

**INTEGRATED GEOPHYSICAL, HYDROGEOCHEMICAL
AND ARTIFICIAL INTELLIGENCE TECHNIQUES
FOR GROUNDWATER STUDY IN THE
LANGAT BASIN, MALAYSIA**

MAHMOUD KHAKI

**FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2014

UNIVERSITY MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **MAHMOUD KHAKI**

Passport No: **B19472127**

Registration/Matric No: **SHC 100098**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Thesis: **“INTEGRATED GEOPHYSICAL, HYDROGEOCHEMICAL AND ARTIFICIAL INTELLIGENCE TECHNIQUES FOR GROUNDWATER STUDY IN THE LANGAT BASIN, MALAYSIA”**

Field of Study: **GEOPHYSICS**

I do solemnly and sincerely declare that:

- (1) I am the sole author of this work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

(Candidate Signature)

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name **ASSOC. PROF. DR ISMAIL BIN YUSOFF**

Designation

Witness’s Signature

Date:

Name **DR. NUR ISLAMI BIN RAHMAN**

Designation

Abstract

Geophysical, hydrogeochemical and artificial intelligence techniques were used to study the groundwater characteristics and their associated problems in Langat Basin, Malaysia. Resistivity surveys and geochemical analyses were used to delineate regions of Langat Basin that are contaminated by brackish water. Hydrogeochemical data of groundwater samples collected from seventeen wells from 2008 to 2013 were analysed. Ninety eight geoelectrical resistivity survey measurements were conducted to obtain subsurface resistivity data. The Wenner array was selected because of its sensitivity in detecting vertical changes in subsurface resistivity. The resistivity imaging results show that the upper layer is usually clay and below this layer is an aquifer with various depths of 10 to 30 m, and the layer thickness changes from 10 to 45 m, respectively, from east to west across the study area. The depth to bedrock varies from 30 m up to 65 m. The results learned from the resistivity survey confirmed the pattern of a continuous structure of layers, as detected from the borehole and geological information. Chemical analyses show the total dissolved solid exceeds 1000 mg/L in the west and is less than 1000 mg/L in the east of study area. Furthermore, the results of the resistivity survey and those from the hydrogeochemical analyses show that the groundwater within the study area is a mixture of brackish and freshwater zones.

A novel investigation in modelling of groundwater level and quality using Artificial Neural Networks (ANNs) and Adaptive neuro fuzzy inference systems (ANFIS) methods was developed in the study area. Water table modelling based on ANNs and ANFIS technology were developed to simulate the water table fluctuations based on the relationship between the variations of rainfall, humidity, evaporation, minimum and maximum temperature and water table depth. The mean square errors and correlation coefficient of the water table depth models for 84 months were between 0.0043 to 0.107 and 0.629 to 0.99 respectively for all models. Evaluating the results of the various kinds

of models show the earned results of the ANFIS model are superior to those gained from ANNs in which they are both more precise and with less error. Furthermore, four common training functions; Gradient descent with momentum and adaptive learning rate back propagation, Levenberg-Marquardt algorithm, Resilient back propagation, Scaled conjugate gradient were compared for the modelling of groundwater level. These results confirm that, for all the networks the Levenberg-Marquardt algorithm is the most effective algorithm to model the groundwater level. This study also developed the potential of the ANFIS and ANN to simulate total dissolved solid (TDS) and electrical conductivity (EC) by employing the values of other existing water quality parameters from five sampling stations over six years from 2008 to 2013. A good agreement between simulate values and their respective measured values in the quality of the groundwater were found. TDS and EC values predicted from the model accompanying obtained result present increasing in concentration continually in the future whereas for well no 3 (the nearest well to costal line) reaches at 32625.51 mg/L and 69501.76 $\mu\text{S}/\text{cm}$ in 2025, respectively.

Abstrak

Intergrasi teknik geofizik, hidrogeokimia dan kepintaran buatan telah digunakan bagi kajian mengenalpasti ciri-ciri air tanah dan masalah yang berkaitan di Lembangan Langat, Malaysia. Survei pengimejan kerintangan elektrik dan analisa geokimia telah digunakan untuk menentukan kawasan Lembangan Langat yang tercemar oleh air payau. Analisa hidrogeokimia dilakukan kepada sampel air tanah yang diambil dari 17 telaga dari tahun 2008 sehingga 2013. Sebanyak 98 survei kerintangan geoelektrik telah dilakukan bagi mendapatkan data kerintangan bawah permukaan. Konfigurasi Wenner dipilih kerana kejituaannya dalam mengenal pasti perubahan kerintangan bawah permukaan secara menegak. Keputusan pengimejan kerintangan menunjukkan lapisan atas adalah lempung dan lapisan dibawahnya adalah akuifer dengan kedalaman yang pelbagai di antara 10 sehingga 30 m dan ketebalan lapisan yang berubah dari 10 sehingga 44 m, masing-masing dari timur ke barat melintasi kawasan kajian. Kedalaman ke batuan dasar berbeza dari 30 m sehingga menjangkau 65 m. Hasil keputusan dari survei keberintangan mengesahkan corak lapisan struktur yang berterusan seperti yang diperolehi dari maklumat lubang gerudi dan geologi. Analisa kimia menunjukkan jumlah pepejal terlarut melebihi 1000 mg/L dan nilainya semakin meningkat bagi kawasan pantai dan biasanya kurang 1000 mg/L menjauhi pantai. Tambahan lagi, keputusan survey keberintangan dan juga analisa hidrogeokimia sampel air tanah menunjukkan air tanah dan tanah dalam lingkungan kawasan kajian adalah dari percampuran antara zon air payau dan air segar.

Penyiasatan yang novel bagi permodelan aras air tanah dan kualiti menggunakan kaedah Artificial Neural Networks (ANNs) dan Adaptive neuro fuzzy inference systems (ANFIS) telah dibentuk bagi kawasan kajian. Permodelan muka aras air menggunakan teknologi ANNs dan ANFIS telah dibentuk bagi tujuan simulasi perubahan aras muka air berdasarkan perkaitan di antara variasi hujan, kelembapan, sejatan, suhu minimum

dan suhu maksimum dan kedalaman ke aras muka air. Mean square errors dan pekali korelasi bagi model kedalaman ke muka aras air selama 84 bulan adalah di antara 0.0043 sehingga 0.107 dan 0.629 dan 0.629 sehingga 0.99, masing-masing bagi semua model. Berdasarkan penilaian keputusan dari pelbagai model menunjukkan keputusan yang diperoleh dari model ANFIS adalah yang terbaik berbanding dari ANNs walaupun kedua-duanya lebih tepat dengan sedikit ketidakpastian. Tambahan pula, empat training function yang biasa digunakan seperti Gradient descent with momentum and adaptive learning rate backpropagation, Levenberg-Marquardt algorithm, Resilient back propagation, Scaled conjugate gradient telah dibandingkan dalam permodelan muka aras air tanah. Hasil yang diperolehi menunjukkan bahawa dari kesemua rangkaian algoritma hanya algoritma Levenberg-Marquardt yang lebih efektif bagi permodelan aras air tanah. Kajian ini juga telah membentuk potensi ANFIS dan ANN untuk mensimulasi jumlah pepejal terlarut (TDS) dan aras kekonduksian elektrik (EC) dengan menggunakan nilai parameter kualiti air yang sedia ada bagi lima stesen persempelan selama lebih enam tahun dari tahun 2008 sehingga 2013. Perkaitan yang rapat telah ditunjukkan diantara nilai-nilai kualiti air tanah yang diukur satu sama lain. Nilai TDS dan EC yang diramalkan daripada model menunjukkan peningkatan kepekatan secara berterusan sehingga tahun 2025. Sebagai contoh, telaga no.3, telaga yang paling hampir dengan kawasan persisir pantai boleh mencapai kepekatan 32625.51 mg/L dan 69501.76 $\mu\text{S}/\text{cm}$ bagi TDS dan EC, masing-masing pada tahun 2025 seperti yang diramalkan oleh model.

Acknowledgements

It is a pleasure to convey my gratitude to my supervisor, Prof. Dr. Ismail Bin Yusoff, for his supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. Above all and the most needed, he supports me in various ways. His truly scientist intuition have made him as a constant oasis of ideas and passions in science, which exceptionally inspire and enrich my growth as a student, a researcher and a scientist want to be. I am indebted to him more than he knows. Also, I would like to sincerely thank my supervisor, Dr. Nur Islami Bin Rahman, for his guidance and support throughout this study, and especially for his confidence in me. His comments and questions were very beneficial in my completion of this research.

I would like to sincerely thank to Assoc. Prof. Dr. Samsudin Bin Hj Taib for his cooperation and support for the use of geophysical equipments.

I am writing to express my sincere gratitude to Ministry of Education Malaysia for making the Malaysia International Scholarship (MIS) possible. It was an honor to be chosen as a recipient of this prestigious scholarship. In addition, I would like to gratefully thank Ms. Mazatul Akmar Binti Aros and Mr. Tan Chai Oo from the Minerals and Geoscience Department Malaysia for their helps.

The financial support through the University of Malaya research grants no PV112-2012A is gratefully acknowledged. Thanks are extended to the Department of Geology, Faculty of Science, and University of Malaya for all the provided facilities and assistance. Thank to the Department of Mineral and Geosciences of Malaysia.

Table of Contents

| | |
|---|-------------|
| Abstract | ii |
| Abstrak | iv |
| Acknowledgements | vi |
| Table of Contents | vii |
| List of Figures | xi |
| List of Tables | xvii |
| CHAPTER 1: Introduction | 1 |
| 1.1 Introduction..... | 1 |
| 1.2 Description of the Area..... | 3 |
| 1.3 Groundwater Condition in the Selangor State | 3 |
| 1.4 Problem Statements | 5 |
| 1.5 Aims and Objectives..... | 6 |
| 1.6 Survey Method and Analyses | 6 |
| 1.7 Thesis Outline | 7 |
| CHAPTER 2: Study Area | 10 |
| 2.1 Introduction..... | 10 |
| 2.2 Geomorphology..... | 12 |
| 2.2.1 Mountainous Area..... | 13 |
| 2.2.2 Hilly area | 13 |
| 2.2.3 Flat lowlands area | 13 |
| 2.3 Geology..... | 14 |
| 2.4 Hydrology and Hydrogeology..... | 18 |
| 2.4.1 Hydrology..... | 19 |
| 2.4.1.1 Climate..... | 19 |
| 2.4.1.2 Temperature | 20 |
| 2.4.1.3 Humidity | 20 |
| 2.4.1.4 Rainfall..... | 20 |
| 2.4.1.5 Water Balance | 22 |
| 2.4.2 Hydrogeology | 23 |
| 2.4.2.1 Aquifer | 23 |
| 2.4.2.2 Water recharge | 24 |
| 2.4.2.3 Water Quality | 24 |
| 2.4.2.4 Water Demand..... | 26 |

| | |
|--|-----------|
| Chapter 3: Research Methodology..... | 28 |
| 3.1 Introduction..... | 28 |
| 3.2 Geoelectrical Resistivity Method..... | 29 |
| 3.2.1 Basic Principles..... | 29 |
| 3.2.2 Geometric Factor and Apparent Resistivity..... | 34 |
| 3.2.4 Basic Inversion Theory..... | 37 |
| 3.2.5 Electrode Array..... | 38 |
| 3.2.6 Resistivity surveying equipment..... | 39 |
| 3.2.7 Two Dimensional Resistivity Modelling..... | 42 |
| 3.2.8 Data Acquisition..... | 43 |
| 3.2.9 Depth of Investigation..... | 44 |
| 3.2.10 Considerations and Limitations of the Electrical Resistivity Imaging Method..... | 46 |
| 3.2.11 Electrical Resistivity and Porosity..... | 47 |
| 3.3 Conceptual Model..... | 48 |
| 3.3.1 Borehole Data..... | 48 |
| 3.3.2 Geoelectrical Resistivity and Lithological Correlation in the Study Area..... | 50 |
| 3.3.3 Fence Diagram..... | 52 |
| 3.4 Hydrogeological Method..... | 53 |
| 3.4.1 Water Level Monitoring..... | 53 |
| 3.4.2 Hydrogeochemical Method..... | 53 |
| 3.4.2.1 In-situ Parameters..... | 54 |
| 3.4.2.1.1 Temperature..... | 54 |
| 3.4.2.1.2 pH..... | 54 |
| 3.4.2.1.3 Total Dissolved Solids..... | 55 |
| 3.4.2.1.4 Electrical Conductivity..... | 56 |
| 3.4.2.1.5 Salinity..... | 56 |
| 3.4.2.2 Cation and Anion..... | 57 |
| 3.4.2.2.1 Major Cations..... | 58 |
| 3.4.2.2.2 Major Anions..... | 60 |
| 3.4.2.3 Field Procedures..... | 62 |
| 3.5 Artificial Intelligence Techniques..... | 64 |
| 3.5.1 Introduction..... | 65 |
| 3.5.2 A Framework for an Artificial Network..... | 66 |
| 3.5.2.1 Proceeding Elements..... | 67 |
| 3.5.2.2 Activation Functions..... | 68 |
| 3.5.2.3 Learning Rules..... | 70 |
| 3.5.2.4 The Effects of the Number of Hidden Layers..... | 71 |
| 3.5.2.5 Neural Network Architectures..... | 71 |
| 3.5.2.6 Back Propagation Algorithm..... | 75 |
| 3.5.2.7 Artificial Neural Network Algorithms..... | 76 |
| 3.5.2.8 Training and Testing Neural Networks..... | 78 |
| 3.5.3 Adaptive Neuro Fuzzy Inference System (ANFIS)..... | 81 |

