50
6-03-96	11:05 AM	2.2	31.70	712	8
13-03-96	11:25 AM	4.0	26.00	714	13
20-03-96	11:20 AM	3.8	65.00	668	17
3-04-96	10:55 AM	6.5	44.60	628	9
10-04-96	10:55 AM	6.0	30.80	643	8
17-04-96	11:10 AM	5.6	239.00	532	82
24-04-96	10:55 AM	8.1	165.00	570	37
1-05-96	10:40 AM	8.8	151.00	514	49
8-05-96	11:05 AM	12.2	35.70	626	20
15-05-96	11:05 AM	11.7	46.80	600	11
21-05-96	11:10 AM	18.2	204.00	409	221
29-05-96	11:25 AM	16.8	30.00	611	11
5-06-96	10:40 AM	17.5	21.90	649	4
19-06-96	10:20 AM	20.2	72.70	576	25
26-06-96	11:15 AM	20.3	50.00	603	17
3-07-96	11:05 AM	22.3	16.40	656	5
10-07-96	11:15 AM	21.9	14.70	616	18
17-07-96	11:30 AM	23.8	26.60	605	9
31-07-96	10:40 AM	22.0	10.90	697	14
7-08-96	10:50 AM	25.3	9.11	665	7
14-08-96	11:10 AM	24.7	8.95	662	6
10-09-96	11:35 AM	19.2	49.50	386	38
18-09-96	11:20 AM	17.6	37.60	547	90
2-10-96	11:05 AM	16.1	93.20	574	28
10-10-96	11:10 AM	13.5	28.00	684	14
16-10-96	10:45 AM	13.5	20.70	668	8
24-10-96	11:05 AM	10.7	85.30	621	21
31-10-96	11:15 AM	9.0	98.00	618	32
6-11-96	10:45 AM	8.8	47.20	678	32
13-11-96	11:40 AM	5.0	47.80	696	11
20-11-96	10:35 AM	4.4	60.50	719	7
27-11-96	10:50 AM	3.0	28.10	774	8
4-12-96	10:45 AM	5.0	69.80	674	11
11-12-96	10:55 AM	4.5	41.90	710	16
17-12-96	10:40 AM	5.6	115.00	706	40
7-01-97	10:55 AM	1.6	184.00	533	33
15-01-97	11:00 AM	1.5	33.80	675	9

21-01-97	11:00 AM	1.4	30.80	744	3
29-01-97	1:00 PM	3.0	43.70	692	5
5-02-97	10:50 AM	2.4	53.80	705	16
12-02-97	11:10 AM	2.0	35.20	684	3
19-02-97	10:50 AM	4.0	41.60	868	22
26-02-97	11:05 AM	2.0	139.00	560	22
5-03-97	10:20 AM	3.5	104.00	564	16
12-03-97	10:50 AM	3.2	86.80	592	29
19-03-97	10:45 AM	2.9	85.10	598	12
9-04-97	10:50 AM	5.5	34.10	636	143
16-04-97	11:15 AM	9.2	28.10	620	6
23-04-97	11:05 AM	9.8	26.40	620	10
30-04-97	10:50 AM	13.8	21.80	617	6
7-05-97	11:40 AM	11.6	105.00	550	34
28-05-97	11:05 AM	16.7	25.40	486	9
18-03-98	10:25 AM	4.7	47.90	762	21
25-03-98	10:40 AM	4.8	69.10	672	16
9-04-98	10:05 AM	9.6	35.90	500	23
15-04-98	10:45 AM	10.1	27.90	500	13
22-04-98	11:10 AM	13.8	31.90	500	18
6-05-98	10:45 AM	17.0	18.50	520	5
14-05-98	10:30 AM	18.3	9.54	590	15
20-05-98	10:15 AM	22.7	11.50	600	13
27-05-98	10:50 AM	19.5	9.78	590	16
3-06-98	10:15 AM	19.7	9.50	670	16
17-06-98	11:00 AM	21.5	8.40	610	6
8-07-98	11:00 AM	22.6	13.20	360	24
15-07-98	10:30 AM	25.6	7.58	600	18
22-07-98	10:30 AM	24.8	8.37	800	6
5-08-98	10:15 AM	23.4	6.18	700	11
12-08-98	12:50 AM	24.1	13.50	400	16
19-08-98	12:45 PM	23.9	7.08	400	13
2-09-98	10:20 AM	24.1	5.10	720	17
9-09-98	10:30 AM	21.0	5.17	710	11
16-09-98	10:10 AM	21.6	6.05	740	11
23-09-98	11:35 AM	20.6	5.03	780	17
30-09-98	10:15 AM	19.8	6.60	740	5
7-10-98	9:45 AM	16.7	7.72	720	12
14-10-98	11:00 AM	15.2	8.49	680	11
21-10-98	10:50 AM	14.1	5.79	730	5
28-10-98	10:50 AM	14.4	6.52	690	5
4-11-98	10:10 AM	8.8	6.84	670	28
11-11-98	10:30 AM	9.1	11.30	680	47
18-11-98	10:20 AM	8.0	6.64	710	21
25-11-98	10:20 AM	7.9	7.33	670	6

2-12-98	10:20 AM	9.2	7.87	620	18
9-12-98	10:25 AM	7.0	8.70	680	14
27-01-99	10:25 AM	1.7	105.00	460	16
3-02-99	10:30 AM	2.4	45.70	590	18
10-02-99	10:00 AM	3.0	40.30	490	12
18-02-99	10:50 AM	2.9	109.00	500	27
24-02-99	10:40 AM	1.7	18.70	460	16
3-03-99	10:45 AM	3.3	29.80	590	13
10-03-99	10:45 AM	2.9	13.80	470	17
17-03-99	9:55 AM	5.7	23.90	510	8
24-03-99	10:20 AM	4.3	38.80	420	7
7-04-99	10:35 AM	10.3	22.40	540	6
14-04-99	11:15 AM	10.9	19.70	470	2
21-04-99	10:20 AM	11.2	18.20	530	2
28-04-99	12:15 PM	12.8	21.90	510	6
5-05-99	10:25 AM	17.6	15.80	520	14
12-05-99	10:25 AM	16.6	9.72	570	11
19-05-99	10:25 AM	20.2	6.62	650	5
26-05-99	10:45 AM	15.5	11.30	540	7
2-06-99	10:50 AM	20.2	20.70	530	17
9-06-99	10:30 AM	24.6	9.65	580	14
23-06-99	10:40 AM	23.0	5.87	620	13
28-06-99	1:50 PM	26.1	8.37	610	10
7-07-99	2:50 PM	26.5	6.20	570	11
14-07-99	11:15 AM	24.2	6.05	540	11
21-07-99	10:10 AM	24.6	6.45	540	13
28-07-99	10:45 AM	26.5	5.51	510	5
4-08-99	10:35 AM	24.2	5.46	480	9
11-08-99	10:00 AM	22.2	6.21	570	8
18-08-99	1:15 AM	22.8	5.68	540	10
1-09-99	10:15 AM	21.7	5.49	600	8
8-09-99	10:15 AM	22.9	8.62	630	8
15-09-99	10:40 AM	20.6	6.32	580	11
22-09-99	10:00 AM	17.3	5.82	630	13
29-09-99	10:10 AM	20.1	7.56	600	8
6-10-99	1:45 AM	14.0	7.10	610	14
13-10-99	10:10 AM	16.0	8.87	610	11
19-10-99	10:20 AM	14.6	7.17	640	11
27-10-99	10:30 AM	10.9	8.52	610	19
3-11-99	10:15 AM	7.1	21.10	610	43
10-11-99	9:55 AM	11.1	19.30	560	48
17-11-99	10:05 AM	5.9	12.50	560	10
24-11-99	10:40 AM	12.2	9.56	660	19
1-12-99	10:00 AM	3.4	9.72	620	10
8-12-99	10:35 AM	5.9	31.40	747	19

15-12-99	9:55 AM	5.2	47.40	910	39
5-01-00	9:50 AM	1.7	47.20	777	25
12-01-00	10:20 AM	2.4	35.20	779	17
19-01-00	10:30 AM	1.1	17.50	809	9
26-01-00	10:30 AM	0.8	12.50	903	3
2-02-00	10:30 AM	1.2	9.50	1,117	4
8-02-00	9:40 AM	1.2	8.50	1,145	3
16-02-00	10:10 AM	3.4	8.45	1,280	10
23-02-00	10:30 AM	5.9	14.90	1,352	5
2-03-00	10:40 AM	4.6	69.00	635	8
8-03-00	10:05 AM	8.3	29.80	754	6
15-03-00	10:00 AM	7.4	25.80	805	11
22-03-00	10:50 AM	8.9	23.80	779	6
29-03-00	10:30 AM	8.2	25.90	773	5
6-04-00	11:50 AM	8.2	31.00	780	9
13-04-00	11:00 AM	7.2	23.90	726	13
19-04-00	10:40 AM	10.3	17.60	749	11
26-04-00	10:25 AM	11.2	46.90	653	15
3-05-00	10:20 AM	14.2	21.80	689	5
10-05-00	10:20 AM	18.4	27.50	588	27
25-05-00	10:15 AM	15.8	82.50	612	16
31-05-00	10:40 AM	18.4	25.00	688	7
5-06-00	10:35 AM	17.7	17.30	698	11
14-06-00	11:05 AM	19.1	279.00	524	49
22-06-00	10:10 AM	20.4	53.10	681	32
26-06-00	10:45 AM	21.9	157.00	672	34
6-07-00	10:50 AM	21.9	24.50	724	19
19-07-00	10:55 AM	19.9	50.80	594	21
26-07-00	10:35 AM	21.9	20.50	643	10
2-08-00	11:15 AM	22.0	208.00	537	152
9-08-00	10:55 AM	23.6	56.70	688	13
16-08-00	10:50 AM	22.5	22.20	697	19
23-08-00	10:25 AM	21.5	13.00	741	5
30-08-00	10:50 AM	22.9	11.30	696	11
6-09-00	10:20 AM	18.9	15.60	651	23
13-09-00	10:20 AM	20.7	13.90	675	11
20-09-00	10:40 AM	19.9	23.10	653	13
27-09-00	10:30 AM	15.0	58.20	652	32
4-10-00	11:15 AM	16.7	27.60	708	11
11-10-00	10:10 AM	12.6	23.10	741	14
18-10-00	10:10 AM	14.0	18.60	756	15
25-10-00	10:25 AM	14.2	14.60	744	10
1-11-00	9:50 AM	10.5	15.20	730	20
8-11-00	11:30 AM	11.6	10.80	757	12
15-11-00	10:20 AM	7.2	27.60	713	41

22-11-00	10:50 AM	3.0	25.00	815	7
29-11-00	11:10 AM	2.0	113.00	709	31
6-12-00	10:10 AM	0.0	22.00	844	23
13-12-00	11:00 AM	1.2	15.20	964	2
10-01-01	10:30 AM	1.1	18.50	834	8
17-01-01	11:20 AM	2.1	38.30	909	6
24-01-01	10:50 AM	2.5	22.50	842	13
31-01-01	11:20 AM	2.6	49.00	897	28
7-02-01	10:30 AM	2.7	42.50	773	7
14-02-01	10:40 AM	2.4	230.00	453	23
21-02-01	10:55 AM	1.0	39.00	676	8
28-02-01	10:40 AM	1.7	178.00	485	5
7-03-01	10:55 AM	3.4	48.00	683	11
14-03-01	10:30 AM	3.3	70.10	679	30
21-03-01	10:30 AM	5.2	104.00	685	22
28-03-01	10:20 AM	4.1	81.00	580	8
4-04-01	10:30 AM	6.1	110.00	570	5
11-04-01	10:50 AM	9.5	74.70	558	17
18-04-01	10:30 AM	8.6	42.80	633	8
25-04-01	10:05 AM	10.5	31.50	656	11
7-05-01	7:55 AM	15.6	16.90	687	6
14-05-01	8:10 AM	15.4	15.20	685	9
23-05-01	10:25 AM	15.6	19.20	639	14
28-05-01	7:55 AM	14.4	75.00	566	16
6-06-01	7:45 AM	14.8	19.40	601	17
13-06-01	10:45 AM	21.0	21.00	584	11
27-06-01	10:15 AM	23.5	15.20	593	17
4-07-01	11:00 AM	21.5	9.34	590	13
16-07-01	8:05 AM	23.1	6.06	660	28
23-07-01	7:35 AM	25.0	7.67	631	14
8-08-01	8:45 AM	26.8	5.66	642	20
13-08-01	9:50 AM	24.8	5.73	660	11
20-08-01	8:05 AM	21.5	7.90	658	26
5-09-01	10:25 AM	21.2	5.94	641	7
12-09-01	11:00 AM	22.7	6.60	622	12
17-09-01	7:55 AM	17.7	6.40	598	15
1-10-01	7:50 AM	15.0	8.33	573	8
10-10-01	10:20 AM	13.4	12.50	545	8
17-10-01	10:50 AM	11.5	85.30	519	20
22-10-01	8:05 AM	12.0	28.30	627	23
5-11-01	8:05 AM	8.7	49.80	583	44
12-11-01	8:05 AM	7.1	27.40	618	14
19-11-01	8:05 AM	10.9	22.30	624	12
26-11-01	8:00 AM	9.6	61.30	566	83
3-12-01	7:55 AM	7.7	121.00	503	27

10-12-01	8:05 AM	5.5	30.50	624	8
2-01-02	8:05 AM	-0.8	20.70	719	2
14-01-02	7:50 AM	2.5	24.30	693	7
21-01-02	8:05 AM	1.9	31.40	733	3
28-01-02	8:10 AM	3.6	85.50	671	5
4-02-02	8:30 AM	1.3	59.60	760	7
11-02-02	8:10 AM	1.6	59.10	741	18
18-02-02	8:15 AM	0.6	39.70	783	3
4-03-02	8:15 AM	1.3	171.00	579	29
11-03-02	8:10 AM	2.1	136.00	533	55
18-03-02	8:10 AM	5.5	43.70	622	5
25-03-02	7:50 AM	3.1	30.80	654	3
8-04-02	7:15 AM	7.1	56.10	584	4
15-04-02	8:10 AM	11.4	200.00	469	56
22-04-02	8:05 AM	11.0	28.70	628	7
6-05-02	7:40 AM	14.0	25.80	653	11
13-05-02	7:50 AM	11.6	67.50	600	34
27-05-02	7:45 AM	14.7	22.80	647	11
3-06-02	7:45 AM	17.8	21.50	622	13
10-06-02	7:40 AM	20.3	21.60	567	8
17-06-02	7:40 AM	17.6	18.60	628	9
24-06-02	8:00 AM	23.7	41.50	458	18
2-07-02	8:05 AM	26.0	13.90	499	4
8-07-02	8:10 AM	24.6	8.53	605	12
15-07-02	8:10 AM	24.3	7.67	672	12
22-07-02	7:20 AM	25.5	7.95	707	3
12-08-02	9:05 AM	24.6	8.56	653	16
19-08-02	7:30 AM	23.8	8.87	677	8
26-08-02	7:40 AM	21.9	8.78	670	1
9-09-02	7:40 AM	22.6	6.86	695	3
16-09-02	7:50 AM	21.0	8.34	692	5
23-09-02	7:35 AM	20.1	8.42	418	3
7-10-02	7:40 AM	17.2	8.66	654	5
21-10-02	7:40 AM	11.5	8.57	719	4
28-10-02	7:40 AM	10.6	8.85	687	5
18-11-02	7:40 AM	5.5	20.10	788	4
25-11-02	7:30 AM	6.3	15.10	691	5
2-12-02	7:45 AM	2.4	11.00	1,161	2
9-12-02	7:30 AM	1.6	9.88	912	8
6-01-03	7:40 AM	2.8	21.10	891	5
13-01-03	7:40 AM	0.9	14.30	965	5
20-01-03	7:50 AM	1.2	9.23	1,037	6
3-02-03	7:40 AM	3.4	10.60	1,341	4
10-02-03	8:50 AM	1.3	17.60	1,055	6
17-02-03	9:40 AM	0.5	10.70	991	7

24-02-03	9:20 AM	0.5	14.20	1,202	8
3-03-03	8:05 AM	0.1	11.7	1,659	4
10-03-03	7:35 AM	0.6	13.8	1,183	13
17-03-03	7:40 AM	5.0	36.0	986	34
24-03-03	7:50 AM	3.7	186	460	20
7-04-03	7:55 AM	3.8	59	781	5
14-04-03	7:45 AM	7.7	101	587	17
28-04-03	7:45 AM	13.6	21	714	6
5-05-03	7:45 AM	13.4	54.8	695	12
12-05-03	7:40 AM	13.1	60.4	698	16
26-05-03	7:40 AM	14.9	59.2	676	11
2-06-03	7:55 AM	15.3	21.5	722	10
16-06-03	7:55 AM	20.0	20.2	720	10
23-06-03	7:45 AM	21.1	18.3	687	12
7-07-03	8:30 AM	24.4	14.3	738	11
14-07-03	7:50 AM	20.6	9.02	788	13
21-07-03	7:50 AM	22.2	13	479	15
28-07-03	7:45 AM	22.0	9.54	679	9
11-08-03	7:55 AM	23.2	20.6	644	18
18-08-03	7:40 AM	22.4	14.9	595	13
25-08-03	7:50 AM	22.9	8.5	674	11
8-09-03	7:50 AM	20.7	6.44	736	16
15-09-03	7:45 AM	22.1	9.76	740	11
22-09-03	7:45 AM	18.1	16.1	504	7
29-09-03	7:45 AM	15.7	14.1	599	10
6-10-03	7:45 AM	11.9	16.3	636	9
20-10-03	7:45 AM	10.6	15.8	663	14
27-10-03	7:45 AM	11.1	18.8	657	17
3-11-03	7:55 AM	11.1	114	677	49
10-11-03	7:40 AM	5.9	38.1	694	22
17-11-03	7:40 AM	6.9	80.3	686	38
24-11-03	7:45 AM	9.9	60.2	642	30
1-12-03	7:45 AM	4.8	114	633	21
8-12-03	7:45 AM	2.0	40	728	9
15-12-03	7:50 AM	2.5	46.9	761	12
12-01-04	8:05 AM	1.4	30	900	10
16-02-04	8:05 AM	0.0	14.8	968	6
23-02-04	8:10 AM	3.0	34	1,034	20
1-03-04	7:45 AM	2.4	52.2	920	26
8-03-04	7:00 AM	1.6	372	413	38
15-03-04	7:20 AM	2.6	71.4	696	17
22-03-04	7:30 AM	1.1	74.7	767	18
29-03-04	7:35 AM	7.8	147	563	25
5-04-04	7:50 AM	4.9	70.2	656	19
19-04-04	7:45 AM	14.4	30.4	718	7

26-04-04	7:50 AM	12.3	33.3	679	12
3-05-04	7:45 AM	8.8	55.9	671	16
17-05-04	7:45 AM	14.5	49.4	666	4
31-05-04	7:50 AM	15.6	39.8	689	17
7-06-04	7:45 AM	18.5	21.9	744	16
21-06-04	8:00 AM	16.7	22.1	721	12
28-06-04	7:55 AM	19.5	15.6	751	7
12-07-04	7:50 AM	21.7	14.2	714	16
19-07-04	7:50 AM	19.9	13.1	788	10
26-07-04	7:55 AM	21.2	11.3	711	17
9-08-04	8:10 AM	20.9	12.2	717	9
16-08-04	7:30 AM	20.6	8.9	754	6
23-08-04	7:50 AM	20.2	8.28	774	21
13-09-04	7:55 AM	20.4	10.2	669	9
20-09-04	7:50 AM	15.8	8.14	758	10
27-09-04	7:50 AM	16.7	7.9	757	9
4-10-04	7:40 AM	16.0	7.68	764	10
18-10-04	7:50 AM	9.7	9.39	726	4
25-10-04	7:55 AM	11.5	10	776	11
1-11-04	7:45 AM	11.2	20.3	669	14
8-11-04	7:40 AM	7.3	31.8	682	19
15-11-04	8:00 AM	5.0	12.9	789	8
22-11-04	7:50 AM	7.3	10.6	810	16
29-11-04	7:50 AM	5.0	23.2	757	19
6-12-04	7:50 AM	3.2	30.7	734	14
20-12-04	8:00 AM	0.0	15.8	936	13
10-01-05	8:30 AM	1.6	51.6	678	7
17-01-05	8:00 AM	0.0	72.4	608	15
24-01-05	8:00 AM	0.0	22	864	5
7-02-05	7:50 AM	1.4	18	980	12
14-02-05	8:00 AM	1.0	41.7	991	13
7-03-05	8:00 AM	2.0	27.4	922	11
14-03-05	7:50 AM	0.0	27	758	8
21-03-05	8:05 AM	0.0	25.6	842	12
11-04-05	8:00 AM	7.9	36.9	642	11
18-04-05	8:05 AM	8.9	20.5	694	10
25-04-05	7:55 AM	4.3	78	632	29
2-05-05	7:55 AM	5.8	54	632	14
9-05-05	7:55 AM	11.7	24.2	661	7
16-05-05	7:50 AM	8.2	26.4	718	10
30-05-05	7:50 AM	11.6	21.9	628	37
6-06-05	7:55 AM	17.2	14.9	697	5
13-06-05	7:50 AM	21.3	10.3	669	12
20-06-05	8:10 AM	13.5	14.7	625	11
27-06-05	8:00 AM	20.4	8.36	692	14

11-07-05	8:15 AM	18.8	6.7	556	18
18-07-05	8:00 AM	20.2	56.2	418	22
25-07-05	7:50 AM	19.9	9.43	493	17
8-08-05	7:50 AM	17.4	7.76	474	14
15-08-05	8:10 AM	17.3	8.84	660	16
22-08-05	7:50 AM	17.6	18.3	551	25
29-08-05	7:55 AM	16.8	7.83	677	12
12-09-05	8:05 AM	16.2	6.02	702	14
19-09-05	8:00 AM	15.0	8.16	619	10
26-09-05	7:55 AM	14.2	15	618	122
17-10-05	7:45 AM	15.1	13.8	629	12
24-10-05	7:55 AM	12.0	8.81	744	19
31-10-05	7:55 AM	11.1	14.8	737	12
7-11-05	7:50 AM	11.3	8.67	730	36
14-11-05	7:45 AM	9.5	12.4	696	30
21-11-05	7:50 AM	5.9	35.7	610	31
28-11-05	7:50 AM	5.1	27.8	940	31
5-12-05	8:10 AM	3.0	42.7	672	16
12-12-05	8:10 AM	1.2	16	1,020	11
19-12-05	8:00 AM	1.2	12.4	888	6
9-01-06	7:50 AM	3.0	55.3	688	16
16-01-06	8:25 AM	1.5	73.8	657	19
23-01-06	7:55 AM	2.5	107.6	602	21
30-01-06	7:50 AM	4.0	220.3	586	83
6-02-06	8:10 AM	1.5	146.3	685	27
13-02-06	7:50 AM	1.5	45.2	823	1
20-02-06	7:50 AM	0.6	154.2	569	11
27-02-06	8:00 AM	0.4	33.3	626	7
6-03-06	8:40 AM	2.2	27.3	702	17
13-03-06	8:30 AM	8.7	450.5	615	64
20-03-06	8:45 AM	8.7	59.1	649	12
27-03-06	8:00 AM	8.6	32.3	684	4
10-04-06	8:00 AM	6.6	48.2	600	15
24-04-06	7:50 AM	11.2	48.77	524	37
8-05-06	8:00 AM	13.2	14.5	779	6
15-05-06	7:50 AM	14.2	21.1	714	8
29-05-06	8:00 AM	18.9	16.8	520	14
5-06-06	8:05 AM	16.1	53.3	441	35
19-06-06	8:00 AM	21.0	13.1	751	10
26-06-06	8:15 AM	21.0	15.8	714	11
10-07-06	8:05 AM	20.4	10.1	752	10
17-07-06	8:00 AM	22.9	13.9	589	12
24-07-06	8:00 AM	19.0	12.6	684	7
31-07-06	8:00 AM	21.7	47.7	544	21
14-08-06	8:00 AM	20.3	11.9	698	26

21-08-06	7:55 AM	19.6	9.8	620	14
28-08-06	7:45 AM	19.5	10.0	681	9
11-09-06	7:50 AM	16.7	9.5	712	9
18-09-06	7:55 AM	17.5	33.7	712	16
25-09-06	8:00 AM	13.2	25.0	565	53
16-10-06	7:45 AM	7.7	59.0	626	22
23-10-06	7:30 AM	8.8	106.6	529	20
30-10-06	7:55 AM	5.7	211.6	524	37
6-11-06	7:55 AM	5.7	41.5	654	13
13-11-06	7:50 AM	6.3	51.1	648	12
20-11-06	8:00 AM	5.0	81.1	575	22
27-11-06	7:55 AM	6.1	34.0	672	10
4-12-06	7:55 AM	2.1	143.0	472	65
11-12-06	7:55 AM	2.0	53.6	908	20
18-12-06	7:55 AM	5.2	64.5	674	22
8-01-07	8:00 AM	4.3	141.8	598	40
15-01-07	8:00 AM	2.9	105.8	699	31
22-01-07	8:00 AM	0.0	38.5	948	13
29-01-07	8:00 AM	-0.9	24.9	917	10
5-02-07	8:10 AM	-1.0	17.0	994	7
12-02-07	8:00 AM	0.2	17.1	952	9
19-02-07	8:00 AM	-0.8	14.8	1,123	22
26-02-07	8:05 AM	0.7	15.7	1,129	7
5-03-07	8:00 AM	0.0	19.5	1,307	7
12-03-07	7:55 AM	0.3	41.7	933	13
19-03-07	8:05 AM	1.0	105.1	537	15
26-03-07	8:05 AM	2.0	237.0	444	51
16-04-07	8:00 AM	3.7	47.2	650	12
23-04-07	7:05 AM	12.3	29.5	612	4
30-04-07	8:00 AM	12.6	47.2	472	7
7-05-07	8:00 AM	14.2	20.7	491	6
14-05-07	8:05 AM	14.9	17.2	762	17
28-05-07	7:55 AM	16.8	15.9	667	9
4-06-07	8:00 AM	20.4	20.3	759	6
11-06-07	7:55 AM	19.7	15.7	670	11
25-06-07	8:00 AM	21.3	9.4	699	8
9-07-07	8:00 AM	23.2	7.3	750	14
16-07-07	7:50 AM	17.1	7.0	495	16
23-07-07	8:05 AM	17.0	6.3	715	21
30-07-07	8:00 AM	22.0	6.4	452	14
13-08-07	8:05 AM	22.3	7.6	401	15
20-08-07	7:55 AM	18.1	10.2	762	19
27-08-07	1:12 AM	19.7	8.8	696	20
10-09-07	8:00 AM	19.4	7.7	757	11
17-09-07	8:10 AM	14.7	7.5	855	14

24-09-07	7:55 AM	17.1	7.2	818	10
15-10-07	8:00 AM	13.6	7.6	852	14
22-10-07	8:00 AM	15.1	7.5	789	6
29-10-07	8:00 AM	11.1	7.9	787	13
5-11-07	8:00 AM	9.3	7.7	839	8
12-11-07	8:00 AM	9.9	6.4	895	10
19-11-07	8:00 AM	6.5	6.5	775	6
26-11-07	7:50 AM	5.8	14.5	752	8
3-12-07	8:00 AM	3.9	32.0	960	48
10-12-07	8:00 AM	2.7	18.8	896	5
7-01-08	8:00 AM	4.9	116.4	814	33
14-01-08	8:00 AM	4.1	89.0	700	27
21-01-08	7:55 AM	1.4	29.6	887	12
28-01-08	8:00 AM	2.3	21.2	1,228	11
25-02-08	8:00 AM	1.0	47.3	781	11
3-03-08	8:20 AM	2.7	45.8	810	4
17-03-08	7:55 AM	2.0	58.4	810	7
7-04-08	7:55 AM	6.6	157.2	732	15
14-04-08	7:55 AM	7.3	87.1	891	1
21-04-08	7:50 AM	13.3	29.2	805	9
28-04-08	7:55 AM	13.6	21.1	782	5
5-05-08	7:55 AM	11.9	21.6	645	6
26-05-08	7:50 AM	15.3	17.0	684	4
2-06-08	8:05 AM	14.8	21.4	612	34
9-06-08	7:55 AM	21.2	13.5	684	6
16-06-08	8:35 AM	16.8	18.7	657	7
23-06-08	8:05 AM	18.2	31.2	671	29
30-06-08	8:25 AM	15.9	54.5	507	23
7-07-08	8:35 AM	17.4	14.8	684	7
14-07-08	8:15 AM	17.1	38.0	500	18
28-07-08	9:00 AM	17.5	15.4	635	4
11-08-08	8:05 AM	15.2	27.2	508	40
18-08-08	7:50 AM	18.0	11.3	624	17
25-08-08	7:25 AM	19.4	8.2	655	13
8-09-08	7:55 AM	15.7	10.1	526	17
15-09-08	8:10 AM	17.1	95.4	563	62
22-09-08	8:00 AM	15.5	15.6	566	20
26-09-08	8:05 AM	15.7	15.0	626	6
6-10-08	7:50 AM	11.0	29.0	559	9
20-10-08	7:55 AM	8.0	17.3	651	1
3-11-08	7:55 AM	7.7	25.9	659	10
10-11-08	8:05 AM	5.6	53.2	666	6
17-11-08	8:05 AM	4.9	219.7	562	16
1-12-08	8:05 AM	2.3	72.2	587	14
15-12-08	7:55 AM	5.5	144.6	732	42