| | |
|--|------------|
| 3.5.3.1 Fuzzy Set and Membership Function | 82 |
| 3.5.3.2 Fuzzy Inference System | 84 |
| 3.5.3.3 ANFIS Architecture and Training | 86 |
| 3.5.3.4 Learning Algorithm of ANFIS | 90 |
| 3.6 Evaluation of the Model | 93 |
| 3.6.1 Data Pre-Processing | 93 |
| 3.6.2 Performance Criteria | 93 |
| Chapter 4: Geophysical and Hydrogeochemical Results and Discussion | 95 |
| 4.1 Introduction | 95 |
| 4.2 Geophysical Study | 96 |
| 4.2.1 Geoelectrical Resistivity Result | 96 |
| 4.2.2 2D Electrical Resistivity Imaging Sections | 97 |
| 4.2.3 Geoelectrical Maps | 110 |
| 4.2.4 Geoelectrical Cross-sections Model | 112 |
| 4.2.5 Time Lapse Electrical Resistivity Monitoring of Groundwater Level | 117 |
| 4.3 Hydrogeochemical Study | 120 |
| 4.3.1 Water Chemistry | 123 |
| 4.4 Groundwater Level Monitoring | 130 |
| 4.5 Summary | 132 |
| Chapter 5: Prediction and Assessment of Groundwater Depth and Quality | 135 |
| 5.1 Introduction | 135 |
| 5.2 A Neural Network and Adaptive Neuro Fuzzy Inference System for Modelling of Groundwater Level (Single Well) | 136 |
| 5.2.1 Data Description | 137 |
| 5.2.2 Building and Training the Network | 142 |
| 5.2.3 Membership Functions (MFs) | 143 |
| 5.2.4 Training and Testing Results for ANN Model | 143 |
| 5.2.5 ANFIS for the Prediction of Water Level | 156 |
| 5.2.6 Training and Testing Results for ANFIS Model | 156 |
| 5.3 A Neural Network and Adaptive Neuro Fuzzy Inference System for Modelling of Groundwater Quality | 160 |
| 5.3.1 The Water Quality Parameters | 161 |
| 5.3.2 Statistical Analyses | 163 |
| 5.3.3 Input Parameters | 164 |
| 5.3.4 Training and Testing Results | 165 |
| 5.4 Prediction Groundwater Level and Quality with Experimental Data | 171 |
| 5.5 Summary | 176 |
| Chapter 6: Conclusions and Recommendations | 179 |

| | |
|---|------------|
| 6.1 Conclusions..... | 179 |
| 6.2 Recommendations | 183 |
| Reference..... | 185 |
| List of Publications and Papers Presented..... | 192 |

List of Figures

| | |
|---|----|
| Figure 1.1: Location map of the study area..... | 4 |
| Figure 1.2: Research outline chart and procedure | 9 |
| Figure 2.1: Geographical position of Langat Basin..... | 11 |
| Figure 2.2: Surface features of Selangor State, Malaysia | 12 |
| Figure 2.3: Map of Peninsular Malaysia showing the three belts (Hutchison et al., 2009) | 16 |
| Figure 2.4: Geological and potential aquifer map of Selangor State Department of Mineral and Geosciences of Malaysia (2007) | 17 |
| Figure 2.5: The location and topographical map of study area | 18 |
| Figure 2.6: Annual precipitation of Telok Datok station 2000-2013 | 21 |
| Figure 2.7: Average of monthly rainfall between 2000 to 2013 in Telok Datok station | 22 |
| Figure 3.1: Parameters used to define Ohm's law for a straight conductor with length (L) and cross section (A)..... | 30 |
| Figure 3.2: The flow of current from a point current source and the potential distribution | 32 |
| Figure 3.3: A generalized four-electrode array | 34 |
| Figure 3.4: Resistivity of common rocks, soil materials and chemicals (Loke, 2013)..... | 37 |
| Figure 3.5: Common electrode arrays used in geoelectrical resistivity surveys | 39 |
| Figure: 3.6: ABEM Terrameter SAS 4000 with its accessories..... | 41 |
| Figure 3.7: (a) Equipment set up for four wheel cables; (b) A typical field arrangement for 2D electrical imaging surveys; and (c) Data cover of standard Wenner using WENNER_L and WENNER_S for roll- along with three stations (ABEM, 2010)..... | 42 |

| | |
|---|-----|
| Figure 3.8: The sensitivity function and median depth of investigation for the Wenner array (Loke, 2013)..... | 45 |
| Figure 3.9: Typical stratigraphic log of the geological environment in the study area | 49 |
| Figure 3.10: Correlation of ERI results based on direct comparison to lithologic logs at two wells. | 51 |
| Figure 3.12: Hand auger with diameter of 11 centimetres..... | 63 |
| Figure 3.13: Typical piezometer with length of 3 meters | 64 |
| Figure 3.14: An input – output mapping..... | 66 |
| Figure 3.15: A typical artificial neuron..... | 68 |
| Figure 3.16: Sigmoid function..... | 69 |
| Figure 3.17: Structure of typical Feed-Forward Neural Network (FNN)..... | 73 |
| Figure 3.18: Structure of typical Elman Neural Network (RNN) | 74 |
| Figure 3.19: Structure of typical Cascade Forward Network (CFN)..... | 75 |
| Figure 3.20: A training process flowchart..... | 80 |
| Figure 3.21: (a) Crisp sets; (b) and fuzzy sets of tall man | 83 |
| Figure 3.22: Example of membership functions for fuzzy sets..... | 84 |
| Figure 3.23: Functional block of a fuzzy inference system diagram (Jang, 1993) | 85 |
| Figure 3.24: Two inputs first-order Sugeno fuzzy model with two rules and architecture of ANFIS..... | 89 |
| Figure 3.25: Flowchart of ANFIS system | 91 |
| Figure 3.26: Methodology flowchart for prediction model | 92 |
| Figure 4.1: Location map of geoelectrical resistivity survey in the study area..... | 97 |
| Figure 4.2: Selected 2D resistivity model of the apparent resistivity data profile 1 | 98 |
| Figure 4.3: Selected 2D resistivity model of the apparent resistivity data profile 2 | 99 |
| Figure 4.4: Selected 2D resistivity model of the apparent resistivity data profile 3 | 100 |

| | |
|---|-----|
| Figure 4.5: Selected 2D resistivity model of the apparent resistivity data profile 4 | 100 |
| Figure 4.6: Selected 2D resistivity model of the apparent resistivity data profile 5 | 101 |
| Figure 4.7: Selected 2D resistivity model of the apparent resistivity data profile 6 | 102 |
| Figure 4.8: Selected 2D resistivity model of the apparent resistivity data profile 7 | 103 |
| Figure 4.9: Selected 2D resistivity model of the apparent resistivity data profile 8 | 104 |
| Figure 4.10: Selected 2D resistivity model of the apparent resistivity data profile 9..... | 105 |
| Figure 4.11: Selected 2D resistivity model of the apparent resistivity data profile 10..... | 106 |
| Figure 4.12: Selected 2D resistivity model of the apparent resistivity data profile 11 | 107 |
| Figure 4.13: Selected 2D resistivity model of the apparent resistivity data profile 12..... | 107 |
| Figure 4.14: Selected 2D resistivity model of the apparent resistivity data profile 13..... | 108 |
| Figure 4.15: Selected 2D resistivity model of the apparent resistivity data profile 14..... | 109 |
| Figure 4.16: Selected 2D resistivity model of the apparent resistivity data profile 15..... | 110 |
| Figure 4.18: Contour distribution map corresponding to aquifer thickness based on the resistivity interpreted..... | 112 |
| Figure 4.19: Geoelectrical cross sections of the study area along line 1 to 3 | 113 |
| Figure 4.20: Geoelectrical cross sections of the study area along line 4 and 5..... | 114 |
| Figure 4.21: Panel diagrams represent the 3D view of the resistivity cross- sections along the study area..... | 115 |
| Figure 4.22: Fence diagram of study area..... | 116 |

| | |
|---|-----|
| Figure 4.23: Field set up groundwater level monitoring..... | 118 |
| Figure 4.24: Resistivity model for monitoring profile 3 from February to August 2013 | 119 |
| Figure 4.25: Well locations for groundwater sampling in the study area..... | 121 |
| Figure 4.26: Contour map of Total dissolved solid level in Langat basin from 2008 to 2013..... | 125 |
| Figure 4.27: Piper diagram showing the average chemical composition of the groundwater samples along different wells in the study area from 2008 to 2013. The detailed chemical data is listed in Table 4.3 | 127 |
| Figure 4.28: Stiff diagram showing the average chemical composition of the groundwater samples along different wells in the study area from 2008 to 2013. The detailed chemical data is listed in Table 4.3 | 129 |
| Figure 4.29: Areal display of Stiff diagrams for wells from the study area. Diagrams are centered on their well locations. | 130 |
| Figure 4.30: Groundwater level (masl) in the study area in March, 2008 to 2013 | 131 |
| Figure 4.31: Groundwater level (masl) in the study area in October, 2008 to 2012 | 132 |
| Figure 5.1: Location of the selected well for modelling of groundwater level in the study area..... | 137 |
| Figure 5.2: Monthly rainfall in the study area from 2007 to 2013. | 139 |
| Figure 5.3: Monthly precipitation (mm) and depth to groundwater level (m) for observation well from 2007 to 2013 in the study area. | 139 |
| Figure 5.4: Minimum, maximum and average monthly temperature (°C) | 140 |
| Figure 5.5: Average monthly evaporation (mm) in the study area..... | 141 |
| Figure 5.6: Average monthly humidity (%) in study area | 141 |
| Figure 5.7: General conceptual neural network for the water level computation in the study area..... | 144 |

| | |
|---|-----|
| Figure 5.8: Scatter plots of the observed and simulated water levels at training period for FNN with different algorithms..... | 147 |
| Figure 5.9: Scatter plots of the observed and simulated water levels at training period for CFN with different algorithms..... | 148 |
| Figure 5.10: Scatter plots of the observed and simulated water levels at training period for RNN with different algorithms | 149 |
| Figure 5.11: The result of the models and the actual value for the FNN approach with LM and RP algorithms..... | 150 |
| Figure 5.12: The result of the models and the actual value for the FNN approach with GDX and SCG algorithms. | 151 |
| Figure 5.13: The result of the models and the actual value for the CFN approach with LM and RP algorithms..... | 152 |
| Figure 5.14: The result of the models and the actual value for the CFN approach with GDX and SCG algorithms. | 153 |
| Figure 5.15: The result of the models and the actual value for the RNN approach with LM and RP algorithms..... | 154 |
| Figure 5.16: The result of the models and the actual value for the RNN approach with GDX and SCG algorithms..... | 155 |
| Figure 5.17: The result of the models and the actual value for the ANFIS approach with Gaussian and Generalized bell MFs. | 158 |
| Figure 5.18: The result of the models and the actual value for the ANFIS with Gaussian and Generalized bell MFs. | 159 |
| Figure 5.19: Location of the selected wells for modelling of groundwater quality (TDS and EC) in the study area. | 163 |
| Figure 5.20: Non-linear model of a neuron..... | 166 |
| Figure 5.21: Scatter plots of the observed and forecasted water quality at training | |

List of Tables

| | |
|--|-----|
| Table 2.1 Statistical data for annual precipitation, Langat Basin..... | 22 |
| Table 3.1: Geometric factor for some common configurations | 35 |
| Table 3.2 Table of existing well information | 49 |
| Table 4.1. Statistical Data for borehole no 3..... | 120 |
| Table 4.2: List of existing well stations which is used for groundwater sampling (Mineral and Geosciences Department of Malaysia) | 121 |
| Table 4.3. Chemical composition of groundwater samples from 2008 to 2013 | 122 |
| Table 4.4. Groundwater classification based on Total dissolved solids (Freeze and Cherry, 1977). | 124 |
| Table 5.1. Characteristic of monthly data sets | 142 |
| Table 5.2. Learning factors at the maximum R and minimum MSE of data set..... | 144 |
| Table 5.3. Comparison of performance of ANN models developed for all, training and testing periods..... | 145 |
| Table 5.4. Comparison of performance of ANFIS models developed for all, training and testing periods..... | 157 |
| Table 5.5. Results of the statistical analyses to show the effect of the factors on samples properties | 164 |
| Table 5.6. The statistics of geochemical parameter variables for five boreholes in the study area..... | 165 |
| Table 5.7. Comparison of performance of models developed for all, training and testing periods..... | 167 |

List of Symbols and Abbreviations

| | |
|-------|---|
| ERI | Electrical Resistivity Imaging |
| AI | Artificial Intelligence |
| ANNs | Artificial Neural Networks |
| ANFIS | Adaptive Neuro Fuzzy Inference System |
| TDS | Total Dissolved Solid |
| EC | Electrical Conductivity |
| MFs | Membership Functions |
| FNN | Feed-Forward Neural Network |
| CFN | Cascade Forward Network |
| RNN | Elman neural network |
| LM | Levenberg-Marquardt |
| RBP | Resilient Back Propagation |
| GDX | Gradient Descent with Momentum and Adaptive Learning Rate Back Propagation |
| SCG | Scaled Conjugate Gradient |
| FIS | Fuzzy Inference Systems |
| R | Correlation Coefficient |
| MSE | Mean Square Error |

CHAPTER 1: Introduction

1.1 Introduction

Recently, accessible water resources have been depleted extravagantly because of increasing demand by domestic and industry water consumption as well as water qualities deterioration caused by pollution. Moreover, the population of Peninsular Malaysia is growing rapidly and it is expected to reach in 2050 more than double of that in 2000. The Malaysian economy should also be taken in account where it has already turned from an agricultural to an industrial one and is projected to grow rapidly to the target of Vision 2020. This momentum of growth in the economy and population will consume much more water (JICA, 2002).

Considering the above-mentioned situations in Peninsular Malaysia especially in the Langat Basin, the current PhD thesis was carried out on planning, development and management for the water resources of Langat Basin.

The Electrical Resistivity Imaging (ERI) method selected for this study is capable of monitoring the water content, internal structure, depth of bedrock and layer thickness from subsurface image. It is also used in complex geological and noisy areas when other geophysical techniques such as seismic refraction and GPR techniques cannot be used (Cosenza et al., 2006; De Vita et al., 2006; Heincke et al., 2010). Geographic Information System (GIS) were also employed to discussion about these factors and the delineation of the study area. The ERI technique is considered to be very sensitive to variation of the water in the subsurface materials (Niesner and Weidinger, 2008).

Multidisciplinary expertise in combination with advanced computational modelling methods is required to understand and analyse the complex relationships that affect the groundwater. The estimation of groundwater resources potential and prediction of

environmental impacts caused by groundwater extraction are the general goals of groundwater simulation. The elaborated simulation model will be employed as a tool to plan for the development of sustainable groundwater resources. The model could make the authorities able to formulate optimal management strategies leading to predict the water resources and develop the area ecologically. Numerical models enable pre-visualization of effects from suggested growth before it occurs can be enabled, which contributes to a more sustainable and thorough environmental management plan with optimal cost and deteriorative implications. Models can be applied as an investigative tool in relation to assessment and development of abatement measures that acquire to be carried out to achieve a specific target quality.

Artificial intelligence models including mathematical formulations, equations and together from different relevant branches of science and engineering to generate a final end product that can be easily used without the arduous task of manual calculations. The accuracy of these models depends on the mathematical formulations' accuracy which is incorporated with the assumption of course that the input data is not a variable of concern. Moreover, the used mathematical formulations may be the primary limiting factor behind a particular computer model.

The present study as an original research project was fulfilled in the most important area in Selangor State, Malaysia. The subject has significantly demonstrated three major purposes and has highlighted the related methodology. Additionally, presented an important technique to improve and operate management in Malaysia. The thesis employed geophysical and hydrogeochemical and computational methodology supported by a high-tech technique. The following section concisely introduces different parts of this study including the research concerned, problem statement, aims and objectives, research outline, and the importance of selected study area.

1.2 Description of the Area

Malaysia is located in the Southeast Asia, in a humid tropical zone with heavy rainfall and high temperatures. It consists of two regions: Peninsular Malaysia and East Malaysia (Malaysian Borneo). The study area is placed in West Malaysia which is called Peninsular Malaysia (Omar and Mansor, 2004; Ramli et al., 2005).

The centre of the study area is located on latitude 02° 48' 59" N and longitude 101° 35' 36" E (Figure 1.1). The study area is known as the Selangor State located in the Langat Basin with an aggregate area of around 300 km². The area for this research is located at the flat lowland in the downstream part of the Langat Basin with the surface elevation of less than 15 above mean sea level. The accessibility within the study area is moderate. And the land is mainly used for agricultural activities such as oil palm and coconut plantation. Quaternary sediments, which consist of unconsolidated gravel, sand, silt and clay of the Simpang Formation in the Pleistocene, and Gula and Beruas Formations in the Holocene, constitute the main aquifer in the study area.

1.3 Groundwater Condition in the Selangor State

The surface waters, such as rivers and streams, are used as main water supplies for domestic, agricultural, and industrial purposes. Groundwater is used on a conjunctive basis where surface supplies are inadequate, or as the only supply source where surface supplies are non-existent. It is known as an important and integral part of hydrological cycle and its existence is related to the rainfall and recharge conditions. In Malaysia, less than 10% of water usage is extracted from groundwater resources. Of the groundwater used, about 70% serves domestic supply, and about 25% and 5% is used for industrial and agricultural purposes, respectively (Karim, 2006). In the Selangor state, 103 groundwater wells have been drilled. The wells are distributed within three

districts of the Langat Basin: Hulu Langat, Kuala Langat, and Sepang. There are five different categories of well: domestic, industrial, observation, test well, and unknown (JMG, 2002). Most of the groundwater abstracted in the Selangor state is utilized for industrial purposes. The largest abstraction of groundwater for industrial use in the country is at the Megasteel/Amsteel factory at Brooklands Estate, in Kuala Langat District. The distributed groundwater potential is categorized into four main categories of aquifer in Malaysia: alluvial, limestone/carbonate, sedimentary and volcanic rocks, and crystalline igneous rocks (Manap et al., 2013).

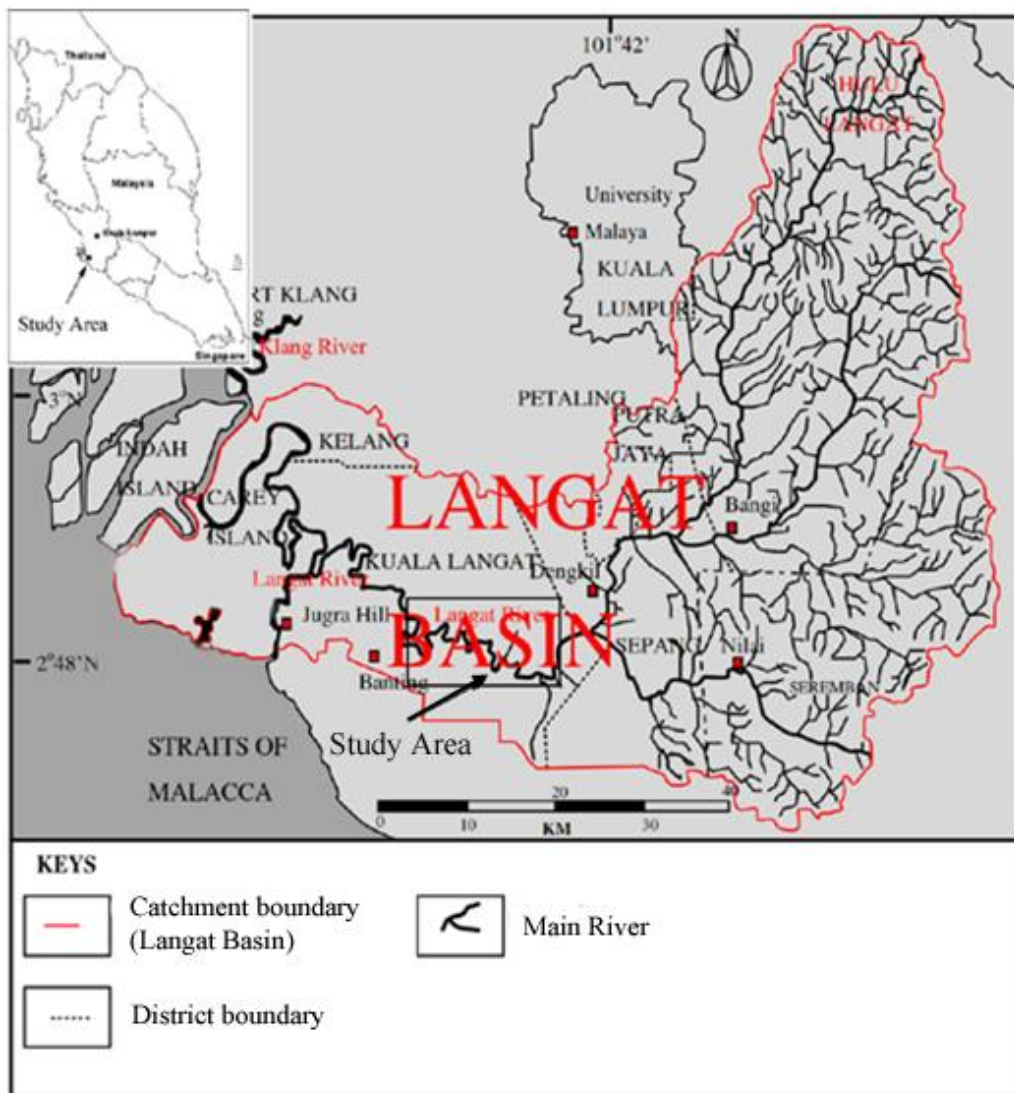


Figure 1.1: Location map of the study area

1.4 Problem Statements

The risk of the groundwater quality hazard can be increased due to several factors such as agricultural activities, sea water intrusion to the aquifer and others activities in the specific area. Groundwater quality is also affected by increasing industrialization and population growth in coming years in the Langat Basin.

The groundwater in the south and west of the study area possibly has effected polluting because of the agricultural activities. Therefore, a region situated below this region is prone to contamination. Furthermore, most of the communities within the area utilize the shallow groundwater for their domestic uses.

In the west part of study area, the occurrence of salt/brackish water in the subsurface is important to be detected. This is necessary because the water resources are obtained from groundwater.

The lack of a conceptual model in the area of Langat Basin is one of the challenging issues.

The high cost of conventional drilling is a significant obstacle in evaluating the depth and thickness of subsurface structure. Therefore, electrical resistivity measurements obtained with the Wenner array provides a more cost-effective method of estimating the depth and thickness of aquifer and depth of bedrock at Langat Basin site in comparison with conventional drilling. Due to the large expense of drilling, boreholes should only be conducted to confirm the thickness and depth of aquifer and depth of bedrock estimated by electrical resistivity imaging.

1.5 Aims and Objectives

The aim of this study is to investigate the applicability of integrated geophysical, hydrogeochemical and artificial intelligence methods to solve the groundwater associated problems. Developing a technique to get a quick financial evaluation of the physical and chemical characteristics of groundwater resources on the research area is the main purpose of this dissertation. The physical resource knowledge will cause to the development of clear, logical, and concise community plans that demonstrates the sustainability concepts in a cost-effective manner. The overall objectives for the research include the following:

- To develop an aquifer conception model, determine the depth of bedrock and study the depth of water table based and geophysical survey results.
- To study hydrogeochemistry and contamination of groundwater and develop the map of interference zone salt and brackish water in the aquifer.
- Integrate the geophysical data with the hydrochemical analyses
- To determine an efficient neural network architecture and learning algorithm.
- To develop computationally efficient and robust approaches for the automatic calibration of groundwater level and quality model. These methods consist of artificial neural network and adaptive neuro fuzzy inference system.
- To provide methodologically standard factors for studying on groundwater in the areas based on the individual geological targets that could be transferred to other area with the same hydrogeological and geological conditions.
- To predict the water quality parameters by ANN and ANFIS models for future.

1.6 Survey Method and Analyses

Variation in a material's physical properties is measured by geophysical methods. The properties of soils and rocks can be classified into a matrix component or

framework and the pore content component. Various materials show different parameter signatures for example their resistivity or its inverse, acoustic velocity, conductivity, density and magnetic permeability. The mineral type, grain packing arrangement, permeability, porosity, and pore content (i.e. gas or fluid type) affect on these (signature) parameters. Generally, no one property is unique to any material; rather a material is explained by ranges of each property. Consequently In most geophysical surveys, it is important that the changes or contrasts in geophysical parameters are measured and that the target presents large differences in property with surrounding material.

In any hydrogeological investigation, a full support from geoelectrical resistivity is desirable and hydrogeochemical analyses methods is required due to obtain subsurface information of the area concerned. However, some subsurface information was found from drilled boreholes in the study area. Other information related to the current research for example rainfall data were obtained by the courtesy of Malaysian Meteorological Department (MMD, 2013).

1.7 Thesis Outline

The general outline of this research is in the following chart (Figure 1.2). Description of research stages is presented in research methodology chapter. The thesis is organized into six chapters. Chapter 1 introduced the background and problem of the study. The aims and objectives of this study are also defined in this chapter.

Chapter 2 description the study area. This chapter deals mainly with the geological, hydrological and hydrogeological of the study area. The first aspect provides the general geology background derived from several references. The hydrology and hydrogeology of the study area is presented at the following chapter. They also consist of climatology, quality of aquifer and so on.

Chapter 3 description the methodology and data acquisition. All the methodologies used in this study are provided in this chapter including geophysical (electrical resistivity imaging), hydrogeochemical and artificial intelligence (artificial neural networks and adaptive neuro fuzzy inference system) techniques. Electrical resistivity theory, equipment, data acquisition and processing are the first presentation, followed by hydrogeochemical and artificial intelligence techniques.