9-02-09	7:50 AM	1.9	55.0	885	18
9-03-09	8:05 AM	2.7	331.6	320	124
16-03-09	7:40 AM	4.3	70.2	467	17
23-03-09	8:00 AM	5.3	47.0	616	12
30-03-09	7:55 AM	7.3	47.0	647	10
6-04-09	8:00 AM	7.5	115.9	602	41
20-04-09	8:00 AM	10.6	39.4	690	5
11-05-09	7:55 AM	12.6	42.4	653	11
25-05-09	7:55 AM	16.8	20.9	733	14
1-06-09	7:35 AM	15.1	64.3	636	15
15-06-09	7:55 AM	18.7	24.1	704	7
29-06-09	8:00 AM	20.9	18.8	727	2
20-07-09	7:50 AM	19.1	12.7	772	12
27-07-09	7:50 AM	20.0	21.6	727	21
17-08-09	8:05 AM	24.2	13.4	723	16
24-08-09	7:55 AM	19.3	11.9	756	19
31-08-09	8:00 AM	16.4	13.2	699	14
14-09-09	7:55 AM	19.3	10.7	721	13
28-09-09	7:55 AM	17.4	14.6	772	29
5-10-09	8:00 AM	12.4	15.6	706	9
19-10-09	8:10 AM	8.7	15.7	746	12
26-10-09	8:00 AM	10.7	14.4	799	13
2-11-09	8:00 AM	9.8	15.7	802	6
9-11-09	8:00 AM	10.3	16.6	788	9
16-11-09	7:50 AM	8.5	14.8	798	7
7-12-09	7:55 AM	3.2	24.3	792	1
14-12-09	8:00 AM	3.4	38.5	798	23
18-01-10	7:55 AM	2.7	13.7	948	5
25-01-10	7:55 AM	4.4	50.1	825	23
8-02-10	7:50 AM	0.6	18.4	855	6
22-02-10	8:00 AM	2.9	14.3	790	10
1-03-10	8:00 AM	3.4	14.3	1,230	8
8-03-10	8:00 AM	4.3	21.3	882	14
15-03-10	8:05 AM	4.0	329.8	477	47
29-03-10	8:05 AM	7.8	25.7	759	24
12-04-10	8:30 AM	9.8	66.3	691	18
19-04-10	7:50 AM	10.0	22.7	750	6
26-04-10	8:00 AM	11.7	19.0	735	11
3-05-10	8:00 AM	17.4	14.1	804	3
10-05-10	8:05 AM	10.0	61.5	631	20
17-05-10	8:00 AM	13.3	36.6	714	13
31-05-10	8:30 AM	21.6	25.5	748	2
7-06-10	8:00 AM	18.0	155.3	568	52
14-06-10	8:00 AM	19.3	26.4	738	11
21-06-10	8:00 AM	22.1	12.8	741	2

28-06-10	7:50 AM	22.9	11.8	709	7
5-07-10	8:30 AM	24.5	10.2	711	1
12-07-10	8:30 AM	23.6	23.6	699	12
19-07-10	7:00 AM	23.3	17.6	653	22
26-07-10	7:50 AM	22.6	18.6	669	13
9-08-10	8:00 AM	22.5	10.1	704	7
23-08-10	7:55 AM	21.5	11.7	671	14
30-08-10	8:25 AM	23.1	8.1	720	5
13-09-10	7:55 AM	17.0	8.7	750	8
20-09-10	8:05 AM	15.5	10.2	711	5
27-09-10	8:00 AM	15.1	8.6	775	6
4-10-10	7:55 AM	11.7	18.5	689	13
18-10-10	7:55 AM	11.8	22.0	732	10
25-10-10	7:55 AM	13.6	23.6	734	18
1-11-10	8:05 AM	7.5	21.5	771	5
8-11-10	9:45 AM	6.7	17.8	781	5
15-11-10	7:55 AM	7.1	14.9	785	10
22-11-10	7:55 AM	8.8	19.0	795	8
29-11-10	7:50 AM	3.5	41.6	791	12
14-03-11	7:45 AM	5.2	197.7	610	18
21-03-11	7:55 AM	4.9	218.3	630	45
28-03-11	8:35 AM	3.0	59.5	700	10
4-04-11	8:15 AM	7.0	70.5	753	18
11-04-11	8:10 AM	11.5	62.1	648	15
18-04-11	7:55 AM	6.8	42.6	698	4
2-05-11	8:05 AM	10.8	83.6	618	19
9-05-11	7:50 AM	13.8	47.2	671	7
16-05-11	7:50 AM	12.5	161.9	516	42
30-05-11	7:55 AM	17.6	67.1	529	138
6-06-11	8:10 AM	18.9	36.0	602	18
13-06-11	7:50 AM	16.2	27.8	621	20
20-06-11	7:55 AM	19.5	15.4	639	8
27-06-11	8:00 AM	19.6	42.8	563	20
4-07-11	8:35 AM	22.8	16.6	620	10
11-07-11	7:20 AM	23.9	10.0	670	16
18-07-11	7:55 AM	24.9	9.1	650	8
25-07-11	7:55 AM	24.3	18.3	460	26
8-08-11	7:50 AM	24.1	9.7	730	18
15-08-11	7:50 AM	20.8	9.8	700	18
22-08-11	8:40 AM	20.2	11.4	644	12
29-08-11	8:45 AM	20.4	14.0	669	8
19-09-11	7:55 AM	16.8	8.4	680	10
26-09-11	7:50 AM	18.4	15.7	610	21
3-10-11	7:50 AM	12.8	29.6	510	30
17-10-11	8:00 AM	11.4	115.0	464	20

24-10-11	8:00 AM	11.1	92.4	490	27
31-10-11	8:05 AM	9.2	46.5	520	12
7-11-11	7:50 AM	9.6	21.5	550	6
14-11-11	7:50 AM	10.3	19.5	560	7
21-11-11	7:50 AM	7.7	19.2	523	16
28-11-11	7:55 AM	8.6	35.4	500	46
5-12-11	8:00 AM	7.4	128.1	520	42
12-12-11	8:00 AM	2.9	77.5	540	21
9-01-12	8:10 AM	2.7	72.6	540	21
16-01-12	8:20 AM	2.5	51.1	540	18
23-01-12	7:45 AM	3.2	46.6	650	22
6-02-12	8:00 AM	3.0	66.3	540	8
13-02-12	7:55 AM	1.4	28.7	391	9
27-02-12	7:55 AM	3.0	51.5	600	4
5-03-12	8:05 AM	1.7	138.6	330	24
12-03-12	8:00 AM	5.9	47.4	430	8
19-03-12	7:55 AM	12.7	41.5	503	10
26-03-12	7:55 AM	10.8	24.3	511	1
2-04-12	8:00 AM	8.1	18.0	500	8
16-04-12	7:55 AM	14.1	13.9	600	6
23-04-12	8:05 AM	9.8	12.7	540	<3
4-06-12	7:55 AM	16.6	37.7	511	38
11-06-12	8:00 AM	22.4	13.3	700	8
18-06-12	7:00 AM	21.8	9.7	702	10
9-07-12	9:00 AM	24.0	6.1	890	14
13-08-12	8:15 AM	21.8	12.2	680	26
27-08-12	8:50 AM	23.7	6.8	730	1
10-09-12	8:45 AM	18.3	11.0	590	12
17-09-12	8:15 AM	18.3	7.6	650	14
1-10-12	8:05 AM	14.0	9.8	800	7
15-10-12	8:05 AM	14.0	9.2	816	12
29-10-12	8:15 AM	10.1	23.7	629	28
5-11-12	8:05 AM	6.6	36.5	680	17
12-11-12	8:00 AM	10.3	11.6	880	10
19-11-12	8:00 AM	6.7	10.7	790	10
26-11-12	8:10 AM	4.8	9.9	803	6
3-12-12	8:05 AM	7.6	13.2	760	11
10-12-12	7:55 AM	5.5	25.5	737	8
14-01-13	7:45 AM	4.8	225.7	460	89
21-01-13	7:40 AM	1.1	46.3	550	11
25-02-13	7:45 AM	2.1	49.1	644	12
4-03-13	7:45 AM	2.4	35.8	700	4
11-03-13	7:35 AM	4.0	64.1	665	108
18-03-13	7:40 AM	2.2	95.8	690	8
8-04-13	8:00 AM	6.3	39.8	605	10

15-04-13	8:45 AM	5.8	264.8	448	26
29-04-13	7:50 AM	11.5	51.3	690	12
6-05-13	8:00 AM	15.7	28.0	810	1
13-05-13	7:45 AM	10.7	22.3	750	10
27-05-13	7:55 AM	14.8	13.0	887	8
3-06-13	8:00 AM	17.8	76.4	753	35
10-06-13	7:40 AM	18.4	20.3	758	14
24-06-13	7:50 AM	23.8	21.6	750	8
8-07-13	8:35 AM	22.1	37.7	583	12
15-07-13	7:55 AM	24.6	24.5	665	12
29-07-13	7:55 AM	18.4	11.9	608	11
12-08-13	8:45 AM	20.0	9.6	637	8
19-08-13	8:00 AM	20.1	7.7	630	<3
26-08-13	7:45 AM	22.2	6.7	670	6
9-09-13	8:00 AM	18.7	7.0	640	8
23-09-13	8:00 AM	15.2	211.7	620	31
30-09-13	7:50 AM	17.3	21.8	700	8
7-10-13	8:20 AM	16.5	61.7	390	94
21-10-13	8:05 AM	12.0	58.6	520	16
28-10-13	8:00 AM	8.0	210.4	410	25
4-11-13	8:00 AM	7.7	216.3	480	24
11-11-13	8:15 AM	6.5	74.4	570	14
18-11-13	7:40 AM	8.0	47.2	520	10
17-03-14	7:40 AM	0.2	58.00	430	5
31-03-14	7:45 AM	2.2	172.70	452	44
7-04-14	8:00 AM	3.5	261.63	557	12
14-04-14	7:55 AM	9.9	118.86	550	14
28-04-14	7:45 AM	9.8	27.36	486	1
5-05-14	7:55 AM	9.0	79.96	440	10
12-05-14	7:55 AM	15.6	30.52	550	1
26-05-14	8:05 AM	16.1	56.83	535	8
2-06-14	7:55 AM	19.6	19.97	640	2
16-06-14	7:05 AM	20.8	18.22	570	12
23-06-14	8:00 AM	21.8	9.92	680	8
7-07-14	8:40 AM	22.2	9.19	570	33
14-07-14	7:55 AM	21.3	25.16	580	13
28-07-14	7:55 AM	20.9	10.73	560	22
11-08-14	8:00 AM	21.5	11.70	650	9
18-08-14	7:55 AM	19.5	14.94	600	13
25-08-14	8:35 AM	20.6	13.92	640	12
8-09-14	8:00 AM	18.6	107.36	450	37
15-09-14	7:55 AM	14.6	72.98	600	28
29-09-14	8:05 AM	16.6	12.81	640	5
6-10-14	8:00 AM	13.3	81.68	550	18
20-10-14	8:10 AM	11.3	52.42	560	10

27-10-14	8:05 AM	9.3	21.21	570	4
10-11-14	7:50 AM	7.0	32.30	530	1
24-11-14	9:05 AM	5.5	54.50	470	76
8-12-14	8:10 AM	2.0	27.57	490	6
16-03-15	7:30 AM	1.5	50.6	441	39
23-03-15	7:45 AM	0.5	17.3	508	26
13-04-15	8:05 AM	8.0	23.3	540	39
27-04-15	7:50 AM	9.4	17.5	590	0.5
4-05-15	8:00 AM	15.3	79.9	700	6
11-05-15	8:50 AM	20.4	12.9	840	6
17-08-15	8:45 AM	23.9	6.7	890	11
14-09-15	7:50 AM	16.1	13.0	680	9
21-09-15	8:30 AM	16.2	6.6	690	7
28-09-15	7:50 AM	19.4	4.8	810	10
5-10-15	7:50 AM	13.6	8.5	650	12
26-10-15	7:25 AM	11.0	10.4	620	10
2-11-15	7:50 AM	9.8	20.2	568	12
9-11-15	8:10 AM	8.1	17.3	455	0.5
16-11-15	7:55 AM	7.7	23.3	440	8
23-11-15	7:50 AM	3.9	17.5	504	6
30-11-15	7:30 AM	4.2	79.9	370	21
14-12-15	7:40 AM	8.0	14.7	500	4
4-01-16	7:50 AM	0.2	36.7	480	13
11-01-16	8:00 AM	0.0	163.2	430	62
25-01-16	8:05 AM	0.9	31.3	500	8
1-02-16	7:50 AM	3.0	134.3	550	36
8-02-16	7:55 AM	4.1	59.9	470	15
22-02-16	8:05 AM	1.9	160.0	400	33
7-03-16	7:50 AM	3.4	34.5	440	8
14-03-16	7:55 AM	7.3	45.4	430	6
21-03-16	7:55 AM	5.0	35.2	420	10
4-04-16	8:00 AM	3.2	124.8	420	28
11-04-16	8:05 AM	4.0	141.1	460	17
18-04-16	7:50 AM	11.2	46.2	460	7
25-04-16	7:55 AM	11.1	27.1	480	1
2-05-16	7:35 AM	10.6	29.3	480	8
16-05-16	7:35 AM	9.3	21.5	450	1
30-05-16	7:50 AM	20.9	12.2	710	1
6-06-16	7:45 AM	18.9	16.8	560	6
20-06-16	7:55 AM	22.5	8.1	750	6
27-06-16	8:05 AM	23.2	7.0	820	10
4-07-16	8:00 AM	20.3	6.1	790	8
11-07-16	7:55 AM	22.2	6.4	760	12
18-07-16	7:55 AM	22.6	7.3	670	14
25-07-16	7:55 AM	23.8	5.6	800	17

8-08-16	8:05 AM	21.6	4.9	790	12
15-08-16	8:40 AM	22.7	12.3	620	22
22-08-16	8:50 AM	21.4	13.2	620	24
29-08-16	8:15 AM	22.0	15.1	630	24
12-09-16	8:00 AM	18.5	7.4	660	16
19-09-16	8:10 AM	19.7	8.1	660	14
26-09-16	7:55 AM	17.2	6.7	740	10
17-10-16	8:30 AM	16.2	7.6	690	12
24-10-16	7:50 AM	12.0	14.5	520	12
31-10-16	8:00 AM	8.9	7.5	620	15
21-11-16	8:05 AM	4.1	9.0	740	10
28-11-16	7:35 AM	4.9	12.2	530	8
5-12-16	8:05 AM	4.4	22.8	500	14