Chapter 4 description results and discussion of electrical resistivity surveys and hydrogeochemical analyses. In this chapter, the study of the groundwater characteristics for the study area will be discussed. The first discussion is on the soil resistivity correlation with different soil characteristics. These results are useful in the geoelectrical interpretation especially with emphasis on the occurrence of salt/brackish water in the aquifer. The next focuses on time lapse electrical resistivity imaging monitoring in the part of study area to monitoring of groundwater level. The last discussion is focuses on hydrogeochemical study and monitoring of water levels in the study area.

Chapter 5 consists of two parts. The first part discusses the results obtained from ANNs and ANFIS methods to the groundwater level for the selected well in the study area. The comparison and efficiency of these methods are also considered in this part. The second part presents and discusses the results of the ANNs and ANFIS methods of modelling of groundwater quality (total dissolved solids and electrical conductivity) in the study area.

Chapter 6 consists of two sections. The first section presents the conclusions of this study. The second section involves some recommendations for future work.

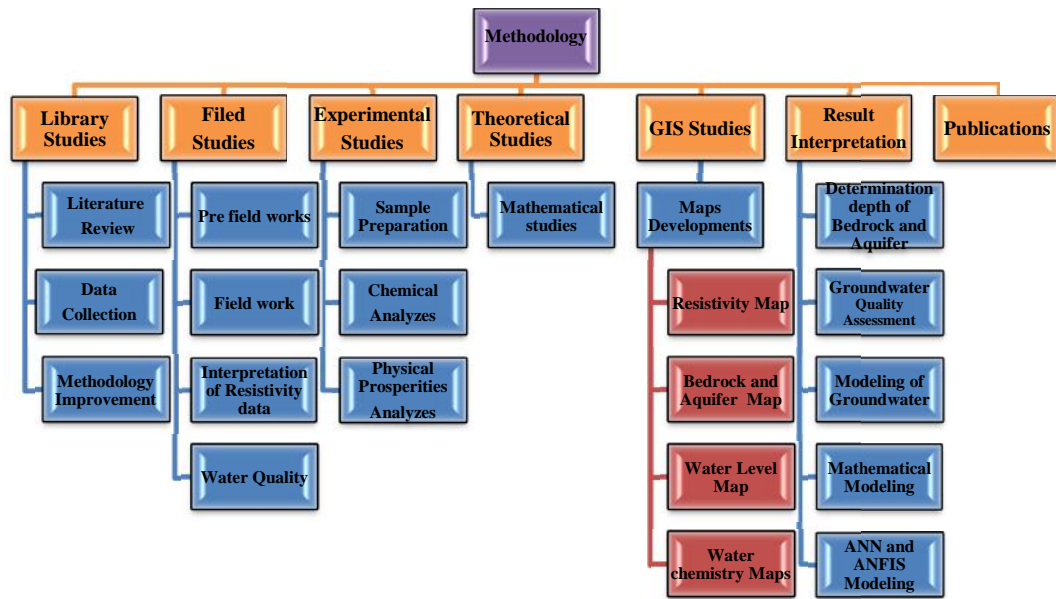


Figure 1.2: Research outline chart and procedure

CHAPTER 2: Study Area

2.1 Introduction

The goal of this chapter is to study the relevant fundamental geological composition of the study area. This section consists of a detailed discussion on geology, hydrology and hydrogeology of the study area. The geology, and hydrogeology as well as past climate states in the area are significant parameters in Hydrogeophysical investigation. Geological mapping provides a fundamental knowledge of controlling of the accessibility of groundwater resources used by the local geology. In addition, the geological information supplied an extra data set to help interpreting and interpolation of the resistivity data. The data was relieved to a greater knowledge of the groundwater investigation for Langat basin.

The research area found on Langat Basin is in the south-western part of Selangor state of Malaysia (Figure 2.1), between 2° 45' 22" N - 2° 54' 0.0" latitude and 101° 30' 11" to 101° 39' 42" longitude. The zone is surrounded by the hilly area to the north and east and the Straits of Malacca to the west. The digital topographic map and GIS media made the newest physiographic characteristics of research. The overall of study area is around 300 km² which consists of cleared land, rubber and oil palm plantations. Langat Basin is an important water catchment area, which provides a supply of raw water and other benefits for around one million people within the basin area. The Langat River Basin occupies the south and southeast parts of the State of Selangor. It is about 78 km long and varies from 20 km to 51.5 km wide. The origin of the Langat River is at the Pahang-Selangor border in which the height of hilly land reaches up to 1,500 m above mean sea level. It last drains into the Straits of Malacca on the Southwest of Selangor state (Figure. 2.1). This river gathers water in the mountainous areas and commonly

streams southwest ward in the mountainous terrain. It progressively changes its route towards the west after streaming in the hilly areas near Kajang and Bangi, respectively, and enters the flat land near Dengkil and flows westward along the Paya Indah Wetland and discharge in to the Straits of Malacca. The main tributaries of the Langat River are the Semenyih and Labu Rivers. The typical stream of the Langat River is in north and northeast towards the south and southwest in the eastern half of the basin and westward on the western part. Langat River water source is used for water supply besides various uses for example entertainment, fishing, effluent discharge, irrigation as well as sand mining. Different possible uses of this significant water source attract many industrial factories to invest in this area. The residual area consists of wetlands, rain forest and reed swamps.

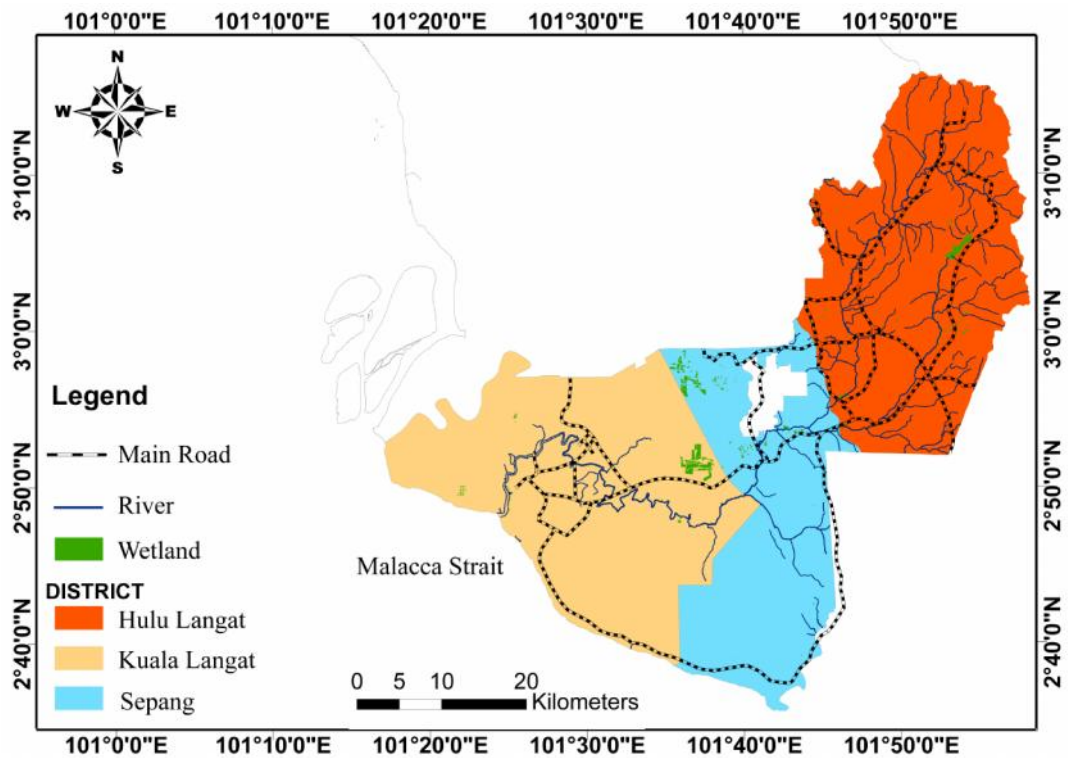


Figure 2.1: Geographical position of Langat Basin

2.2 Geomorphology

The geological condition and topographic formation control the geomorphology of drainage pattern in Langat Basin. Topographically, the Langat Basin consists of three zones; namely, the mountainous area, the hilly area and the lowlands from the upstream to the downstream. The groundwater recharging areas are in the upstream mountainous and hilly areas, and an aquifer spreads broadly in the flat lowlands. The surface feature of the Selangor State is displayed in Figure 2.2 and the study area is found in the southwestern area dominated by flat lowland.

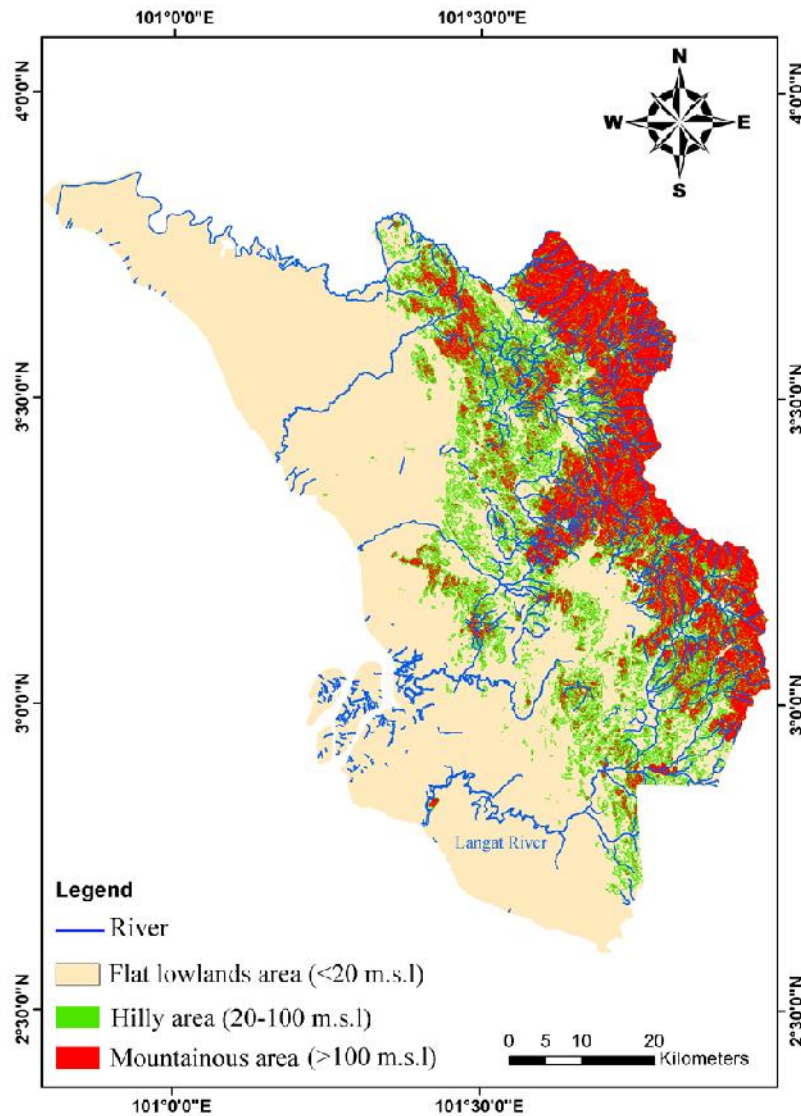


Figure 2.2: Surface features of Selangor State, Malaysia

2.2.1 Mountainous Area

The mountainous area is in the north-eastern part of the Basin. According to geographical information, the mountainous area with more than 20% slope is around 1362 km². The Langat River and its main tributary, the Semenyih River, start from the western slope of the mountain area infiltrating the Malay Peninsula. These rivers gather water within mountainous places and stream usually south-westward in the mountainous terrain. Riverbed ingredients are typically boulders of granite. The height of mountains at their riverhead area is roughly 1,000 m above sea level whereas almost of this area is below 500 m above sea level and ridges are steeper than those in the hilly area. The water from small channels collects from river and small waterfalls streaming from slopes towards the river.

2.2.2 Hilly area

Mild slopes spreading north to east in the centre part of the Langat Basin characterize the particular topography of the hilly area. Height of the hills is less than 100 m above sea level. The slope for hilly area was estimated to be 1324 km² and 5-20%, using geographical information. Langat River streams softly within the hilly area. The River sediments changes make up progressively from boulder and gravel within the mountainous area to sand and then silt in the hilly area.

2.2.3 Flat lowlands area

The lowland area with an alluvial plain is in the south-western part of the Langat Basin. The area surrounded by hill extends on the north and east, and through the Straits

of Malacca on the west. The Langat River direction changes to the west at the boundary of the hilly area and the lowland, and receives a tributary of the Semenyih River. After joining another branch, Labu River, the Langat River meandering begins. Close to the seacoast, the Langat River flows around a hill named Bt. Jugra, twist in a peaty swamp and after that flows into the sea. The riverbed sediments vary from silty at the Langat and Semenyih Rivers joint to clayey at Banting. Several small hills are available at the upper and eastern border of the lowland. The overall of the lowlands area and slope for this area was determined to be 5582 km² and less than 5%, using geographical information. The height of the lowlands is less than 20 m above sea level. The highest elevation is up to 15 m above sea level and the lowest elevation is 10 m in the study area.

2.3 Geology

The geological setting is one of the principal features of Langat Basin for contribution in sediments and growth of basin. Langat Basin is found in the geological Western Belt of Malaysia. The Western Belt significantly is different from the Central and the Eastern Belts at historical evolution, tectonic, structural and stratigraphy settings. The largely Lower Palaeozoic rocks extend southwards along the Western Belt into Selangor and the Federal Territory of Kuala Lumpur (Hutchison et al., 2009). Figure 2.3 shows the boundary between the Central margin and the Western Belt, which covers the entire state of Perak, Selangor, Malacca, the western and the central part of Negeri Sembilan (Hutchison et al., 2009). Figure 2.4 shows the geological map of Langat Basin gained from Department of Mineral and Geosciences of Malaysia (2007). Banting on the west and the Dengkil on the east side surround the research area which is in the flat lowland covered with alluvium sediment. The Quaternary alluvium extended

to the coastal plain. Quaternary sediments consist of unconsolidated gravel, sand, silt and clay of the Simpang formation in the Pleistocene, and Gula and Beruas formations in the Holocene form the geological condition of the lowlands. These formations deposit under fluvial or marine sediment conditions to fill waves of the bedrock of Kenny Hill Formation. The uppermost layer is the Beruas formation and includes, principally, the peat of Holocene fluvial-estuarine depositing with a maximum thickness of about 1–5 m. Soft clayey soils of light and greenish-gray to gray marine silty clay comprise the Gula Formation. The layer thickness increases from the northern part of the basin (several meters) towards the coast (more than 25 m). The Simpang formation is the thickest and oldest Quaternary sediment which has placed above the eroded surface of the older hard rock. The expected thickness is ranging from several meters in upperparts to over 100 meters at the coast. The monotonous sequence of interbedded shale, mudstones, and sandstones of the bedrock have found in the Kenny Hill Formation (bedrock). Figure 2.5 displays the location and topographic contour map of the study area.

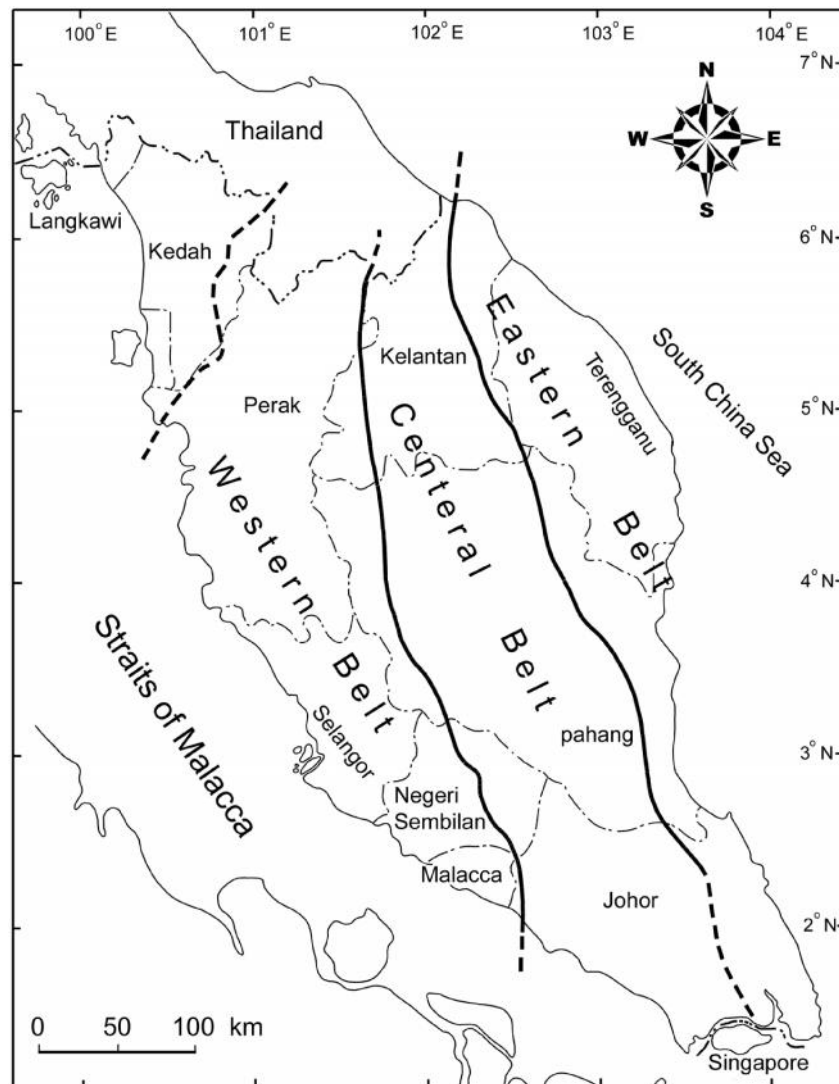


Figure 2.3: Map of Peninsular Malaysia showing the three belts (Hutchison et al., 2009)

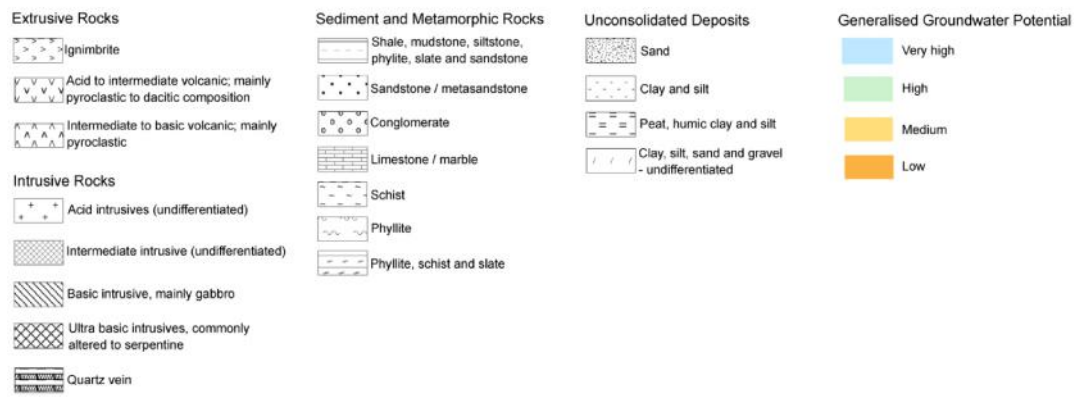
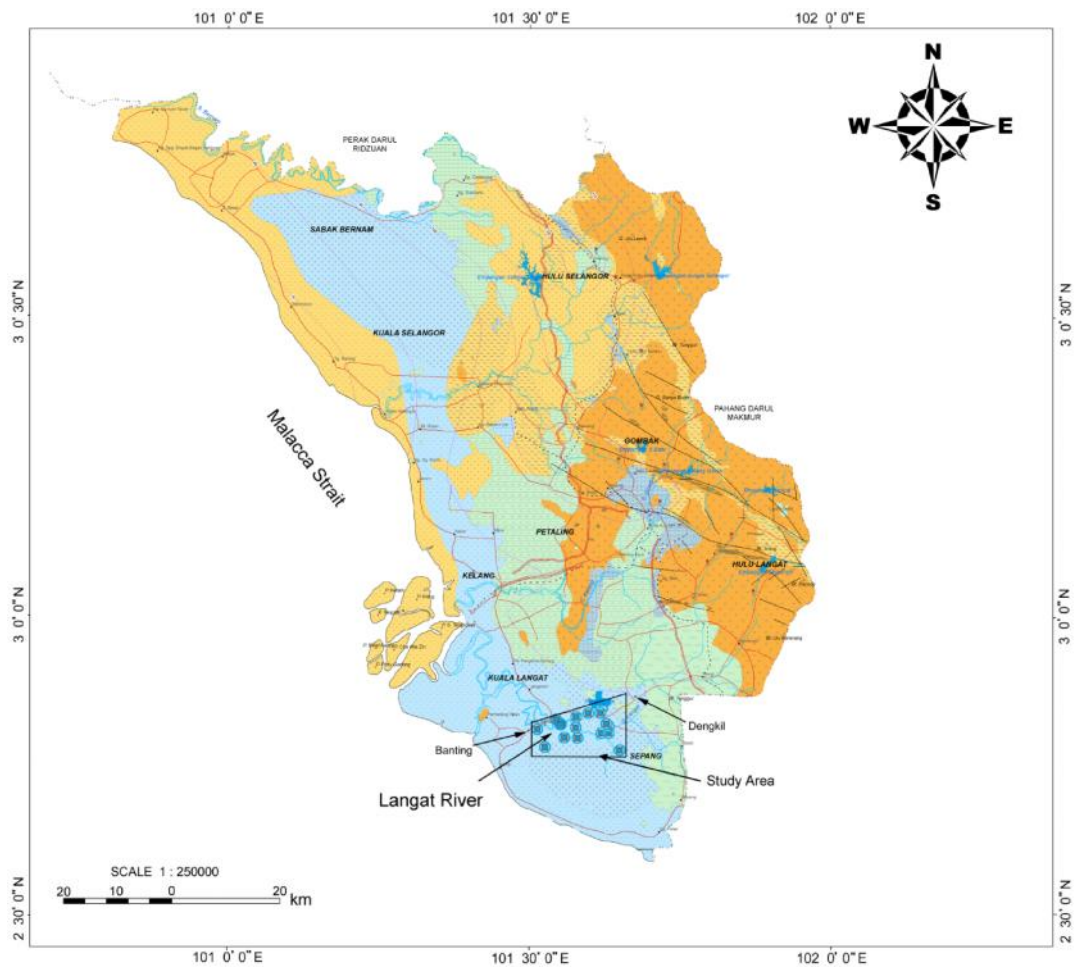


Figure 2.4: Geological and potential aquifer map of Selangor State Department of Mineral and Geosciences of Malaysia (2007)

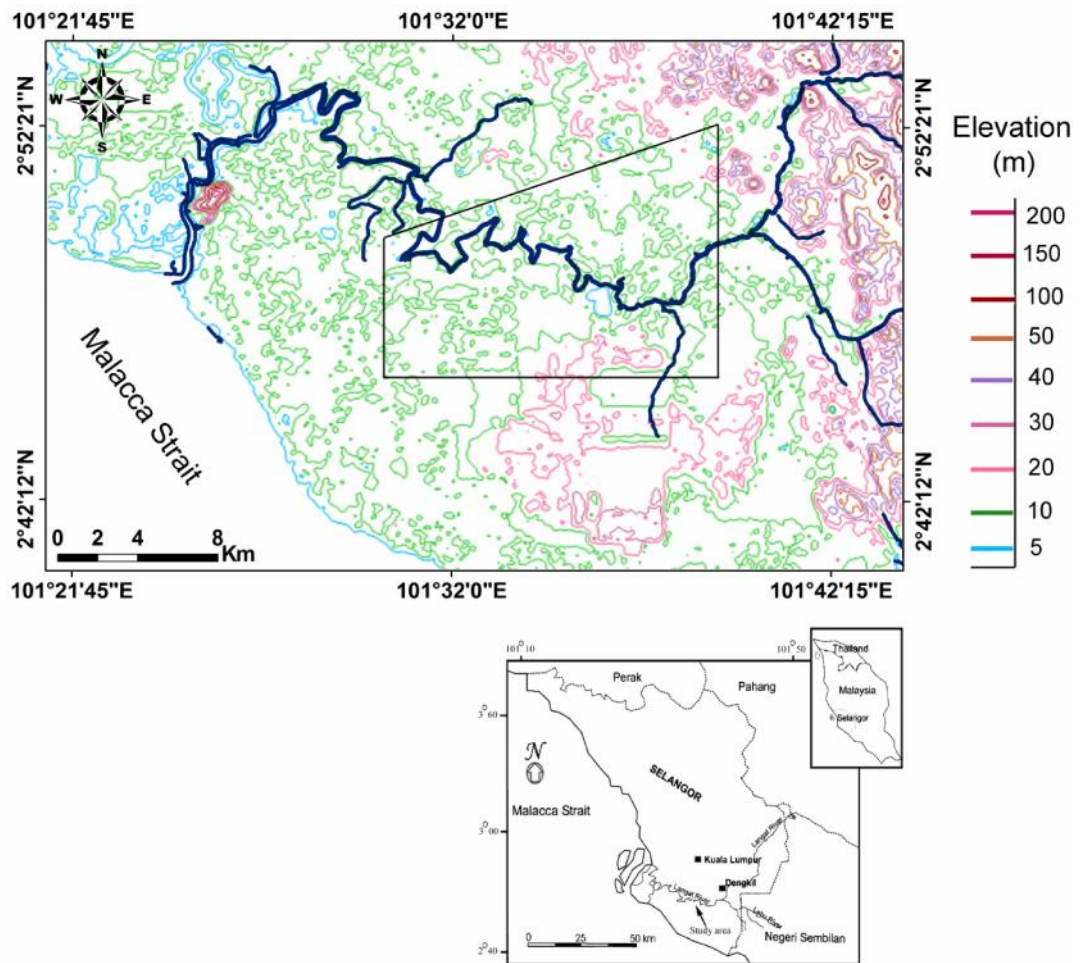


Figure 2.5: The location and topographical map of study area

2.4 Hydrology and Hydrogeology

Groundwater is one of the most important sources of water supply in the future. It is a valuable natural resource which without that there could be no life on Earth. This fact has led to a steady increase in demand for this finite resource, which has resulted in declining water levels, saltwater encroachment, dewatering, and land subsidence in some areas. Therefore, we need to improve the basic knowledge of groundwater. Hydrology has been developed as a science in reply to the desire to understand complicated problems of earth's water and help to clear up water these problems.