Appendix B Grubb's test MATLAB Code

```
function [b,idx,outliers] =
deleteoutliers(a,alpha,rep);
% [B, IDX, OUTLIERS] = DELETEOUTLIERS(A, ALPHA, REP)
%
% For input vector A, returns a vector B with outliers
(at the significance
% level alpha) removed. Also, optional output argument
idx returns the
% indices in A of outlier values. Optional output
argument outliers returns
% the outlying values in A.
%
% ALPHA is the significance level for determination of
outliers. If not
% provided, alpha defaults to 0.05.
%
% REP is an optional argument that forces the
replacement of removed
% elements with NaNs to presereve the length of a.
(Thanks for the
% suggestion, Urs.)
%
% This is an iterative implementation of the Grubbs
Test that tests one
% value at a time. In any given iteration, the tested
value is either the
```

```
% highest value, or the lowest, and is the value that
is furthest
% from the sample mean. Infinite elements are discarded
if rep is 0, or
% replaced with NaNs if rep is 1 (thanks again, Urs).
%
% Appropriate application of the test requires that
data can be reasonably
% approximated by a normal distribution. For reference,
see:
% 1) "Procedures for Detecting Outlying Observations in
Samples," by F.E.
%    Grubbs; Technometrics, 11-1:1--21; Feb., 1969, and
% 2) _Outliers in Statistical Data_, by V. Barnett and
%    T. Lewis; Wiley Series in Probability and
Mathematical Statistics;
%    John Wiley & Sons; Chichester, 1994.
% A good online discussion of the test is also given in
NIST's Engineering
% Statistics Handbook:
%
http://www.itl.nist.gov/div898/handbook/eda/section3/ed
a35h.htm
%
% ex:
% [B,idx,outliers] = deleteoutliers([1.1 1.3 0.9 1.2 -
6.4 1.2 0.94 4.2 1.3 1.0 6.8 1.3 1.2], 0.05)
%    returns:
%    B = 1.1000    1.3000    0.9000    1.2000    1.2000
0.9400    1.3000    1.0000    1.3000    1.2000
%    idx =  5     8    11
%    outliers = -6.4000    4.2000    6.8000
%
% ex:
% B = deleteoutliers([1.1 1.3 0.9 1.2 -6.4 1.2 0.94 4.2
1.3 1.0 6.8 1.3 1.2
% Inf 1.2 -Inf 1.1], 0.05, 1)
% returns:
% B = 1.1000  1.3000  0.9000  1.2000  NaN  1.2000
0.9400  NaN  1.3000  1.0000  NaN  1.3000  1.2000  NaN
1.2000  NaN  1.1000
% Written by Brett Shoelson, Ph.D.
% shoelson@helix.nih.gov
% 9/10/03
```

```
% Modified 9/23/03 to address suggestions by Urs
Schwartz.
% Modified 10/08/03 to avoid errors caused by duplicate
"maxvals."
%    (Thanks to Valeri Makarov for modification
suggestion.)

if nargin == 1
    alpha = 0.05;
    rep = 0;
elseif nargin == 2
    rep = 0;
elseif nargin == 3
    if ~ismember(rep,[0 1])
        error('Please enter a 1 or a 0 for optional
argument rep.')
    end
elseif nargin > 3
    error('Requires 1,2, or 3 input arguments.');
end

if isempty(alpha)
    alpha = 0.05;
end

b = a;
b(isinf(a)) = NaN;

%Delete outliers:
outlier = 1;
while outlier
    tmp = b(~isnan(b));
    meanval = mean(tmp);
    maxval = tmp(find(abs(tmp-mean(tmp))==max(abs(tmp-
mean(tmp)))));
    maxval = maxval(1);
    sdval = std(tmp);
    tn = abs((maxval-meanval)/sdval);
    critval = zcritical(alpha,length(tmp));
    outlier = tn > critval;
    if outlier
        tmp = find(a == maxval);
        b(tmp) = NaN;
    end
```

```
end
if nargout >= 2
    idx = find(isnan(b));
end
if nargout > 2
    outliers = a(idx);
end
if ~rep
    b=b(~isnan(b));
end
return

function zcrit = zcritical(alpha,n)
%ZCRIT = ZCRITICAL(ALPHA,N)
% Computes the critical z value for rejecting outliers
(GRUBBS TEST)
tcrit = tinv(alpha/(2*n),n-2);
zcrit = (n-1)/sqrt(n)*(sqrt(tcrit^2/(n-2+tcrit^2)));
```

Appendix C Processed data after removing the duplications and outliers

Date	River Temp (^{C})	Flow (cms)	Conductivity (uS/cm)	Suspended Solids
28-Sep-2015	19.4	4.8	810	10
8-Aug-2016	21.6	4.9	790	12
23-Sep-98	8.2	5.03	780	17
2-Sep-98	8.1	5.1	720	17
9-Sep-98	8.1	5.17	710	11
4-Aug-99	8.1	5.46	480	9
1-Sep-99	7.6	5.49	600	8
28-Jul-99	7.6	5.51	510	5
25-Jul-2016	23.8	5.6	800	17
8-Aug-01	8	5.66	642	20
18-Aug-99	7.7	5.68	540	10
13-Aug-01	8	5.73	660	11
21-Oct-98	8.1	5.79	730	5
22-Sep-99	7.6	5.82	630	13
23-Jun-99	7.9	5.87	620	13
5-Sep-01	8	5.94	641	7
12-Sep-05	8.4	6.02	702	14
14-Jul-99	8	6.05	540	11
16-Sep-98	8.2	6.05	740	11

9-Jul-2012	24	6.1	890	14
5-Aug-98	8.2	6.18	700	11
7-Jul-99	7.8	6.2	570	11
11-Aug-99	7.8	6.21	570	8
23-Jul-2007	17	6.3	715	21
15-Sep-99	7.7	6.32	580	11
11-Jul-2016	22.2	6.4	760	12
8-Sep-03	8.4	6.44	736	16
21-Jul-99	7.7	6.45	540	13
19-Nov-2007	6.5	6.5	775	6
28-Oct-98	8.1	6.52	690	5
30-Sep-98	8.1	6.6	740	5
19-May-99	8.1	6.62	650	5
26-Aug-2013	22.2	6.7	670	6
9-Sep-02	8.1	6.86	695	3
9-Sep-2013	18.7	7	640	8
6-Sep-95	7.9	7.05	620	10
19-Aug-98	7.8	7.08	400	13
6-Oct-99	8.1	7.1	610	14
19-Oct-99	7.7	7.17	640	11
24-Sep-2007	17.1	7.2	818	10
18-Jul-2016	22.6	7.3	670	14
9-Jul-2007	23.2	7.3	750	14
30-Aug-95	8.2	7.37	648	12
12-Sep-2016	18.5	7.4	660	16
17-Sep-2007	14.7	7.5	855	14
29-Sep-99	7.7	7.56	600	8
15-Jul-98	8.1	7.58	600	18
17-Sep-2012	18.3	7.6	650	14
15-Oct-2007	13.6	7.6	852	14
23-Jul-01	8	7.67	631	14
4-Oct-04	8	7.68	764	10
10-Sep-2007	19.4	7.7	757	11
7-Oct-98	8.1	7.72	720	12
8-Aug-05	8.1	7.76	474	14
29-Aug-05	8	7.83	677	12
2-Dec-98	8.2	7.87	620	18
29-Oct-2007	11.1	7.9	787	13
4-Oct-95	7.8	8	693	10
19-Sep-2016	19.7	8.1	660	14
20-Sep-04	8	8.14	758	10

19-Sep-05	8	8.16	619	10
25-Aug-2008	19.4	8.2	655	13
1-Oct-01	7.8	8.33	573	8
27-Jun-05	8.3	8.36	692	14
28-Jun-99	7.9	8.37	610	10
19-Sep-2011	16.8	8.4	680	10
14-Oct-98	8.1	8.49	680	11
25-Aug-03	8	8.5	674	11
27-Oct-99	7.9	8.52	610	19
8-Jul-02	8.3	8.53	605	12
21-Sep-94	7.8	8.57	697	13
27-Sep-2010	15.1	8.6	775	6
8-Sep-99	8	8.62	630	8
7-Oct-02	7.8	8.66	654	5
13-Sep-2010	17	8.7	750	8
27-Aug-2007	19.7	8.8	696	20
24-Oct-05	8	8.81	744	19
15-Aug-05	7.8	8.84	660	16
13-Oct-99	7.3	8.87	610	11
20-Sep-95	7.9	8.89	664	8
16-Aug-04	8.1	8.9	754	6
14-Aug-96	7.4	8.95	662	6
21-Nov-2016	4.1	9	740	10
14-Jul-03	8.1	9.02	788	13
18-Jul-2011	24.9	9.1	650	8
7-Aug-96	7.9	9.11	665	7
12-Oct-94	8.1	9.17	699	14
15-Oct-2012	14	9.2	816	12
20-Jan-03	8.1	9.23	1037	6
4-Jul-01	6.9	9.34	590	13
18-Oct-04	7.4	9.39	726	4
25-Jun-2007	21.3	9.4	699	8
25-Jul-05	7.9	9.43	493	17
3-Jun-98	8.3	9.5	670	16
6-Sep-94	8.2	9.5	691	16
14-May-98	8.2	9.54	590	15
24-Nov-99	8.6	9.56	660	19
20-Jun-95	8.1	9.57	678	13
12-Aug-2013	20	9.6	637	8
9-Jun-99	8.1	9.65	580	14
18-Jun-2012	21.8	9.7	702	10

12-May-99	8.2	9.72	570	11
9-Aug-93	7.8	9.73	590	11
15-Sep-03	8.1	9.76	740	11
27-May-98	8.2	9.78	590	16
5-Oct-94	7.9	9.78	677	16
26-Jul-95	7.9	9.79	664	8
1-Oct-2012	14	9.8	800	7
9-Dec-02	8	9.88	912	8
26-Nov-2012	4.8	9.9	803	6
23-Jun-2014	21.8	9.92	680	8
28-Aug-06	7.9	10	681	9
10-Jul-06	8.2	10.1	752	10
13-Sep-04	8.1	10.2	669	9
13-Jun-05	8.1	10.3	669	12
26-Oct-2015	11	10.4	620	10
11-Jan-95	7.7	10.5	864	7
17-Feb-03	7.9	10.7	991	7
8-Nov-00	8.2	10.8	757	12
24-Aug-94	7.8	10.9	693	12
10-Sep-2012	18.3	11	590	12
30-Aug-00	8.3	11.3	696	11
22-Aug-2011	20.2	11.4	644	12
20-May-98	8.2	11.5	600	13
18-Oct-95	7.9	11.6	670	10
12-Nov-2012	10.3	11.6	880	10
11-Aug-2014	21.5	11.7	650	9
28-Jun-2010	22.9	11.8	709	7
29-Jul-2013	18.4	11.9	608	11
9-Aug-04	8.4	12.2	717	9
5-Jan-94	7.8	12.4	825	8
10-Oct-01	8.1	12.5	545	8
24-Jul-06	7.9	12.6	684	7
20-Jul-2009	19.1	12.7	772	12
10-Aug-94	7.8	12.8	683	18
15-Nov-04	7.9	12.9	789	8
14-Sep-2015	16.1	13	680	9
19-Jun-06	8.1	13.1	751	10
19-Jul-04	8.2	13.1	788	10
3-Dec-2012	7.6	13.2	760	11
11-Jun-2012	22.4	13.3	700	8
9-Jun-2008	21.2	13.5	684	6

18-Jan-2010	2.7	13.7	948	5
17-Oct-05	7.7	13.8	629	12
17-Jul-06	8.2	13.9	589	12
25-Aug-2014	20.6	13.92	640	12
29-Aug-2011	20.4	14	669	8
29-Sep-03	7.9	14.1	599	10
26-Oct-2009	10.7	14.4	799	13
24-Oct-2016	12	14.5	520	12
20-Jun-05	8.1	14.7	625	11
18-Aug-03	8	14.9	595	13
18-Aug-2014	19.5	14.94	600	13
26-Sep-2008	15.7	15	626	6
25-Nov-02	8.2	15.1	691	5
14-May-01	7.6	15.2	685	9
20-Jun-2011	19.5	15.4	639	8
5-Oct-2009	12.4	15.6	706	9
11-Jun-2007	19.7	15.7	670	11
26-Jun-06	8.2	15.8	714	11
28-May-2007	16.8	15.9	667	9
12-Dec-05	7.9	16	1020	11
22-Sep-03	7.5	16.1	504	7
6-Oct-03	8	16.3	636	9
3-Jul-96	8.3	16.4	656	5
9-Nov-2009	10.3	16.6	788	9
6-Jun-2016	18.9	16.8	560	6
7-May-01	7.7	16.9	687	6
5-Feb-07	7.9	17	994	7
12-Feb-07	8	17.1	952	9
5-Jun-00	7.4	17.3	698	11
19-Jan-00	8.4	17.5	809	9
27-Oct-93	8.1	17.6	686	10
8-Feb-95	8.1	17.7	888	5
8-Nov-2010	6.7	17.8	781	5
2-Apr-2012	8.1	18	500	8
16-Jun-2014	20.8	18.22	570	12
23-Feb-95	8.3	18.3	832	11
8-Feb-2010	0.6	18.4	855	6
10-Jan-01	8.1	18.5	834	8
17-Jun-02	7.7	18.6	628	9
16-Jun-2008	16.8	18.7	657	7
10-Dec-2007	2.7	18.8	896	5

26-Apr-2010	11.7	19	735	11
24-May-95	8	19.1	675	11
19-Apr-95	8.4	19.2	722	10
6-Jun-01	7.7	19.4	601	17
14-Nov-2011	10.3	19.5	560	7
14-Apr-99	8.2	19.7	470	2
2-Jun-2014	19.6	19.97	640	2
18-Nov-02	8.1	20.1	788	4
16-Jun-03	8.1	20.2	720	10
4-Jun-2007	20.4	20.3	759	6
26-Jul-00	8.2	20.5	643	10
18-Apr-05	8.1	20.5	694	10
11-Aug-03	8.1	20.6	644	18
16-Oct-96	8.2	20.7	668	8
13-Jun-01	7.9	21	584	11
15-May-06	8.2	21.1	714	8
27-Oct-2014	9.3	21.21	570	4
13-Dec-95	7.8	21.3	788	8
2-Jun-03	7.5	21.5	722	10
10-Jun-02	8	21.6	567	8
24-Jun-2013	23.8	21.6	750	8
24-May-94	7.9	21.7	665	11
30-Sep-2013	17.3	21.8	700	8
28-Apr-99	8.2	21.9	510	6
18-Oct-2010	11.8	22	732	10
29-Mar-95	8.2	22.1	705	11
16-Aug-00	8.3	22.2	697	19
13-May-2013	10.7	22.3	750	10
7-Apr-99	8.2	22.4	540	6
24-Jan-01	8.1	22.5	842	13
12-Apr-95	8	22.6	718	17
19-Apr-2010	10	22.7	750	6
5-Dec-2016	4.4	22.8	500	14
11-Oct-00	7.7	23.1	741	14
29-Nov-04	7.9	23.2	757	19
16-Nov-2015	7.7	23.3	440	8
12-Jul-2010	23.6	23.6	699	12
29-Oct-2012	10.1	23.7	629	28
22-Mar-00	8.7	23.8	779	6
17-Mar-99	8.1	23.9	510	8
15-Jun-2009	18.7	24.1	704	7

9-May-05	8.1	24.2	661	7
14-Jan-02	8	24.3	693	7
15-Jul-2013	24.6	24.5	665	12
29-Jan-07	8.1	24.9	917	10
31-May-00	8.1	25	688	7
22-Nov-00	8.3	25	815	7
14-Jul-2014	21.3	25.16	580	13
28-May-97	8.1	25.4	486	9
10-Dec-2012	5.5	25.5	737	8
21-Mar-05	8	25.6	842	12
29-Mar-2010	7.8	25.7	759	24
6-May-02	8.3	25.8	653	11
15-Mar-00	8.6	25.8	805	11
3-Nov-2008	7.7	25.9	659	10
13-Mar-96	7.9	26	714	13
23-Apr-97	8.3	26.4	620	10
16-May-05	8.1	26.4	718	10
17-Jul-96	8.3	26.6	605	9
14-Mar-05	8.1	27	758	8
6-Mar-06	7.7	27.3	702	17
7-Mar-05	8.1	27.4	922	11
17-May-95	8.3	27.5	317	9
8-Dec-2014	2	27.57	490	6
4-Oct-00	8	27.6	708	11
13-Jun-2011	16.2	27.8	621	20
15-Apr-98	8.2	27.9	500	13
10-Oct-96	8.2	28	684	14
27-Nov-96	8	28.1	774	8
13-Feb-2012	1.4	28.7	391	9
6-Oct-2008	11	29	559	9
21-Apr-2008	13.3	29.2	805	9
2-May-2016	10.6	29.3	480	8
23-Apr-2007	12.3	29.5	612	4
21-Jan-2008	1.4	29.6	887	12
3-Mar-99	8.2	29.8	590	13
29-May-96	8.4	30	611	11
19-Apr-04	8.3	30.4	718	7
10-Dec-01	7.9	30.5	624	8
6-Dec-04	8	30.7	734	14
10-Apr-96	8	30.8	643	8
6-Apr-00	8.2	31	780	9