Hydrogeology is the study of geology which deals with the occurrence, movement and distribution of water in the rocks and soil of the Earth's crust (Hiscock, 2009)

Precipitation as a part of climatic factors as a mode of occurrence including temperature, humidity and wind affects the process of evaporation. The topography of the area will impact the precipitation process which will indirectly affect the excessive and low rates of runoff. Furthermore, the geology of an area will impact the topography of the area. Which three factors of climate, topography and geology of an area play essential role in the hydrological process.

2.4.1 Hydrology

One of the essential aspects of hydrology is the study of groundwater. Groundwater hydrology is a science of occurrence, distribution, and movement of water in the earth. It strongly depends on other natural sciences. Knowing rainfall and evaporation needs an understanding of climatology, which is a main influence on the hydrological conditions. The principal surface water body in the study area consists of Langat River and its tributaries. Shallow agricultural ditches form the drainage, which are discharged into Langat River or into its tributaries.

2.4.1.1 Climate

The climate of the Peninsular Malaysia is identified as four seasons namely, the northeast monsoon, southwest monsoons and two shorter periods of inter-monsoon seasons. The northeast monsoon occurs from early November to March and the southwest monsoon occurs from May to September (MMD, 2013).

During the two inter-monsoon seasons, the winds are variable. In the Langat Basin there is not any significant prevailing wind direction. The maximum wind speed in this area is varies from 5.5 to 7.9 m/s and this occurs less than 3% of the time.

2.4.1.2 Temperature

Temperature is not directly related to the water balance equation however it is an effective factor for evaporation. Mean annual temperature is roughly 27° C in the range of 24° to 32° C (MMD, 2013). The highest and lowest temperature reached during the noon and night with an average of 24° and 32° C, respectively.

2.4.1.3 Humidity

The average monthly relative humidity in the range of 77% to 85% varies from place to place of research area and from month to month. The minimum range of average relative humidity is varying from 67% in February to 79% in November. The maximum range of mean relative humidity is varying from 82% in June to 89% in November. In Peninsular Malaysia, the lowest relative humidity occurs in January and February while the highest relative humidity normally happens in November (MMD, 2013).

2.4.1.4 Rainfall

The precipitation data are helpful in showing the seasonal variation of precipitation on Langat Basin. Precipitation data were earned from the Telok Datok

station in the study area. Figure 2.6 shows the annual precipitation records from 2000 to 2013. The data show that the nature of precipitation gradually increased during the time. The range of annual precipitation is from 1585 and 2729 mm. Table 2.1 shows the basic statistical data for rainfall for the time period of 2000 to 2013. Figure 2.7 displays the minimum, maximum and mean monthly rainfall in the study area. The maximum rainfall occurs in September to December with a mean of 433 mm. The highest amount of rainfall (480 mm) occurred in December 2012. The minimum rainfall occurs in February with a mean of 81 mm. In addition the number of rainy days in a year ranges from 140 to 210 days.

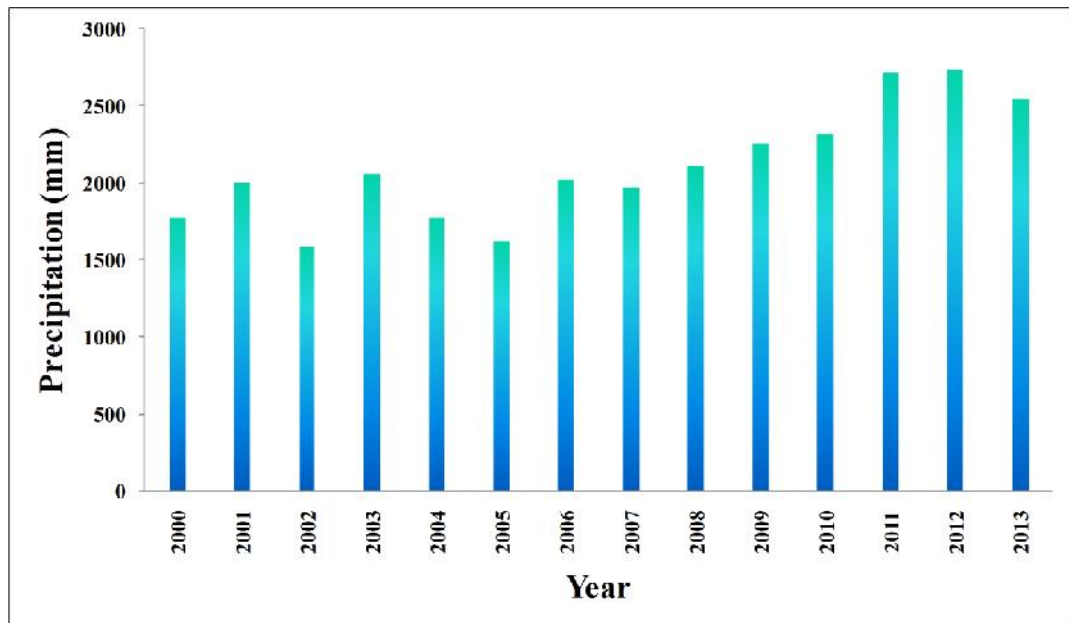


Figure 2.6: Annual precipitation of Telok Datok station 2000-2013

Table 2.1 Statistical data for annual precipitation, Langat Basin

| | |
|--------------------------|----------------|
| Average Rainfall | 2107 mm |
| Median Rainfall | 2041 mm |
| Standard Deviation | 370 mm |
| Range in Annual Rainfall | 1585 - 2729 mm |

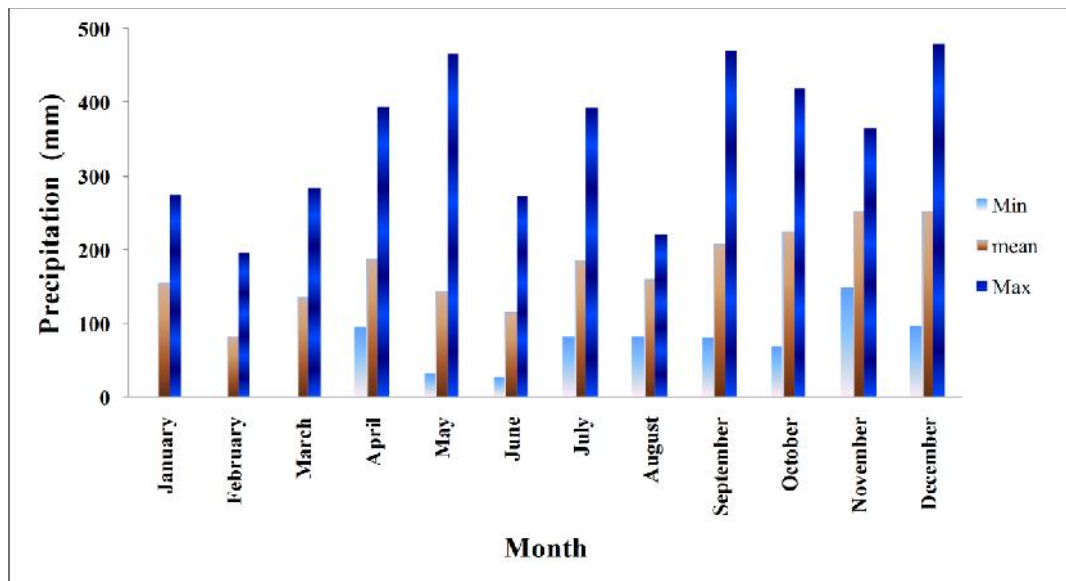


Figure 2.7: Average of monthly rainfall between 2000 to 2013 in Telok Datok station

2.4.1.5 Water Balance

The interpolation of discharge, the examination of evapotranspiration and the annual water balance from 1980 to 1999 in the study area were estimated based on the JICA (2002) report on the previous rainfall analyses. According to this report the annual water balance in the study area is as summarised below:

- 1) Annual average groundwater recharge in the upper + Semenyih + middle Langat River Basin is estimated at 108 mm/year. This is equivalent to $139 \times 10^6 \text{ m}^3/\text{year}$ and is 4.8% the rainfall in the area.
- 2) The time series of annual groundwater recharges change irregularly in the upper + Semenyih + middle Langat River Basin. The annual groundwater recharge in 1988-1990, 1992 and 1997 fell below zero in sub-basins. Groundwater recharge seems to draw down after drought years of one or two-year period. This may suggest that the effect of surface water on the groundwater appears approximately two years later.
- 3) The runoff coefficient for the upper + Semenyih + middle Langat River (upper reach from Dengkil Town) is roughly estimated at 37.8%.

2.4.2 Hydrogeology

The purpose of most hydrogeological studies is to find potential sites for development of enough quantity of reasonably good quality groundwater for particular uses: domestic, irrigation or industrial etc.

2.4.2.1 Aquifer

An aquifer is a body involving geological material which provides valuable volumes of groundwater to natural springs and water boreholes. The aquifer in the lowlands of the Langat Basin (study area) is alluvial sediments consisting of sand and gravel by study of the information and exploration wells the Mineral and Geosciences Department of Malaysia (JICA, 2002). Based on drilling data the gravelly sand (Simpang formation) aquifer layer are covered by clayey layers and then, at some

locations, by peat layers, which affect the aquifer depending on the location and make it confined. This Simpang Formation is the thickest and oldest Quaternary sediment in the Langat Basin. The depth of aquifer is less than 15 m at the north-western of the study area to more than 25 m at the west of the area. The thickness of this formation varies significantly because of the undulating nature of the top of the formation and uneven bedrock surface, on which the formation is in direct contact. The expected thickness is from several meters in upperparts of the area to over 45 meters at the west of the area.

2.4.2.2 Water recharge

To recharge is to refill an aquifer's groundwater. Estimation of recharge is of great significance for evaluating groundwater of an area. In tropical areas, rainfall is the major source of groundwater recharge. The rainfall recharges groundwater in the study area by show downward flow through which aquitard, infiltration of water around the edges of bedrock outcrop where the aquitard is thin or absent, Flowing from relatively more permeable bedrock, infiltration from, riverbed occurs in stretches where the river bottom is in contact with more permeable sandy horizons infiltration from ponds and wetland areas where the upper aquifer and the aquitard have been removed and replaced by more permeable materials (JICA, 2002).

2.4.2.3 Water Quality

The quality of water is as significant as its accessible quantity. The issue of groundwater contamination is highly challenging for growing countries because of the inadequate appropriate sanitary qualifications and piped water supply. Currently, this issue is increasing solemn since the last few decades due to rapid industrialization,

urbanization and agricultural activities that are leading to acute deterioration in the quality of both surface water and groundwater (Singhal and Gupta, 2010). Recently in Peninsular Malaysia the accessible water resources have been under pressure because of growing demand by industry and domestic water consumption, and leads to the destruction in water quality due to contamination. Groundwater quality has been developed into a major problem as the water demand increases in the Langat Basin. Agricultural development is known as the cause of groundwater contamination in the study area. These are essentially the most critical indicators that lead to contamination of groundwater because of higher using inorganic fertilizers and pesticides for gaining greater harvest produces. Other activities such as mining are another main factor in increasing groundwater contamination sources in the Langat Basin. In the study area, the disposal water infiltrates in the groundwater system by the drainage seepage or directly by pore space of top soil within the surrounding disposal water sources. The most dependable water quality survey was published by Malaysian-Japanese committee to environmental management project (JMG, 2002). These water quality analyses revealed that the area of water especially at the west of the study area is degraded.

The change of groundwater quality due to saline intrusion from the west of the study area near the coastal plain which is steadily increased is recognized as the main problem in the area. The main natural contamination is saltwater or brackish water. Pumping well stations at the west of the study area close to the coastal plain, around 10 km from the sea, are facing the danger of saltwater intrusion. In general, the excessive chloride values of groundwater take place because of seawater intrusion. It is assumed that the seawater flows into the freshwater aquifer near the coastal areas. Saline intrusion will be the biggest impact in this area. Moreover, brackish water is detected in the Langat River. Banting is the point transition between brackish and fresh water according to Nasiman and Nazan (1997).

National Water Quality Standards for Malaysia (NWQS) (DOE, 2006) were utilized to consider Langat Basin quality. Generally speaking, Langat Basin water quality as a result of environmental management project is classified as IV and V (exceeds standards level) that is suitable for irrigation only and requires extensive treatment for drinking.

2.4.2.4 Water Demand

Robins et al. (Robins et al., 1999) describe water management as it is a need to manage the whole quantity of water gain from environment employing measures to control both waste and consumption. Nowadays, growing demand of industry, domestic water consumption and destruction in water quality by pollution have resulted in putting pressure on accessible water resources. Increasing water supply due to decrease in water demand, thus in many respects, water demand is simply a subset of water supply. The population have effects on water supply as well as demand. The economy in Malaysia has turned from agricultural to an industrial which is important to take in account. This momentum of development within growth and economy will need added water supply. To reduce that status, numerous regions are starting to rely on organizing, development and management for the total water resources of Peninsular Malaysia.

Increasing in agricultural activities, domestics and industries cause to an increase in water use tendency in the Langat Basin. However, this region becomes significant because of being the fastest developing economic region in Malaysia. It is necessary to study groundwater resources and develop sustainably within water stressed and isolated regions. Most of the groundwater abstracted in the Langat Basin is for industrial use with the biggest users being Mega Steel Company. The natural groundwater flow is affected by abstraction of water from Mega Steel factory. Mean 23,300 m³/day being

abstracted from some wells at the Mega Steel Company as it is shown in past groundwater pumping record (Minerals and Geosciences Department, 2002).

Water demand management searches for conserving water and optimizes the usage of it and therefore restricts the need for new supplies. Computer processing is capable to growth the hydrological models. It is an essential tool for planning, design, and managing the hydrological associated infrastructure. Through the years, hydrological models can be employed as a tool to modify a system that needed the decision making and policy evaluation.

Chapter 3: Research Methodology

3.1 Introduction

Materials and techniques are the essential section of each study that helps to find the results during research procedures. The main steps of research methodology are literature review, field surveying, experimental analyses, modelling, theoretical study, and interpretation. The detailed approaches on their procedures and their output have comprehensively displayed in this chapter. The chosen approaches were surely assumed to be suitable methods to contribute in knowledge. The techniques introduced here are fundamentally in order and the data are interpreted in chapters 4 and 5.

An integration of electrical resistivity, hydrogeochemical and artificial intelligence techniques were used for studying the groundwater in the selected area. Electrical Resistivity Imaging (ERI) technique is one of the geophysical techniques. It has been chosen for this study due to its ability to image the subsurface structure, bedrock depth and layer thickness. Borehole is as a tool for investigation on subsurface geology which is needed to find geological correlation between the subsurface characteristic changes and geoelectrical resistivity. Besides, the gained data from Borehole are utilized for calibration of the electrical resistivity data to the subsurface structure. The hydrogeochemical approach is used to assess the chemical content of groundwater. The artificial intelligence method is used to model groundwater level and quality. The use of all methods is ideal for investigation of groundwater problem.

3.2 Geoelectrical Resistivity Method

Earth resistivity is special interest in hydrogeological investigations. The most favourable geophysical techniques for groundwater investigation are geoelectrical methods and particularly resistivity surveys because they are comparatively cost-effective (Kirsch, 2007) as they provide a robust response to subsurface structure. The electrical resistivity technique consists of the apparent resistivity of soils and rock measurements as a function of different geological variables for example the mineral and fluid content, permeability, porosity and degree of water saturation in the rock. Employing the resistivity technique for this research was because of its ability to image the subsurface structure, water content, bedrock depth and layer thickness. This method is also popular to apply in complex geological and noisy regions when other geophysical approaches for example seismic and GPR methods are not beneficial (Cosenza et al., 2006; Heincke et al., 2010; Perrone et al., 2004). Therefore, the resistivity method seems to be suitable to achieve the objectives outlined in this research.

This section presents the principles of Electrical Resistivity Imaging (ERI) method and conciseness of the equipment used in the fieldwork. Furthermore, the explanation on the programs employed to interpret and supply the results will also be presented.

3.2.1 Basic Principles

Georg Simon Ohm in 1827 determined that the Electric Current in a conducting wire is proportional to the potential difference across it, as shown in Figure 3.1. The linear relationship is expressed in equation

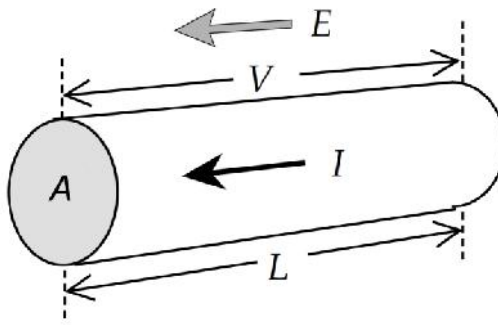


Figure 3.1: Parameters used to define Ohm’s law for a straight conductor with length (L) and cross section (A)

$$V = IR \tag{Equation (3.1)}$$

Where V is the potential difference, I is the electrical current, and R is the resistance. The resistance is proportional to the length L and inversely proportional to the cross-sectional area A of the conductor for a given material (Figure 3.1). These relevance are indicated through following equation

$$R = \rho \frac{L}{A} \tag{Equation (3.2)}$$

The proportionality constant ρ is the resistivity of the conductor. Resistivity is a physical property of the material of the conductor and shows its ability to oppose a flow of charge. Conductivity of the material (σ) is the inverse of ρ .

The basic physical law utilized in geoelectrical resistivity survey is called Ohm’s Law which is expressed by equation (3.3) for current flow in a continuous medium (Telford and Sheriff, 1990)

$$J = \sigma E \quad , \quad \sigma = \frac{1}{\rho} \quad \rightarrow \quad E = \rho J \tag{Equation (3.3)}$$

Where E is the electric field intensity and J is the current density. In practice the electric field potential is measured. Equation (3.4) gives the relevance between the electric potential and the field intensity

$$E = -\nabla V \quad \text{Equation (3.4)}$$

The result of combining the equation (3.3) and (3.4) is

$$J = -\sigma \nabla V \quad \text{Equation (3.5)}$$

The current sources have the type of point sources nearly in all researches. Then over an elemental volume ΔV surrounding the current source I which is placed at (x_s, y_s, z_s) , the connection between the current (Dey and Morrison, 1979) is written by

$$\nabla \cdot J = \left(\frac{I}{\Delta V} \right) \delta(x - x_s) \delta(y - y_s) \delta(z - z_s) \quad \text{Equation (3.6)}$$

Where δ is the Dirac delta function. Combination of the equation (3.5) and (3.6) gives:

$$-\nabla \cdot [\sigma(x, y, z) \nabla \phi(x, y, z)] = \left(\frac{I}{\Delta V} \right) \delta(x - x_s) \delta(y - y_s) \delta(z - z_s) \quad \text{Equation (3.7)}$$

This fundamental equation expresses the potential distribution in the ground because of a point current source. To solve this equation many methods have been employed (Loke, 2013). This is the forward modelling problem for example for identification of the potential which was observed over a given structure in the subsurface. In general, the linear filter approach is applied to the one dimensional case,

because the subsurface is limited to a number of horizontal layers (Koefoed and Patra, 1979). The finite difference and finite element methods are applied for the 2D and 3D cases. A homogeneous subsurface and a single point current source on the ground surface is the simplest case as it is shown in Figure 3.2. The current flows radially away from the point source and with distance from the current source the potential varies inversely. The current flows perpendicular to the equipotential surfaces. The potential, V , at a distance, r , from a point current source, I , on the surface of the earth of resistivity, ρ , is expressed by following equation

$$V = \frac{\rho I}{2\pi r} \quad \text{Equation (3.8)}$$

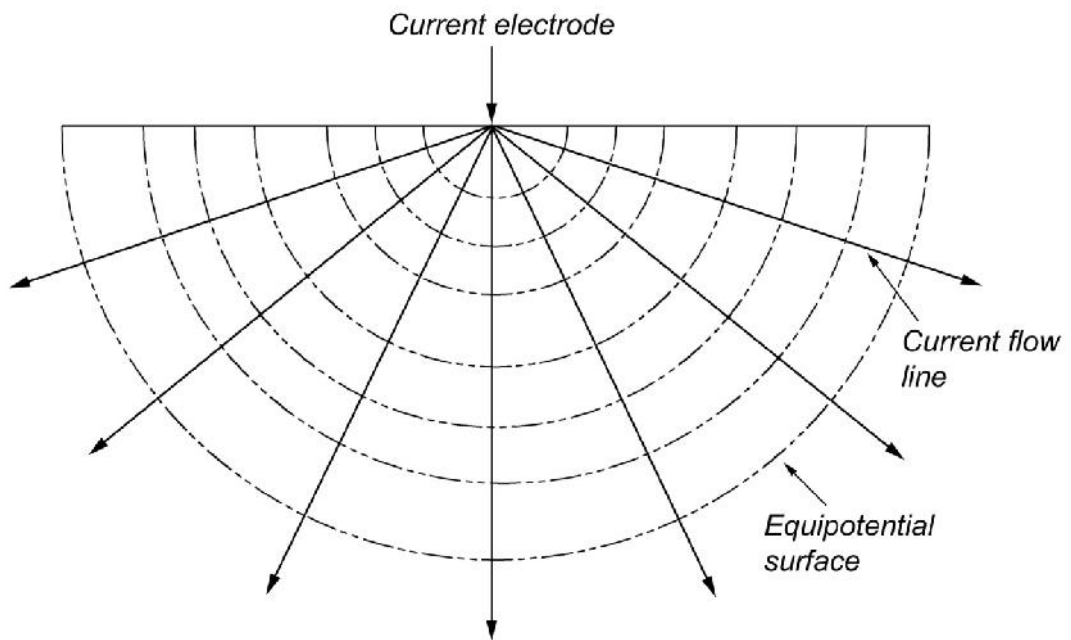


Figure 3.2: The flow of current from a point current source and the potential distribution

The potential difference between two points is measured within almost all electrical resistivity surveys. Figure 3.3 shows a typical arrangement with four electrodes which in this case a four-electrode array including of two current electrodes. Normally it is the difference in potential, ΔV , between two potential electrodes at points on the surface, which is measured using electric voltmeter and this is referred to as the voltage.

The two current electrodes (C_1, C_2) introduce a current $\pm I$ and the potential difference ΔV between the potential electrodes P_1 and P_2 can be given as

$$\Delta V = V_{P_1} - V_{P_2} \quad \text{Equation (3.9)}$$

Where V_{P_1} is the potential at electrode P_1 . The potential at P_1 is the sum of potentials because of current sources and sinks. In the case of being only a single current source and a single current sink, the potential difference is:

$$\Delta V = (V_{P_1C_1} - V_{P_1C_2}) - (V_{P_2C_1} - V_{P_2C_2}) \quad \text{Equation (3.10)}$$

Where $V_{P_1C_1}$ is the potential at P_1 due to current entering at C_1 . As C_2 is a current sink where current leaves the earth, $V_{P_1C_2}$ and $V_{P_2C_2}$ are negative. As a result of combining the equation (3.8) and (3.10) is

$$\Delta V = \frac{\rho I}{2\pi} \left(\frac{1}{r_1} - \frac{1}{r_2} - \frac{1}{r_3} + \frac{1}{r_4} \right) \quad \text{Equation (3.11)}$$

Where $r_1 = C_1P_1$, $r_2 = C_1P_2$, $r_3 = C_2P_1$ and $r_4 = C_2P_2$

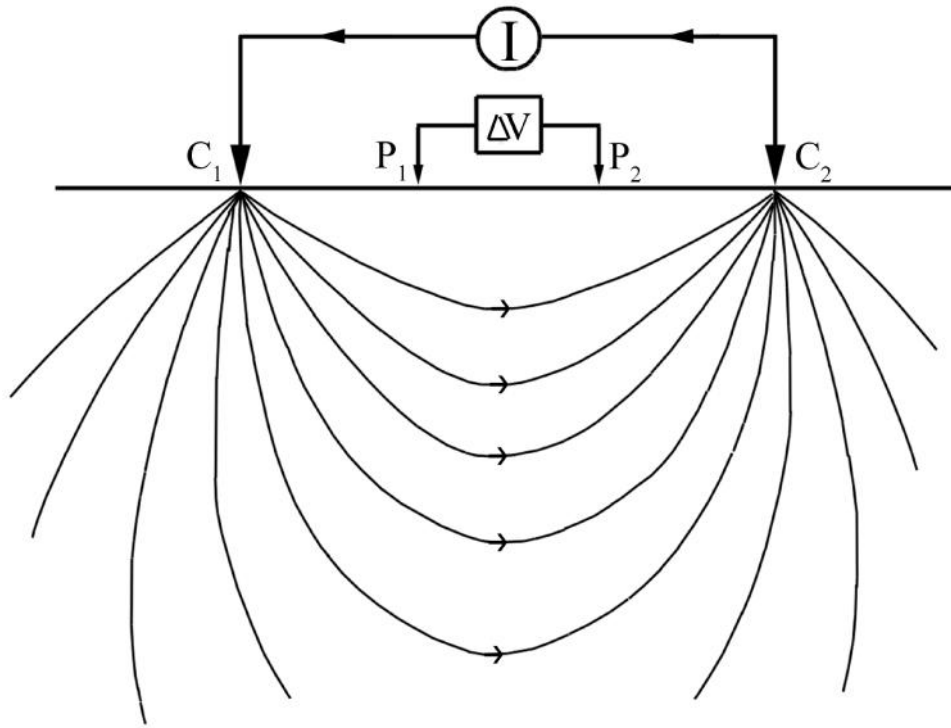


Figure 3.3: A generalized four-electrode array

3.2.2 Geometric Factor and Apparent Resistivity

Equation (3.11) gives a resistivity ($\rho = \frac{1}{\sigma}$) and is equal to the bulk resistivity of the homogenous half space and as a result it is constant for any electrode geometry and injection current

$$\rho = K \times \frac{\Delta V}{I} \quad \text{Equation (3.12)}$$

Where the parameter K indicates a geometric factor

$$K = \frac{2\pi}{\frac{1}{C_1P_1} + \frac{1}{C_1P_2} + \frac{1}{C_2P_1} + \frac{1}{C_2P_2}} \quad \text{Equation (3.13)}$$

The arrangement of four electrodes C_1 , C_2 , P_1 and P_2 affect on the geometric factor K. Table 3.1 demonstrates the geometric factor for the common electrodes array.