23-Jun-2008	18.2	31.2	671	29
25-Jan-2016	0.9	31.3	500	8
8-Dec-99	8.5	31.4	747	19
25-Apr-01	8.2	31.5	656	11
6-Mar-96	8.1	31.7	712	8
8-Nov-04	8	31.8	682	19
22-Apr-98	8.2	31.9	500	18
27-Mar-06	8	32.3	684	4
21-Apr-94	7.8	32.6	764	12
26-Apr-04	8	33.3	679	12
18-Sep-06	7.8	33.7	712	16
15-Jan-97	8.1	33.8	675	9
27-Nov-06	7.9	34	672	10
1-Mar-94	8	34.3	653	8
7-Mar-2016	3.4	34.5	440	8
1-Feb-95	8	34.6	752	13
21-Mar-2016	5	35.2	420	10
8-May-96	8.1	35.7	626	20
4-Mar-2013	2.4	35.8	700	4
9-Apr-98	8.2	35.9	500	23
6-Jun-2011	18.9	36	602	18
5-Nov-2012	6.6	36.5	680	17
17-May-2010	13.3	36.6	714	13
4-Jan-2016	0.2	36.7	480	13
11-Apr-05	8.1	36.9	642	11
8-Jul-2013	22.1	37.7	583	12
14-Jul-2008	17.1	38	500	18
10-Nov-03	8	38.1	694	22
17-Jan-01	8.1	38.3	909	6
22-Jan-07	8.2	38.5	948	13
9-Mar-94	7.6	38.7	616	15
24-Mar-99	8.2	38.8	420	7
21-Feb-01	8.2	39	676	8
20-Apr-2009	10.6	39.4	690	5
18-Feb-02	8.1	39.7	783	3
8-Apr-2013	6.3	39.8	605	10
8-Dec-03	8.3	40	728	9
10-Feb-99	8.1	40.3	490	12
19-Mar-2012	12.7	41.5	503	10
29-Nov-2010	3.5	41.6	791	12
12-Mar-2007	0.3	41.7	933	13

14-Feb-05	7.9	41.7	991	13
11-Dec-96	8	41.9	710	16
3-May-95	8.3	42.3	631	8
11-May-2009	12.6	42.4	653	11
7-Feb-01	8.4	42.5	773	7
18-Apr-2011	6.8	42.6	698	4
5-Dec-05	8	42.7	672	16
18-Apr-01	8	42.8	633	8
31-Jan-96	8.2	43	638	8
18-Mar-02	7.9	43.7	622	5
29-Jan-97	8.1	43.7	692	5
3-Apr-96	8.1	44.6	628	9
13-Feb-06	7.9	45.2	823	1
14-Mar-2016	7.3	45.4	430	6
3-Feb-99	8.2	45.7	590	18
3-Mar-2008	2.7	45.8	810	4
16-Apr-2018	11.2	46.2	460	7
21-Jan-2013	1.1	46.3	550	11
31-Oct-2011	9.2	46.5	520	12
23-Jan-2012	3.2	46.6	650	22
15-May-96	8.3	46.8	600	11
15-Dec-03	7.9	46.9	761	12
23-Mar-2009	5.3	47	616	12
16-Apr-2007	3.7	47.2	650	12
25-Feb-2008	1	47.3	781	11
12-Mar-2012	5.9	47.4	430	8
31-Jul-06	7.7	47.7	544	21
13-Nov-96	8.1	47.8	696	11
18-Mar-98	8	47.9	762	21
7-Mar-01	8.3	48	683	11
10-Apr-06	8.1	48.2	600	15
4-May-94	7.9	48.4	629	10
24-Apr-06	8	48.77	524	37
31-Jan-01	8.1	49	897	28
25-Feb-2013	2.1	49.1	644	12
17-May-04	8	49.4	666	4
10-Sep-96	8.1	49.5	386	38
5-Nov-01	8.1	49.8	583	44
26-Jun-96	8.2	50	603	17
14-Feb-96	7.4	50.1	568	19
16-Mar-2015	1.5	50.6	441	39

19-Jul-00	8.2	50.8	594	21
13-Nov-06	7.9	51.1	648	12
29-Apr-2013	11.5	51.3	690	12
27-Feb-2012	3	51.5	600	4
10-Jan-05	7.9	51.6	678	7
1-Mar-04	7.8	52.2	920	26
20-Oct-2014	11.3	52.42	560	10
22-Jun-00	8.1	53.1	681	32
5-Jun-06	8	53.3	441	35
11-Dec-06	8.4	53.6	908	20
5-Feb-97	8.1	53.8	705	16
2-May-05	8	54	632	14
30-Jun-2008	15.9	54.5	507	23
5-May-03	7.9	54.8	695	12
9-Feb-2009	1.9	55	885	18
9-Jan-06	8	55.3	688	16
3-May-04	8	55.9	671	16
8-Apr-02	8.5	56.1	584	4
18-Jul-05	7.8	56.2	418	22
9-Aug-00	8.3	56.7	688	13
26-May-2014	16.1	56.83	535	8
26-Apr-95	7.8	56.9	555	30
17-Mar-2014	0.2	58	430	5
27-Sep-00	7.7	58.2	652	32
17-Mar-2008	2	58.4	810	7
21-Oct-2013	12	58.6	520	16
16-Oct-06	8	59	626	22
20-Mar-06	7.9	59.1	649	12
26-May-03	7.8	59.2	676	11
28-Mar-2011	3	59.5	700	10
4-Feb-02	8.2	59.6	760	7
7-Dec-93	7.7	59.8	698	26
8-Feb-2016	4.1	59.9	470	15
30-Nov-93	7.7	60.1	702	22
24-Nov-03	8	60.2	642	30
12-May-03	7.3	60.4	698	16
20-Nov-96	8.2	60.5	719	7
17-May-94	8.2	61.5	641	12
11-Apr-2011	11.5	62.1	648	15
1-Jun-2009	15.1	64.3	636	15
18-Dec-06	8.4	64.5	674	22

20-Mar-96	8.1	65	668	17
12-Apr-2010	9.8	66.3	691	18
13-May-02	7.9	67.5	600	34
2-Mar-00	8.4	69	635	8
25-Mar-98	8.1	69.1	672	16
4-Dec-96	7.9	69.8	674	11
14-Mar-01	8.3	70.1	679	30
5-Apr-04	8.2	70.2	656	19
4-Apr-2011	7	70.5	753	18
24-Jan-96	8.2	70.6	716	19
13-Apr-94	7.6	71.3	644	13
15-Mar-04	7.9	71.4	696	17
1-Dec-2008	2.3	72.2	587	14
17-Jan-05	7.9	72.4	608	15
9-Jan-2012	2.7	72.6	540	21
19-Jun-96	8.3	72.7	576	25
15-Sep-2014	14.6	72.98	600	28
16-Jan-06	8.8	73.8	657	19
11-Nov-2013	6.5	74.4	570	14
11-Apr-01	7.9	74.7	558	17
28-May-01	7.7	75	566	16
22-Mar-95	8.2	76	605	17
3-Jun-2013	17.8	76.4	753	35
12-Dec-2011	2.9	77.5	540	21
25-Apr-05	8	78	632	29
25-Jan-95	7.7	79.4	640	34
30-Nov-2015	4.2	79.9	370	21
5-May-2014	9	79.96	440	10
17-Nov-03	8.2	80.3	686	38
28-Mar-01	8.2	81	580	8
20-Nov-06	7.9	81.1	575	22
6-Oct-2014	13.3	81.68	550	18
22-Nov-95	7.8	82.3	671	15
25-May-00	8.1	82.5	612	16
6-Apr-94	7.6	83.4	571	9
2-May-2011	10.8	83.6	618	19
19-Mar-97	8.2	85.1	598	12
17-Oct-01	7.9	85.3	519	20
24-Oct-96	8.1	85.3	621	21
28-Jan-02	7.9	85.5	671	5
12-Mar-97	8.2	86.8	592	29

14-Apr-2008	7.3	87.1	891	1
14-Jan-2008	4.1	89	700	27
24-Oct-2011	11.1	92.4	490	27
2-Oct-96	8	93.2	574	28
15-Nov-95	7.8	93.4	596	43
18-Mar-2013	2.2	95.8	690	8
31-Oct-96	8.1	98	618	32
14-Apr-03	7.9	101	587	17
5-Mar-97	8.2	104	564	16
27-Jan-99	8.2	105	460	16
19-Mar-2007	1	105.1	537	15
15-Jan-07	8	105.8	699	31
23-Oct-06	7.9	106.6	529	20
8-Sep-2014	18.6	107.36	450	37
23-Jan-06	8	107.6	602	21
18-Feb-99	7.6	109	500	27
29-Nov-00	8.3	113	709	31
17-Dec-96	7.9	115	706	40
6-Apr-2009	7.5	115.9	602	41
7-Jan-2008	4.9	116.4	814	33
14-Apr-2014	9.9	118.86	550	14
3-Dec-01	7.9	121	503	27
4-Apr-2016	3.2	124.8	420	28
5-Dec-2011	7.4	128.1	520	42
1-Feb-2016	3	134.3	550	36
5-Mar-2012	1.7	138.6	330	24
26-Feb-97	8.1	139	560	22
11-Apr-2016	4	141.1	460	17
8-Jan-07	8.4	141.8	598	40
15-Dec-2008	5.5	144.6	732	42
6-Feb-06	7.8	146.3	685	27
29-Mar-04	7.7	147	563	25
23-Feb-94	7.9	151	481	23
20-Feb-06	7.6	154.2	569	11
26-Jun-00	8.1	157	672	34
7-Apr-2008	6.6	157.2	732	15
22-Feb-2016	1.9	160	400	33
16-May-2011	12.5	161.9	516	42
24-Apr-96	7.8	165	570	37
4-Mar-02	8.2	171	579	29
31-Mar-2014	2.2	172.7	452	44

Appendix D Sample spreadsheet of calculations implemented to find the various statistical measures adopted for this study over the training period for the best scenario case.

Date	T ^{0}C	Flow cms	C µS/cm	SSC mg/l	SSC$_i$	$(SSC_i - SSC)^2$	$\lvert SSC_i - SSC\rvert$	$(SSC_i - \overline{SSC})^2$
8-Sep-93	18.5	9.73	590	11	10.413	0.344	0.587	6.661
27-Oct-93	11.9	17.6	686	10	11.076	1.158	1.076	12.823
30-Nov-93	3.9	60.1	702	22	16.237	33.218	5.764	70.880
7-Dec-93	5.3	59.8	698	26	19.924	36.924	6.077	154.233
5-Jan-94	1.8	12.4	825	8	7.506	0.244	0.494	31.147
1-Mar-94	2.6	34.3	653	8	12.534	20.561	4.534	31.147
9-Mar-94	1.8	38.7	616	15	12.547	6.020	2.454	2.014
6-Apr-94	5.8	83.4	571	9	18.668	93.476	9.668	20.985
13-Apr-94	9.7	71.3	644	13	22.668	93.466	9.668	0.338
21-Apr-94	10.7	32.6	764	12	10.406	2.542	1.595	2.499
4-May-94	11.4	48.4	629	10	14.554	20.737	4.554	12.823
17-May-94	11.8	61.5	641	12	14.403	5.774	2.403	2.499
24-May-94	16.7	21.7	665	11	10.170	0.690	0.831	6.661
10-Aug-94	20.7	12.8	683	18	9.216	77.152	8.784	19.528
24-Aug-94	22.4	10.9	693	12	9.133	8.220	2.867	2.499
6-Sep-94	18.9	9.5	691	16	11.250	22.559	4.750	5.852
21-Sep-94	20.8	8.57	697	13	10.743	5.096	2.257	0.338
5-Oct-94	14	9.78	677	16	11.580	19.541	4.421	5.852
12-Oct-94	14.9	9.17	699	14	11.720	5.200	2.280	0.176
11-Jan-95	3.8	10.5	864	7	7.762	0.581	0.762	43.309
25-Jan-95	4.2	79.4	640	34	18.930	227.117	15.070	416.938
1-Feb-95	7.3	34.6	752	13	13.809	0.654	0.809	0.338
8-Feb-95	2.2	17.7	888	5	9.479	20.059	4.479	73.633
23-Feb-95	4.2	18.3	832	11	9.603	1.951	1.397	6.661
22-Mar-95	6.7	76	605	17	17.690	0.476	0.690	11.690
29-Mar-95	8.7	22.1	705	11	11.409	0.167	0.409	6.661
12-Apr-95	10.4	22.6	718	17	11.191	33.746	5.809	11.690
19-Apr-95	11.4	19.2	722	10	11.105	1.220	1.105	12.823
3-May-95	11.7	42.3	631	8	13.533	30.614	5.533	31.147
17-May-95	18.1	27.5	317	9	8.938	0.004	0.062	20.985
24-May-95	17.6	19.1	675	11	10.215	0.616	0.785	6.661
20-Jun-95	25.3	9.57	678	13	9.339	13.406	3.661	0.338
26-Jul-95	26	9.79	664	8	8.267	0.071	0.267	31.147
30-Aug-95	24	7.37	648	12	10.041	3.840	1.960	2.499
6-Sep-95	23.7	7.05	620	10	10.723	0.522	0.723	12.823
20-Sep-95	18.6	8.89	664	8	11.201	10.246	3.201	31.147
4-Oct-95	17.7	8	693	10	11.728	2.984	1.728	12.823
18-Oct-95	13.5	11.6	670	10	11.479	2.189	1.479	12.823
15-Nov-95	5.1	93.4	596	43	29.074	193.922	13.926	865.480

22-Nov-95	5.4	82.3	671	15	21.107	37.291	6.107	2.014
13-Dec-95	1.6	21.3	788	8	7.554	0.199	0.446	31.147
24-Jan-96	3.2	70.6	716	19	20.781	3.172	1.781	29.366
14-Feb-96	1.4	50.1	568	19	14.150	23.524	4.850	29.366
6-Mar-96	2.2	31.7	712	8	11.289	10.816	3.289	31.147
13-Mar-96	4	26	714	13	7.704	28.043	5.296	0.338
20-Mar-96	3.8	65	668	17	18.391	1.935	1.391	11.690
10-Apr-96	6	30.8	643	8	9.053	1.109	1.053	31.147
8-May-96	12.2	35.7	626	20	11.338	75.029	8.662	41.204
15-May-96	11.7	46.8	600	11	15.090	16.724	4.090	6.661
29-May-96	16.8	30	611	11	16.548	30.783	5.548	6.661
19-Jun-96	20.2	72.7	576	25	23.054	3.788	1.946	130.395
26-Jun-96	20.3	50	603	17	20.322	11.032	3.322	11.690
3-Jul-96	22.3	16.4	656	5	9.962	24.621	4.962	73.633
17-Jul-96	23.8	26.6	605	9	7.615	1.917	1.385	20.985
7-Aug-96	25.3	9.11	665	7	8.717	2.947	1.717	43.309
14-Aug-96	24.7	8.95	662	6	9.684	13.575	3.684	57.471
10-Sep-96	19.2	49.5	386	38	37.938	0.004	0.062	596.290
2-Oct-96	16.1	93.2	574	28	26.731	1.612	1.270	207.909
10-Oct-96	13.5	28	684	14	8.934	25.662	5.066	0.176
16-Oct-96	13.5	20.7	668	8	10.481	6.155	2.481	31.147
24-Oct-96	10.7	85.3	621	21	23.454	6.023	2.454	55.042
31-Oct-96	9	98	618	32	27.755	18.022	4.245	339.261
13-Nov-96	5	47.8	696	11	11.659	0.434	0.659	6.661
20-Nov-96	4.4	60.5	719	7	17.135	102.722	10.135	43.309
27-Nov-96	3	28.1	774	8	8.098	0.010	0.098	31.147
4-Dec-96	5	69.8	674	11	17.488	42.092	6.488	6.661
11-Dec-96	4.5	41.9	710	16	9.975	36.307	6.026	5.852
17-Dec-96	5.6	115	706	40	43.929	15.440	3.929	697.966
15-Jan-97	1.5	33.8	675	9	12.669	13.460	3.669	20.985
5-Feb-97	2.4	53.8	705	16	10.699	28.098	5.301	5.852
26-Feb-97	2	139	560	22	21.111	0.791	0.889	70.880
5-Mar-97	3.5	104	564	16	11.448	20.724	4.552	5.852
12-Mar-97	3.2	86.8	592	29	22.717	39.472	6.283	237.747
19-Mar-97	2.9	85.1	598	12	20.059	64.952	8.059	2.499
23-Apr-97	9.8	26.4	620	10	10.846	0.715	0.846	12.823
28-May-97	16.7	25.4	486	9	9.747	0.557	0.747	20.985
18-Mar-98	4.7	47.9	762	21	19.011	3.956	1.989	55.042
25-Mar-98	4.8	69.1	672	16	17.912	3.656	1.912	5.852
9-Apr-98	9.6	35.9	500	23	14.837	66.633	8.163	88.718
15-Apr-98	10.1	27.9	500	13	12.859	0.020	0.141	0.338
22-Apr-98	13.8	31.9	500	18	9.407	73.840	8.593	19.528
14-May-98	18.3	9.54	590	15	10.528	19.995	4.472	2.014
20-May-98	22.7	11.5	600	13	13.108	0.012	0.108	0.338
27-May-98	19.5	9.78	590	16	10.008	35.906	5.992	5.852

3-Jun-98	19.7	9.5	670	16	10.511	30.132	5.489	5.852
15-Jul-98	25.6	7.58	600	18	13.826	17.419	4.174	19.528
5-Aug-98	23.4	6.18	700	11	11.704	0.495	0.704	6.661
19-Aug-98	23.9	7.08	400	13	12.912	0.008	0.088	0.338
30-Sep-98	19.8	6.6	740	5	11.309	39.807	6.309	73.633
7-Oct-98	16.7	7.72	720	12	11.605	0.156	0.395	2.499
14-Oct-98	15.2	8.49	680	11	11.829	0.688	0.829	6.661
28-Oct-98	14.4	6.52	690	5	11.306	39.766	6.306	73.633
2-Dec-98	9.2	7.87	620	18	9.813	67.035	8.188	19.528
27-Jan-99	1.7	105	460	16	17.772	3.138	1.772	5.852
10-Feb-99	3	40.3	490	12	13.425	2.031	1.425	2.499
18-Feb-99	2.9	109	500	27	27.431	0.186	0.431	180.071
3-Mar-99	3.3	29.8	590	13	12.231	0.591	0.769	0.338
17-Mar-99	5.7	23.9	510	8	10.181	4.757	2.181	31.147
24-Mar-99	4.3	38.8	420	7	8.482	2.196	1.482	43.309
7-Apr-99	10.3	22.4	540	6	10.366	19.064	4.366	57.471
14-Apr-99	10.9	19.7	470	2	4.946	8.680	2.946	134.118
28-Apr-99	12.8	21.9	510	6	10.889	23.906	4.889	57.471
12-May-99	16.6	9.72	570	11	11.144	0.021	0.144	6.661
19-May-99	20.2	6.62	650	5	10.638	31.789	5.638	73.633
9-Jun-99	24.6	9.65	580	14	12.317	2.833	1.683	0.176
28-Jun-99	26.1	8.37	610	10	14.978	24.783	4.978	12.823
7-Jul-99	26.5	6.2	570	11	11.742	0.551	0.742	6.661
21-Jul-99	24.6	6.45	540	13	12.065	0.874	0.935	0.338
11-Aug-99	22.2	6.21	570	8	8.863	0.745	0.863	31.147
8-Sep-99	22.9	8.62	630	8	10.420	5.858	2.420	31.147
15-Sep-99	20.6	6.32	580	11	10.937	0.004	0.063	6.661
29-Sep-99	20.1	7.56	600	8	9.967	3.870	1.967	31.147
6-Oct-99	14	7.1	610	14	11.433	6.592	2.567	0.176
13-Oct-99	16	8.87	610	11	11.702	0.493	0.702	6.661
19-Oct-99	14.6	7.17	640	11	11.689	0.475	0.689	6.661
27-Oct-99	10.9	8.52	610	19	9.988	81.218	9.012	29.366
24-Nov-99	12.2	9.56	660	19	10.774	67.670	8.226	29.366
8-Dec-99	5.9	31.4	747	19	14.129	23.731	4.871	29.366
19-Jan-00	1.1	17.5	809	9	6.559	5.957	2.441	20.985
2-Mar-00	4.6	69	635	8	13.693	32.408	5.693	31.147
15-Mar-00	7.4	25.8	805	11	11.694	0.482	0.694	6.661
22-Mar-00	8.9	23.8	779	6	11.350	28.621	5.350	57.471
6-Apr-00	8.2	31	780	9	12.487	12.156	3.487	20.985
25-May-00	15.8	82.5	612	16	19.353	11.242	3.353	5.852
31-May-00	18.4	25	688	7	10.658	13.383	3.658	43.309
5-Jun-00	17.7	17.3	698	11	10.645	0.126	0.355	6.661
22-Jun-00	20.4	53.1	681	32	32.617	0.380	0.617	339.261
19-Jul-00	19.9	50.8	594	21	18.477	6.366	2.523	55.042
26-Jul-00	21.9	20.5	643	10	12.806	7.871	2.806	12.823

16-Aug-00	22.5	22.2	697	19	10.970	64.486	8.030	29.366
30-Aug-00	22.9	11.3	696	11	8.927	4.299	2.073	6.661
27-Sep-00	15	58.2	652	32	21.456	111.176	10.544	339.261
4-Oct-00	16.7	27.6	708	11	9.471	2.339	1.529	6.661
11-Oct-00	12.6	23.1	741	14	10.867	9.819	3.134	0.176
8-Nov-00	11.6	10.8	757	12	10.302	2.884	1.698	2.499
22-Nov-00	3	25	815	7	12.473	29.953	5.473	43.309
29-Nov-00	2	113	709	31	33.521	6.354	2.521	303.423
10-Jan-01	1.1	18.5	834	8	5.783	4.915	2.217	31.147
17-Jan-01	2.1	38.3	909	6	14.506	72.344	8.506	57.471
24-Jan-01	2.5	22.5	842	13	11.618	1.911	1.382	0.338
31-Jan-01	2.6	49	897	28	18.677	86.915	9.323	207.909
7-Feb-01	2.7	42.5	773	7	9.689	7.231	2.689	43.309
21-Feb-01	1	39	676	8	12.378	19.165	4.378	31.147
7-Mar-01	3.4	48	683	11	10.281	0.516	0.719	6.661
14-Mar-01	3.3	70.1	679	30	18.417	134.159	11.583	269.585
28-Mar-01	4.1	81	580	8	16.867	78.620	8.867	31.147
11-Apr-01	9.5	74.7	558	17	23.870	47.194	6.870	11.690
25-Apr-01	10.5	31.5	656	11	10.288	0.507	0.712	6.661
7-May-01	15.6	16.9	687	6	11.302	28.111	5.302	57.471
14-May-01	15.4	15.2	685	9	11.598	6.752	2.598	20.985
28-May-01	14.4	75	566	16	14.637	1.857	1.363	5.852
6-Jun-01	14.8	19.4	601	17	10.474	42.583	6.526	11.690
13-Jun-01	21	21	584	11	12.810	3.275	1.810	6.661
4-Jul-01	21.5	9.34	590	13	11.288	2.933	1.713	0.338
23-Jul-01	25	7.67	631	14	11.613	5.697	2.387	0.176
1-Oct-01	15	8.33	573	8	11.429	11.759	3.429	31.147
10-Oct-01	13.4	12.5	545	8	10.472	6.112	2.472	31.147
17-Oct-01	11.5	85.3	519	20	22.455	6.027	2.455	41.204
5-Nov-01	8.7	49.8	583	44	19.346	607.810	24.654	925.318
3-Dec-01	7.7	121	503	27	31.104	16.842	4.104	180.071
10-Dec-01	5.5	30.5	624	8	7.637	0.132	0.363	31.147
14-Jan-02	2.5	24.3	693	7	10.811	14.524	3.811	43.309
28-Jan-02	3.6	85.5	671	5	19.293	204.278	14.293	73.633
4-Feb-02	1.3	59.6	760	7	8.882	3.542	1.882	43.309
18-Feb-02	0.6	39.7	783	3	9.457	41.692	6.457	111.957
6-May-02	14	25.8	653	11	9.363	2.681	1.638	6.661
13-May-02	11.6	67.5	600	34	22.277	137.431	11.723	416.938
10-Jun-02	20.3	21.6	567	8	10.360	5.571	2.360	31.147
17-Jun-02	17.6	18.6	628	9	10.778	3.163	1.778	20.985
8-Jul-02	24.6	8.53	605	12	12.991	0.981	0.991	2.499
9-Sep-02	22.6	6.86	695	3	11.135	66.173	8.135	111.957
7-Oct-02	17.2	8.66	654	5	11.757	45.658	6.757	73.633
18-Nov-02	5.5	20.1	788	4	10.453	41.646	6.453	91.795
25-Nov-02	6.3	15.1	691	5	11.429	41.326	6.429	73.633

9-Dec-02	1.6	9.88	912	8	8.991	0.983	0.991	31.147
20-Jan-03	1.2	9.23	1037	6	5.609	0.153	0.391	57.471
17-Feb-03	0.5	10.7	991	7	7.105	0.011	0.105	43.309
14-Apr-03	7.7	101	587	17	23.040	36.478	6.040	11.690
12-May-03	13.1	60.4	698	16	7.706	68.785	8.294	5.852
26-May-03	14.9	59.2	676	11	13.124	4.509	2.124	6.661
2-Jun-03	15.3	21.5	722	10	10.692	0.479	0.692	12.823
16-Jun-03	20	20.2	720	10	8.938	1.127	1.062	12.823
14-Jul-03	20.6	9.02	788	13	11.311	2.852	1.689	0.338
11-Aug-03	23.8	20.6	644	18	14.230	14.211	3.770	19.528
18-Aug-03	22.4	14.9	595	13	13.127	0.016	0.127	0.338
25-Aug-03	22.9	8.5	674	11	9.643	1.841	1.357	6.661
8-Sep-03	20.7	6.44	736	16	11.662	18.817	4.338	5.852
15-Sep-03	22.1	9.76	740	11	11.095	0.009	0.095	6.661
22-Sep-03	18.1	16.1	504	7	8.355	1.836	1.355	43.309
29-Sep-03	15.7	14.1	599	10	11.007	1.014	1.007	12.823
6-Oct-03	11.9	16.3	636	9	10.815	3.295	1.815	20.985
10-Nov-03	5.9	38.1	694	22	15.644	40.399	6.356	70.880
17-Nov-03	6.9	80.3	686	38	28.110	97.820	9.890	596.290
24-Nov-03	9.9	60.2	642	30	14.231	248.665	15.769	269.585
8-Dec-03	2	40	728	9	10.774	3.145	1.774	20.985
15-Dec-03	2.5	46.9	761	12	10.447	2.411	1.553	2.499
1-Mar-04	2.4	52.2	920	26	19.009	48.878	6.991	154.233
15-Mar-04	2.6	71.4	696	17	18.098	1.205	1.098	11.690
5-Apr-04	4.9	70.2	656	19	15.306	13.646	3.694	29.366
19-Apr-04	14.4	30.4	718	7	8.327	1.761	1.327	43.309
26-Apr-04	12.3	33.3	679	12	8.689	10.961	3.311	2.499
17-May-04	14.5	49.4	666	4	12.836	78.077	8.836	91.795
19-Jul-04	19.9	13.1	788	10	11.190	1.416	1.190	12.823
9-Aug-04	20.9	12.2	717	9	9.927	0.859	0.927	20.985
16-Aug-04	20.6	8.9	754	6	11.420	29.377	5.420	57.471
13-Sep-04	20.4	10.2	669	9	9.856	0.733	0.856	20.985
20-Sep-04	15.8	8.14	758	10	11.142	1.305	1.142	12.823
4-Oct-04	16	7.68	764	10	10.973	0.947	0.973	12.823
18-Oct-04	9.7	9.39	726	4	9.239	27.450	5.239	91.795
8-Nov-04	7.3	31.8	682	19	13.511	30.127	5.489	29.366
15-Nov-04	5	12.9	789	8	9.151	1.324	1.151	31.147
29-Nov-04	5	23.2	757	19	12.190	46.383	6.811	29.366
6-Dec-04	3.2	30.7	734	14	8.350	31.923	5.650	0.176
10-Jan-05	1.6	51.6	678	7	10.200	10.240	3.200	43.309
17-Jan-05	0	72.4	608	15	12.009	8.946	2.991	2.014
14-Feb-05	1	41.7	991	13	12.618	0.146	0.382	0.338
7-Mar-05	2	27.4	922	11	10.656	0.118	0.344	6.661
14-Mar-05	0	27	758	8	9.448	2.095	1.448	31.147
21-Mar-05	0	25.6	842	12	6.926	25.744	5.074	2.499

11-Apr-05	7.9	36.9	642	11	13.570	6.603	2.570	6.661
18-Apr-05	8.9	20.5	694	10	11.165	1.357	1.165	12.823
25-Apr-05	4.3	78	632	29	18.646	107.199	10.354	237.747
2-May-05	5.8	54	632	14	14.566	0.320	0.566	0.176
9-May-05	11.7	24.2	661	7	10.249	10.559	3.249	43.309
16-May-05	8.2	26.4	718	10	12.161	4.669	2.161	12.823
13-Jun-05	21.3	10.3	669	12	9.318	7.196	2.683	2.499
20-Jun-05	13.5	14.7	625	11	11.053	0.003	0.053	6.661
27-Jun-05	20.4	8.36	692	14	10.841	9.977	3.159	0.176
25-Jul-05	19.9	9.43	493	17	16.268	0.537	0.733	11.690
8-Aug-05	17.4	7.76	474	14	13.498	0.252	0.502	0.176
15-Aug-05	17.3	8.84	660	16	11.748	18.077	4.252	5.852
29-Aug-05	16.8	7.83	677	12	11.892	0.012	0.108	2.499
19-Sep-05	15	8.16	619	10	11.777	3.156	1.777	12.823
17-Oct-05	15.1	13.8	629	12	11.319	0.464	0.681	2.499
24-Oct-05	12	8.81	744	19	10.249	76.578	8.751	29.366
12-Dec-05	1.2	16	1020	11	6.506	20.198	4.494	6.661
16-Jan-06	1.5	73.8	657	19	10.748	68.089	8.252	29.366
23-Jan-06	2.5	107.6	602	21	23.255	5.087	2.255	55.042
6-Feb-06	1.5	146.3	685	27	26.098	0.815	0.903	180.071
6-Mar-06	2.2	27.3	702	17	11.291	32.597	5.709	11.690
20-Mar-06	8.7	59.1	649	12	18.693	44.796	6.693	2.499
27-Mar-06	8.6	32.3	684	4	12.320	69.216	8.320	91.795
10-Apr-06	6.6	48.2	600	15	11.880	9.737	3.120	2.014
24-Apr-06	11.2	48.77	524	37	23.356	186.151	13.644	548.452
15-May-06	14.2	21.1	714	8	10.822	7.966	2.822	31.147
5-Jun-06	16.1	53.3	441	35	36.042	1.086	1.042	458.776
19-Jun-06	21	13.1	751	10	10.349	0.122	0.349	12.823
26-Jun-06	21	15.8	714	11	8.758	5.025	2.242	6.661
10-Jul-06	20.4	10.1	752	10	11.249	1.560	1.249	12.823
17-Jul-06	22.9	13.9	589	12	12.866	0.751	0.866	2.499
24-Jul-06	19	12.6	684	7	10.538	12.514	3.538	43.309
31-Jul-06	21.7	47.7	544	21	20.938	0.004	0.063	55.042
28-Aug-06	19.5	10	681	9	10.703	2.901	1.703	20.985
18-Sep-06	17.5	33.7	712	16	12.263	13.964	3.737	5.852
16-Oct-06	7.7	59	626	22	20.814	1.406	1.186	70.880
23-Oct-06	8.8	106.6	529	20	20.526	0.277	0.526	41.204
13-Nov-06	6.3	51.1	648	12	16.008	16.066	4.008	2.499
20-Nov-06	5.6	81.1	575	22	17.070	24.310	4.931	70.880
27-Nov-06	6.1	34	672	10	14.458	19.870	4.458	12.823
11-Dec-06	2	53.6	908	20	19.386	0.377	0.614	41.204
18-Dec-06	5.2	64.5	674	22	18.847	9.943	3.153	70.880
8-Jan-07	4.3	141.8	598	40	37.789	4.891	2.212	697.966
15-Jan-07	2.9	105.8	699	31	29.041	3.838	1.959	303.423
22-Jan-07	0	38.5	948	13	9.384	13.075	3.616	0.338

29-Jan-07	-0.9	24.9	917	10	2.889	50.562	7.111	12.823
5-Feb-07	-1	17	994	7	7.646	0.417	0.646	43.309
12-Feb-07	0.2	17.1	952	9	9.052	0.003	0.052	20.985
12-Mar-2007	0.3	41.7	933	13	10.490	6.300	2.510	0.338
19-Mar-2007	1	105.1	537	15	13.961	1.080	1.039	2.014
16-Apr-2007	3.7	47.2	650	12	10.185	3.293	1.815	2.499
23-Apr-2007	12.3	29.5	612	4	10.102	37.232	6.102	91.795
28-May-2007	16.8	15.9	667	9	10.987	3.947	1.987	20.985
4-Jun-2007	20.4	20.3	759	6	8.833	8.026	2.833	57.471
11-Jun-2007	19.7	15.7	670	11	9.398	2.565	1.602	6.661
25-Jun-2007	21.3	9.4	699	8	10.315	5.358	2.315	31.147
9-Jul-2007	23.2	7.3	750	14	12.937	1.131	1.063	0.176
23-Jul-2007	17	6.3	715	21	11.425	91.688	9.575	55.042
27-Aug-2007	19.7	8.8	696	20	11.062	79.881	8.938	41.204
10-Sep-2007	19.4	7.7	757	11	11.153	0.024	0.153	6.661
17-Sep-2007	14.7	7.5	855	14	9.535	19.937	4.465	0.176
24-Sep-2007	17.1	7.2	818	10	9.793	0.043	0.207	12.823
15-Oct-2007	13.6	7.6	852	14	9.580	19.536	4.420	0.176
29-Oct-2007	11.1	7.9	787	13	9.475	12.429	3.526	0.338
19-Nov-2007	6.5	6.5	775	6	7.557	2.424	1.557	57.471
10-Dec-2007	2.7	18.8	896	5	8.969	15.749	3.969	73.633
7-Jan-2008	4.9	116.4	814	33	31.803	1.434	1.197	377.099
14-Jan-2008	4.1	89	700	27	21.419	31.146	5.581	180.071
21-Jan-2008	1.4	29.6	887	12	11.840	0.025	0.160	2.499
25-Feb-2008	1	47.3	781	11	9.028	3.889	1.972	6.661
17-Mar-2008	2	58.4	810	7	12.297	28.056	5.297	43.309
14-Apr-2008	7.3	87.1	891	1	12.440	130.876	11.440	158.280
21-Apr-2008	13.3	29.2	805	9	10.063	1.130	1.063	20.985
9-Jun-2008	21.2	13.5	684	6	8.795	7.809	2.795	57.471
16-Jun-2008	16.8	18.7	657	7	10.447	11.883	3.447	43.309
23-Jun-2008	18.2	31.2	671	29	15.135	192.230	13.865	237.747
30-Jun-2008	15.9	54.5	507	23	17.565	29.537	5.435	88.718
14-Jul-2008	17.1	38	500	18	17.301	0.489	0.699	19.528
25-Aug-2008	19.4	8.2	655	13	10.745	5.087	2.256	0.338
26-Sep-2008	15.7	15	626	6	11.041	25.416	5.041	57.471
6-Oct-2008	11	29	559	9	11.616	6.842	2.616	20.985
3-Nov-2008	7.7	25.9	659	10	12.435	5.927	2.435	12.823
1-Dec-2008	2.3	72.2	587	14	14.527	0.278	0.527	0.176
15-Dec-2008	5.5	144.6	732	42	42.782	0.612	0.782	807.642
23-Mar-2009	5.3	47	616	12	10.362	2.682	1.638	2.499
6-Apr-2009	7.5	115.9	602	41	38.458	6.460	2.542	751.804
20-Apr-2009	10.6	39.4	690	5	9.997	24.970	4.997	73.633
11-May-2009	12.6	42.4	653	11	12.228	1.507	1.228	6.661
1-Jun-2009	15.1	64.3	636	15	25.121	102.441	10.121	2.014
15-Jun-2009	18.7	24.1	704	7	9.905	8.438	2.905	43.309

20-Jul-2009	19.1	12.7	772	12	11.363	0.405	0.637	2.499
5-Oct-2009	12.4	15.6	706	9	11.223	4.940	2.223	20.985
26-Oct-2009	10.7	14.4	799	13	10.242	7.609	2.759	0.338
9-Nov-2009	10.3	16.6	788	9	10.478	2.183	1.478	20.985
18-Jan-2010	2.7	13.7	948	5	7.092	4.375	2.092	73.633
8-Feb-2010	0.6	18.4	855	6	5.325	0.456	0.675	57.471
29-Mar-2010	7.8	25.7	759	24	11.971	144.706	12.029	108.557
12-Apr-2010	9.8	66.3	691	18	11.808	38.346	6.192	19.528
19-Apr-2010	10	22.7	750	6	11.240	27.462	5.240	57.471
26-Apr-2010	11.7	19	735	11	11.133	0.018	0.133	6.661
17-May-2010	13.3	36.6	714	13	7.705	28.034	5.295	0.338
28-Jun-2010	22.9	11.8	709	7	8.980	3.921	1.980	43.309
12-Jul-2010	23.6	23.6	699	12	12.643	0.413	0.643	2.499
13-Sep-2010	17	8.7	750	8	11.413	11.648	3.413	31.147
27-Sep-2010	15.1	8.6	775	6	10.909	24.099	4.909	57.471
18-Oct-2010	11.8	22	732	10	11.039	1.080	1.039	12.823
8-Nov-2010	6.7	17.8	781	5	9.876	23.779	4.876	73.633
29-Nov-2010	3.5	41.6	791	12	16.761	22.671	4.761	2.499
28-Mar-2011	3	59.5	700	10	13.992	15.933	3.992	12.823
4-Apr-2011	7	70.5	753	18	23.051	25.517	5.051	19.528
11-Apr-2011	11.5	62.1	648	15	11.913	9.530	3.087	2.014
18-Apr-2011	6.8	42.6	698	4	15.723	137.429	11.723	91.795
2-May-2011	10.8	83.6	618	19	23.282	18.339	4.282	29.366
6-Jun-2011	18.9	36	602	18	20.586	6.686	2.586	19.528
13-Jun-2011	16.2	27.8	621	20	13.397	43.604	6.603	41.204
20-Jun-2011	19.5	15.4	639	8	9.789	3.199	1.789	31.147
18-Jul-2011	24.9	9.1	650	8	9.457	2.124	1.457	31.147
22-Aug-2011	20.2	11.4	644	12	9.407	6.724	2.593	2.499
29-Aug-2011	20.4	14	669	8	9.094	1.197	1.094	31.147
19-Sep-2011	16.8	8.4	680	10	11.906	3.632	1.906	12.823
24-Oct-2011	11.1	92.4	490	27	25.409	2.531	1.591	180.071
31-Oct-2011	9.2	46.5	520	12	14.270	5.153	2.270	2.499
14-Nov-2011	10.3	19.5	560	7	10.263	10.645	3.263	43.309
5-Dec-2011	7.4	128.1	520	42	42.856	0.733	0.856	807.642
12-Dec-2011	2.9	77.5	540	21	23.087	4.357	2.087	55.042
9-Jan-2012	2.7	72.6	540	21	21.782	0.611	0.782	55.042
23-Jan-2012	3.2	46.6	650	22	10.494	132.390	11.506	70.880
13-Feb-2012	1.4	28.7	391	9	8.220	0.609	0.780	20.985
27-Feb-2012	3	51.5	600	4	13.063	82.140	9.063	91.795
5-Mar-2012	1.7	138.6	330	24	22.522	2.184	1.478	108.557
12-Mar-2012	5.9	47.4	430	8	9.190	1.416	1.190	31.147
19-Mar-2012	12.7	41.5	503	10	12.752	7.571	2.752	12.823
2-Apr-2012	8.1	18	500	8	6.773	1.507	1.227	31.147
11-Jun-2012	22.4	13.3	700	8	8.517	0.267	0.517	31.147
18-Jun-2012	21.8	9.7	702	10	10.084	0.007	0.084	12.823

10-Sep-2012	18.3	11	590	12	10.488	2.287	1.512	2.499
17-Sep-2012	18.3	7.6	650	14	11.376	6.884	2.624	0.176
1-Oct-2012	14	9.8	800	7	10.603	12.984	3.603	43.309
15-Oct-2012	14	9.2	816	12	10.317	2.832	1.683	2.499
29-Oct-2012	10.1	23.7	629	28	10.770	296.866	17.230	207.909
5-Nov-2012	6.6	36.5	680	17	14.819	4.758	2.181	11.690
12-Nov-2012	10.3	11.6	880	10	9.327	0.453	0.673	12.823
26-Nov-2012	4.8	9.9	803	6	8.550	6.502	2.550	57.471
3-Dec-2012	7.6	13.2	760	11	9.074	3.711	1.927	6.661
10-Dec-2012	5.5	25.5	737	8	12.882	23.835	4.882	31.147
21-Jan-2013	1.1	46.3	550	11	12.644	2.702	1.644	6.661
25-Feb-2013	2.1	49.1	644	12	11.322	0.460	0.678	2.499
4-Mar-2013	2.4	35.8	700	4	11.415	54.982	7.415	91.795
18-Mar-2013	2.2	95.8	690	8	14.611	43.699	6.611	31.147
8-Apr-2013	6.3	39.8	605	10	7.357	6.986	2.643	12.823
29-Apr-2013	11.5	51.3	690	12	12.194	0.038	0.194	2.499
13-May-2013	10.7	22.3	750	10	11.157	1.339	1.157	12.823
3-Jun-2013	17.8	76.4	753	35	33.068	3.733	1.932	458.776
24-Jun-2013	23.8	21.6	750	8	8.755	0.569	0.755	31.147
8-Jul-2013	22.1	37.7	583	12	8.669	11.093	3.331	2.499
15-Jul-2013	24.6	24.5	665	12	14.537	6.435	2.537	2.499
29-Jul-2013	18.4	11.9	608	11	10.373	0.393	0.627	6.661
12-Aug-2013	20	9.6	637	8	9.773	3.145	1.773	31.147
26-Aug-2013	22.2	6.7	670	6	10.503	20.277	4.503	57.471
9-Sep-2013	18.7	7	640	8	11.150	9.920	3.150	31.147
30-Sep-2013	17.3	21.8	700	8	9.981	3.924	1.981	31.147
21-Oct-2013	12	58.6	520	16	16.510	0.260	0.510	5.852
11-Nov-2013	6.5	74.4	570	14	15.451	2.107	1.451	0.176
14-Apr-2014	9.9	118.8	550	14	13.893	0.011	0.107	0.176
5-May-2014	9	79.96	440	10	10.140	0.020	0.140	12.823
2-Jun-2014	19.6	19.97	640	2	11.383	88.043	9.383	134.118
16-Jun-2014	20.8	18.22	570	12	9.781	4.924	2.219	2.499
23-Jun-2014	21.8	9.92	680	8	9.437	2.066	1.437	31.147
14-Jul-2014	21.3	25.16	580	13	9.883	9.716	3.117	0.338
11-Aug-2014	21.5	11.7	650	9	8.889	0.012	0.111	20.985
18-Aug-2014	19.5	14.94	600	13	10.866	4.556	2.134	0.338
25-Aug-2014	20.6	13.92	640	12	9.280	7.398	2.720	2.499
8-Sep-2014	18.6	107.3	450	37	37.025	0.001	0.025	548.452
15-Sep-2014	14.6	72.98	600	28	19.302	75.650	8.698	207.909
6-Oct-2014	13.3	81.68	550	18	17.985	0.000	0.015	19.528
20-Oct-2014	11.3	52.42	560	10	21.124	123.739	11.124	12.823
27-Oct-2014	9.3	21.21	570	4	10.945	48.226	6.945	91.795
8-Dec-2014	2	27.57	490	6	7.929	3.721	1.929	57.471
16-Mar-2015	1.5	50.6	441	39	30.965	64.563	8.035	646.128
14-Sep-2015	16.1	13	680	9	11.788	7.772	2.788	20.985

26-Oct-2015	11	10.4	620	10	10.246	0.060	0.246	12.823
16-Nov-2015	7.7	23.3	440	8	2.250	33.066	5.750	31.147
30-Nov-2015	4.2	79.9	370	21	18.787	4.896	2.213	55.042
4-Jan-2016	0.2	36.7	480	13	10.604	5.742	2.396	0.338
25-Jan-2016	0.9	31.3	500	8	7.971	0.001	0.029	31.147
1-Feb-2016	3	134.3	550	36	38.571	6.612	2.571	502.614
8-Feb-2016	4.1	59.9	470	15	20.045	25.455	5.045	2.014
7-Mar-2016	3.4	34.5	440	8	11.204	10.264	3.204	31.147
21-Mar-2016	5	35.2	420	10	5.637	19.036	4.363	12.823
4-Apr-2016	3.2	124.8	420	28	25.359	6.975	2.641	207.909
11-Apr-2016	4	141.1	460	17	21.087	16.704	4.087	11.690
18-Apr-2016	11.2	46.2	460	7	12.384	28.989	5.384	43.309
2-May-2016	10.6	29.3	480	8	9.897	3.599	1.897	31.147
6-Jun-2016	18.9	16.8	560	6	11.154	26.559	5.154	57.471
11-Jul-2016	22.2	6.4	760	12	12.861	0.742	0.861	2.499
18-Jul-2016	22.6	7.3	670	14	10.152	14.809	3.848	0.176
12-Sep-2016	18.5	7.4	660	16	11.396	21.196	4.604	5.852
19-Sep-2016	19.7	8.1	660	14	10.675	11.054	3.325	0.176
24-Oct-2016	12	14.5	520	12	9.878	4.502	2.122	2.499
21-Nov-2016	4.1	9	740	10	12.032	4.131	2.032	12.823
5-Dec-2016	4.4	22.8	500	14	9.626	19.133	4.374	0.176
N= 420		$\overline{SSC}$ = 13.851		$\sum$	**9421.165**	**1451.709**	**25990.248**	

Appendix E Sample spreadsheet of calculations implemented to find the various statistical measures adopted for this study over the testing period for the best scenario case.

Date	T ^{0}C	Flow cms	C µS/cm	SSC mg/l	SSC$_i$	$(SSC_i - SSC)^2$	$\lvert SSC_i - SSC \rvert$	$(SSC_i - \overline{SSC})^2$
23-Feb-94	2	151	481	23	22.269	0.533	0.730	63.362
26-Apr-95	9	56.9	555	30	23.800	38.435	6.200	223.802
31-Jan-96	1.4	43	638	8	12.234	17.932	4.235	49.562
3-Apr-96	6.5	44.6	628	9	11.173	4.724	2.174	36.482
24-Apr-96	8.1	165	570	37	36.977	0.001	0.023	482.242
29-Jan-97	3	43.7	692	5	10.303	28.123	5.303	100.802
2-Sep-98	24.1	5.1	720	17	13.399	12.966	3.601	3.842
9-Sep-98	21	5.17	710	11	11.881	0.777	0.882	16.322
16-Sep-98	21.6	6.05	740	11	12.273	1.621	1.273	16.322
23-Sep-98	20.6	5.03	780	17	10.957	36.518	6.043	3.842
21-Oct-98	14.1	5.79	730	5	10.686	32.340	5.687	100.802
3-Feb-99	2.4	45.7	590	18	12.693	28.163	5.307	8.762
23-Jun-99	23	5.87	620	13	11.032	3.872	1.968	4.162
14-Jul-99	24.2	6.05	540	11	12.347	1.815	1.347	16.322
28-Jul-99	26.5	5.51	510	5	5.4609	0.212	0.461	100.802
4-Aug-99	24.2	5.46	480	9	8.5614	0.192	0.439	36.482
18-Aug-99	22.8	5.68	540	10	11.598	2.554	1.598	25.402
1-Sep-99	21.7	5.49	600	8	11.519	12.384	3.519	49.562
22-Sep-99	17.3	5.82	630	13	11.739	1.588	1.260	4.162
26-Jun-00	21.9	157	672	34	33.816	0.034	0.183	359.482
9-Aug-00	23.6	56.7	688	13	14.822	3.320	1.822	4.162
18-Apr-01	8.6	42.8	633	8	15.526	56.647	7.526	49.562
8-Aug-01	26.8	5.66	642	20	20.099	0.010	0.099	24.602
13-Aug-01	24.8	5.73	660	11	8.3739	6.896	2.626	16.322
5-Sep-01	21.1	5.94	641	7	10.404	11.589	3.404	64.642
4-Mar-02	1.3	171	579	29	27.973	1.053	1.026	194.882
18-Mar-02	5.5	43.7	622	5	8.7669	14.190	3.767	100.802
8-Apr-02	7.1	56.1	584	4	11.977	63.633	7.977	121.882
5-May-03	13.4	54.8	695	12	9.8108	4.793	2.189	9.242
29-Mar-04	7.8	147	563	25	24.190	0.655	0.809	99.202
3-May-04	8.8	55.9	671	16	18.948	8.691	2.948	0.922
18-Jul-05	20.2	56.2	418	22	22.124	0.016	0.125	48.442
12-Sep-05	16.2	6.02	702	14	11.500	6.250	2.500	1.082
5-Dec-05	3	42.7	672	16	10.635	28.783	5.365	0.922
9-Jan-06	3	55.3	688	16	11.906	16.757	4.094	0.922
13-Feb-06	1.5	45.2	823	1	9.0234	64.375	8.023	197.122
20-Feb-06	0.6	154.2	569	11	13.886	8.329	2.886	16.322
3-Mar-2008	2.7	45.8	810	4	13.529	90.811	9.530	121.882
7-Apr-2008	6.6	157.2	732	15	14.823	0.031	0.177	0.002

9-Feb-2009	1.9	55	885	18	19.661	2.760	1.661	8.762
16-May-2011	12.5	161.9	516	42	42.038	0.001	0.038	726.842
9-Jul-2012	24	6.1	890	14	15.430	2.046	1.431	1.082
17-Mar-2014	0.2	58	430	5	8.2794	10.754	3.279	100.802
31-Mar-2014	2.2	172.7	452	44	39.813	17.528	4.187	838.682
26-May-2014	16.1	56.83	535	8	14.338	40.182	6.339	49.562
28-Sep-2015	19.4	4.8	810	10	9.1174	0.779	0.883	25.402
22-Feb-2016	1.9	160	400	33	34.985	3.941	1.985	322.562
14-Mar-2016	7.3	45.4	430	6	5.0535	0.896	0.947	81.722
25-Jul-2016	23.8	5.6	800	17	15.770	1.513	1.230	3.842
8-Aug-2016	21.6	4.9	790	12	12.024	0.001	0.025	9.242
N= 50		$\overline{SSC}$ = 15.04			$\sum$	**692.012**	**141.129**	**4943.920**

Appendix F Sample spreadsheet of calculations implemented to find the uncertainty analysis measures adopted for this study over the testing period for the best scenario case.

Observed SSC (mg/l)	Predicted S_{MLR}	Predicted S_{ANFIS}	Predicted S_{ANN}	ei MLR	│T-e_i│ MLR	Z MLR	ei ANFIS	│T-ei│ ANFIS	Z ANFIS	ei ANN	│T-ei│ ANN	Z ANN		ei MLR	ei ANFIS	ei ANN
1.000	13.004	10.901	9.023	1.114	1.095		1.037	0.994		0.955	0.953					
4.000	16.094	16.482	11.977	0.605	0.585		0.615	0.572		0.476	0.474					
4.000	13.389	11.768	13.530	0.525	0.505	2.325	0.469	0.426	1.875	0.529	0.527			0.525	0.469	
5.000	12.945	15.518	10.303	0.413	0.394	1.812	0.492	0.449	1.977	0.314	0.312			0.413	0.492	
5.000	8.689	6.596	10.687	0.240	0.221	1.015	0.120	0.077	0.340	0.330	0.328			0.240	0.120	
5.000	11.381	12.790	5.461	0.357	0.338	1.555	0.408	0.365	1.607	0.038	0.036	0.335		0.357	0.408	0.038
5.000	13.470	13.939	8.767	0.430	0.411	1.891	0.445	0.402	1.772	0.244	0.242	2.251		0.430	0.445	0.244
5.000	14.611	17.741	8.279	0.466	0.447	2.054	0.550	0.507	2.234	0.219	0.217	2.020		0.466	0.550	0.219
6.000	14.004	9.307	5.054	0.368	0.349	1.605	0.191	0.148	0.650	-0.075	0.077	-0.716		0.368	0.191	-0.075
7.000	10.304	6.515	10.404	0.168	0.149	0.684	-0.031	0.074	-0.327	0.172	0.170	1.582		0.168	-0.031	0.172
8.000	12.371	13.928	12.235	0.189	0.170	0.782	0.241	0.198	0.871	0.184	0.182	1.698		0.189	0.241	0.184
8.000	10.321	10.916	11.519	0.111	0.091	0.420	0.135	0.092	0.405	0.158	0.156	1.454		0.111	0.135	0.158
8.000	14.063	14.294	15.526	0.245	0.226	1.038	0.252	0.209	0.921	0.288	0.286			0.245	0.252	
8.000	18.342	19.441	14.339	0.360	0.341	1.569	0.386	0.343	1.509	0.253	0.251	2.340		0.360	0.386	0.253
9.000	13.883	14.165	11.174	0.188	0.169	0.777	0.197	0.154	0.678	0.094	0.092	0.854		0.188	0.197	0.094
9.000	10.785	13.525	8.561	0.079	0.059	0.273	0.177	0.134	0.590	-0.022	0.024	-0.224		0.079	0.177	-0.022
10.000	10.555	12.090	11.598	0.023	0.004	0.019	0.082	0.039	0.173	0.064	0.062	0.579		0.023	0.082	0.064
10.000	9.871	6.071	9.117	-0.006	0.025	-0.114	-0.217	0.260	-1.145	-0.040	0.042	-0.395		-0.006	-0.217	-0.040
11.000	10.215	9.923	11.882	-0.032	0.051	-0.236	-0.045	0.088	-0.387	0.033	0.031	0.291		-0.032	-0.045	0.033

11.000	10.554	9.832	12.273	-0.018	0.037	-0.171	-0.049	0.092	-0.405	0.048	0.045	0.422		-0.018	-0.049	0.048
11.000	10.960	12.112	12.347	-0.002	0.021	-0.096	0.042	0.001	-0.006	0.050	0.048	0.446		-0.002	0.042	0.050
11.000	11.178	10.313	8.374	0.007	0.012	-0.056	-0.028	0.071	-0.313	-0.118	0.121	-1.125		0.007	-0.028	-0.118
11.000	32.552	26.576	13.886	0.471	0.452	2.079	0.383	0.340	1.498	0.101	0.099	0.921		0.471	0.383	0.101
12.000	17.494	16.420	9.811	0.164	0.144	0.665	0.136	0.093	0.410	-0.087	0.090	-0.837		0.164	0.136	-0.087
12.000	10.398	9.332	12.025	-0.062	0.081	-0.375	-0.109	0.152	-0.671	0.001	0.001	-0.013		-0.062	-0.109	0.001
13.000	10.726	10.718	11.032	-0.084	0.103	-0.473	-0.084	0.127	-0.559	-0.071	0.074	-0.686		-0.084	-0.084	-0.071
13.000	9.355	10.601	11.740	-0.143	0.162	-0.746	-0.089	0.132	-0.580	-0.044	0.047	-0.434		-0.143	-0.089	-0.044
13.000	20.293	16.968	14.822	0.193	0.174	0.801	0.116	0.073	0.320	0.057	0.055	0.509		0.193	0.116	0.057
14.000	9.206	7.795	11.500	-0.182	0.201	-0.926	-0.254	0.297	-1.311	-0.085	0.088	-0.818		-0.182	-0.254	-0.085
14.000	11.307	6.096	15.431	-0.093	0.112	-0.515	-0.361	0.404	-1.781	0.042	0.040	0.372		-0.093	-0.361	0.042
15.000	34.729	28.916	14.823	0.365	0.345	1.589	0.285	0.242	1.066	-0.005	0.007	-0.069		0.365	0.285	-0.005
16.000	16.562	17.658	18.948	0.015	0.004	-0.019	0.043	0.000	-0.001	0.073	0.071	0.663		0.015	0.043	0.073
16.000	12.739	11.615	10.635	-0.099	0.118	-0.544	-0.139	0.182	-0.803	-0.177	0.180	-1.674		-0.099	-0.139	-0.177
16.000	15.074	10.886	11.907	-0.026	0.045	-0.208	-0.167	0.210	-0.927	-0.128	0.131	-1.217		-0.026	-0.167	-0.128
17.000	10.959	9.853	13.399	-0.191	0.210	-0.965	-0.237	0.280	-1.234	-0.103	0.106	-0.985		-0.191	-0.237	-0.103
17.000	10.170	9.348	10.957	-0.223	0.242	-1.115	-0.260	0.303	-1.334	-0.191	0.193	-1.799		-0.223	-0.260	-0.191
17.000	11.067	9.390	15.770	-0.186	0.206	-0.946	-0.258	0.301	-1.326	-0.033	0.035	-0.325		-0.186	-0.258	-0.033
18.000	13.056	15.198	12.693	-0.139	0.159	-0.730	-0.073	0.117	-0.514	-0.152	0.154	-1.435		-0.139	-0.073	-0.152
18.000	14.971	17.201	19.661	-0.080	0.099	-0.457	-0.020	0.063	-0.277	0.038	0.036	0.336		-0.080	-0.020	0.038
20.000	11.627	10.470	20.099	-0.236	0.255	-1.172	-0.281	0.324	-1.428	0.002	0.000	-0.001		-0.236	-0.281	0.002
22.000	19.084	24.310	22.125	-0.062	0.081	-0.372	0.043	0.000	0.001	0.002	0.000	0.001		-0.062	0.043	0.002
23.000	32.204	22.570	22.270	0.146	0.127	0.584	-0.008	0.051	-0.226	-0.014	0.016	-0.152		0.146	-0.008	-0.014
25.000	32.956	32.604	24.191	0.120	0.101	0.464	0.115	0.072	0.318	-0.014	0.017	-0.155		0.120	0.115	-0.014
29.000	35.821	35.062	27.974	0.092	0.073	0.334	0.082	0.039	0.173	-0.016	0.018	-0.167		0.092	0.082	-0.016
30.000	16.667	12.771	23.800	-0.255	0.274	-1.263	-0.371	0.414	-1.824	-0.101	0.103	-0.958		-0.255	-0.371	-0.101
33.000	33.746	24.833	34.985	0.010	0.010	-0.044	-0.123	0.167	-0.734	0.025	0.023	0.215		0.010	-0.123	0.025
34.000	38.311	38.468	33.817	0.052	0.033	0.150	0.054	0.011	0.046	-0.002	0.005	-0.043		0.052	0.054	-0.002

37.000	36.346	28.677	36.977	-0.008	0.027	-0.124	-0.111	0.154	-0.677	0.000	0.003	-0.024		-0.008	-0.111	0.000
42.000	36.776	41.640	42.038	-0.058	0.077	-0.354	-0.004	0.047	-0.206	0.000	0.002	-0.018		-0.058	-0.004	0.000
44.000	36.211	48.764	39.813	-0.085	0.104	-0.478	0.045	0.002	0.007	-0.043	0.046	-0.426		-0.085	0.045	-0.043

	MLR	ANFIS	ANN
T	0.019	0.043	0.002
MAD	0.147	0.153	0.072
S_{MAD}	0.217	0.227	0.107
$\bar{e}$	0.074	0.045	0.009
S_e	0.209	0.231	0.105
$2S_e$	0.418	0.461	0.210
+	3.099	3.209	1.656
-	0.453	0.384	0.628

Note that the $|Z|$ score values of greater than 2.5 order of magnitude has been removed as per the procedure described at section 4.1.3.4 of this study

Curriculum Vitae

Name:	Issam Mohamed
Post-secondary Education and Degrees:	University of Benghazi Benghazi, Libya 2007 B.Sc. in Civil Engineering
	Western University London, Ontario, Canada 2018 M.ESc. in Civil and Environmental Engineering
Related Work Experience	Master's Research Student Upper Thames River Conservation Authority (UTRCA) London, Ontario, Canada 1993-1995

Publications:

Mohamed, I., and Shah, I. (2018). Modeling of Suspended Sediment Concentration Using Conventional and Machine Learning Approaches, in Thames River, Canada. In International Conference on Water Management Modeling. Toronto, Canada: CHI.

Mohamed, I., and Shah, I. (2018). Suspended Sediment Concentration Modeling Using Conventional and Machine Learning Approaches in Thames River, London, Canada. Manuscript submitted for publication.