Table 3.1: Geometric factor for some common configurations

| Array | Geometric factor (K) |
|---------------|-------------------------------------|
| Wenner | $\frac{2\pi a^2}{\pi a}$ |
| Dipole-dipole | $\frac{2\pi a^2}{\pi n(n+1)(n+2)a}$ |
| Schlumberger | $\frac{(n+1)(n+2)a}{\pi n(n+1)a}$ |
| Pole-Dipole | $\frac{2\pi n(n+1)a}{\pi n(n+1)a}$ |

In general case of subsurface electrodes, the position of the mirror current electrodes C'_1 and C'_2 should be incorporated to compute the geometric factor, yielding.

$$K = \frac{2\pi}{\frac{1}{C_1P_1} + \frac{1}{C'_1P_1} + \frac{1}{C_1P_2} + \frac{1}{C'_1P_2} + \frac{1}{C_2P_1} + \frac{1}{C'_2P_1} + \frac{1}{C_2P_2} + \frac{1}{C'_2P_2}} \quad \text{Equation (3.14)}$$

In case of inhomogeneous earth, various values of ρ are measured when the geometry of the electrode spread is changed or the array is moved. Consequently, the measured quantity by equation (3.12), ρ , is referred to as the apparent resistivity ρ_a which indicate that it reflects the a homogenous half space properties that may not exist in practice (Kuras, 2002). Equation (3.13) and (3.14) can be valid only for a ground with a flat surface. The geometric factor is unknown in case of variable topography, and it can only be evaluated by numerical modelling employing a homogenous resistivity (Günther, 2004).

3.2.3 Relationship between Resistivity and Geology

In general, the variations within the resistivity distribution of the subsurface have

the function of lithology. The electrical properties of minerals and rocks have been found to be an interest of research for around two centuries. In the case of most known variation of minerals, rocks, soil and unconsolidated sediment, and for groundwater, the laboratory measurements of resistivity have been done. Moreover, many field measurements have been carried out with these substances in their natural setting (Robinson, 1988). The resistivity of some rocks, soil materials and chemicals has been illustrated in Figure 3.4 (Loke, 2013).

Typically, igneous and metamorphic rocks have high resistivity values which are determined by the degree of fracturing, and moisture content in the fractures and may vary from about 1,000 to 10 million Ohm.m, depending on saturation level. Lower resistivity values are corresponding to sedimentary rocks which are typically more porous and permeable than igneous and metamorphic rocks and typical resistivity values for these rocks vary from 10 to around 10,000 Ohm.m, with most values below 1,000 Ohm.m. Unconsolidated sediments often have pretty low resistivity varying from about 10 to less than 1,000 Ohm.m, which confirms, the dependency of resistivity value on the porosity and water saturation, and the clay content. There are lower resistivity values for clayey soils which they typically have lower resistivity than sandy soil. Resistivity values for various rock overlap due to variations in some factors such as the porosity, the degree of water saturation and the concentration of dissolved salts which significantly affect on the resistivity. Groundwater resistivity values range from 10 to 100 Ohm.m depending on the concentration of dissolved salts.

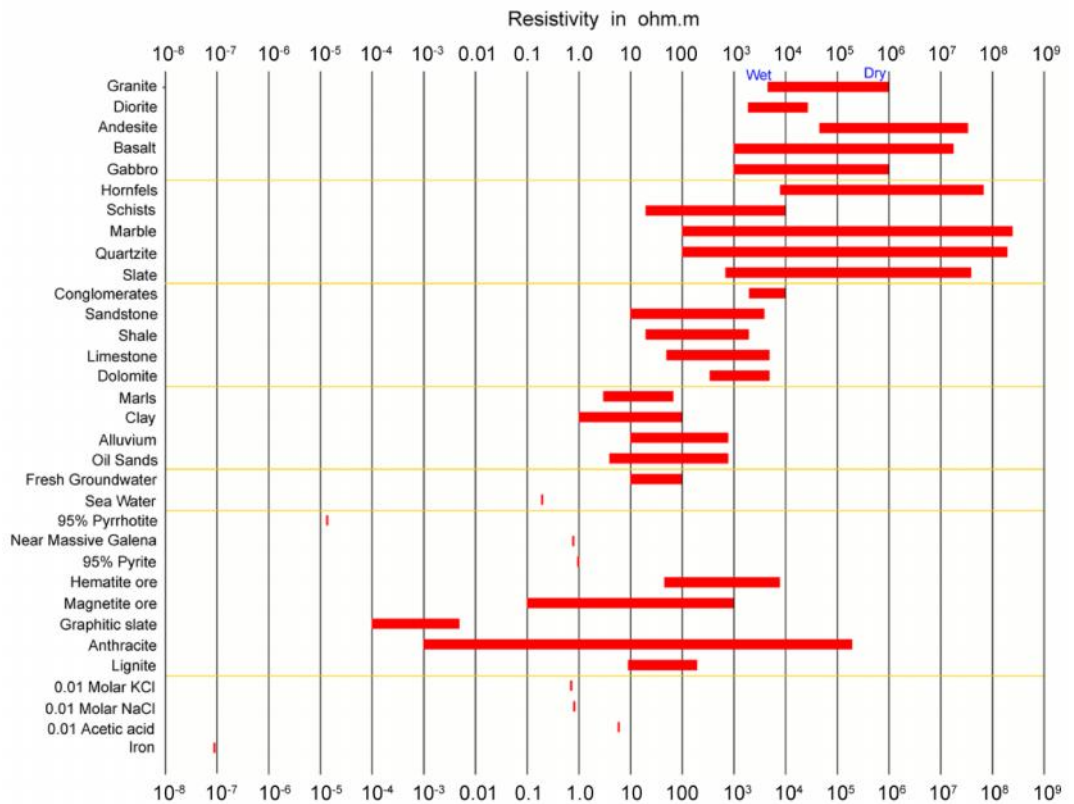


Figure 3.4: Resistivity of common rocks, soil materials and chemicals (Loke, 2013)

3.2.4 Basic Inversion Theory

An explanation of 2D inversion of electrical resistivity data employed in this thesis in RES2DINV inversion software is mentioned in following section. Typically a set of resistance measurements decreased to apparent resistivity values which are resulted from resistivity field data collection. The sensitivity of data is good due to a suitable model. Outlying data points can be "systematic" and which "random" noise. Measurement failures in the field cause systematic noise such as breaks in the cable, poor ground contact and so on. Systematic noise is usually easy to detect. Random noise consists of effects such as telluric currents which have influences on all the readings. The issue of non-uniqueness of a solution established fact in the resistivity sounding inversion and for the identical measured data set, there could be a broad range of models

for the same data set of computed apparent resistivity values. Using some presumptions during a survey can incorporate into the inversion subroutine that causes to narrow down the range of feasible models (Loke, 2013). To determine a model for the subsurface whose response is the same to the measured data, subjecting to certain limitations and within suitable limits is the idea of geophysical inversion. The model is an idealized statistical delegation and has a set of model parameters that are the physical quantities had to be approximated the observed data (Loke, 2013). The RES2DINV software employs the cell-based approach, in which the model parameters are the resistivity values of the model cells, while the data is the calculated apparent resistivity values (Loke, 2013). For the 2D and 3D resistivity models, the finite-difference (Dey and Morrison, 1979) or finite-element approaches (Silvester and Ferrari, 1996) supply the numerical link between the model parameters and the model reaction.

3.2.5 Electrode Array

The apparent resistivity value depends on the geometry of the employed electrode array (Reynolds, 2011), as described through the geometric factor K . 2D imaging of the subsurface includes various types of electrode configuration and Wenner (Figure 3.5(a)), Schlumberger (Figure 3.5(b)), Dipole-Dipole (Figure 3.5(c)) and Pole-Dipole (Figure 3.5(d)) are the most typical arrays in resistivity imaging. Table 3.1 shows the geometric factors for these arrays. In a field survey some parameters such as the subsurface structure to be imaged, the sensitivity of the resistivity meter and the background noise level have effects on the selection of the extremely suitable array. The characteristics of an array consist of the depth of investigation, the sensitivity of the array to vertical and horizontal variation in the subsurface resistivity, the coverage of horizontal data and the signal strength (Loke, 2013).

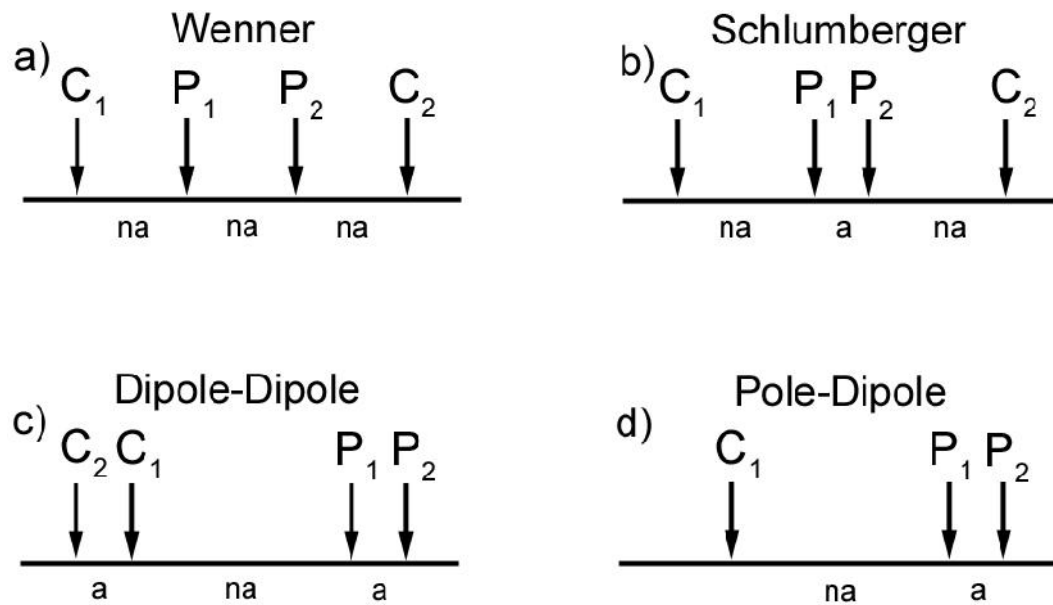


Figure 3.5: Common electrode arrays used in geoelectrical resistivity surveys

In this study, Terrameter SAS4000 manufactured by ABEM with Wenner array is used for the all geoelectrical resistivity data acquisition. In the case of surveys performed in a noisy area, the Wenner array is useful that is because of excessive signal strength and its ability to give dense near-surface cover of resistivity data. On the other hand, the outcomes achieved from this array provide a great resolution for horizontal structure with vertical variation of resistivity and can supply a clear image about the subsurface structure. The geometric factor for the Wenner array is smaller than the geometric factor for other array (Telford and Sheriff, 1990).

3.2.6 Resistivity surveying equipment

The resistance of the subsurface structure was measured by resistivity survey instruments. Surface resistivity conditions and depth of investigation are significant in order to choose a suitable instrumentation. Most resistivity meters measure a low-frequency alternating current rather than direct current which is due to two main

reasons. Firstly, applying direct current result in build-up of anions around the negative electrode and cations around the positive electrode and that means electrolytic polarization occurs, and this inhibits the arrival of further ions at the electrodes. Secondly, applying alternating current overcomes the effects of natural electric currents in the ground (telluric currents) which are flow parallel to the Earth's surface and cause regional potential gradients (Kearey et al., 2009). For 2D resistivity measurements, the basic equipment includes a transmitter, receiver, power supply, switching unit, electrodes and cables.

The 2D geoelectrical resistivity imaging surveys carried out at the proposed sites employing the ABEM Terrameter SAS4000. This system is automatic electric imaging and it uses for induces polarization and resistivity imaging. The Terrameter was connected together with an automatic selector system and multicore cable to the electrodes which were connected to takeouts with equal intervals. Figure 3.6 shows Electrode Selector ES10-64 employ with Terrameter SAS4000 and multi-conductor cables in resistivity imaging. The Wenner array with spread length (with maximum spread length of 400 m) of surveys line which depends on the target and space available in the field was employed to data acquisition. The spread was focused to find deeper targets for instance basements and deep aquifers. The data cover the resistivity surveying. Figure 3.7 (a) shows the equipment set up with the four wheel cables. 61 electrodes were required with 'a' meter electrode spacing in this setting. For this set up the data acquisition is WEENER_S and WENNER_L protocols. Figure 3.7 (b) displays measurement employing electrode combinations that provided standard separations of 'a' spacing for shallowest target and then 2a, 3a, etc to find out more concerning and deeper target. Figure 3.7 (c) shows the data coverage for four wheel cables which employs ABEM Terrameter SAS4000.



Figure: 3.6: ABEM Terrameter SAS 4000 with its accessories

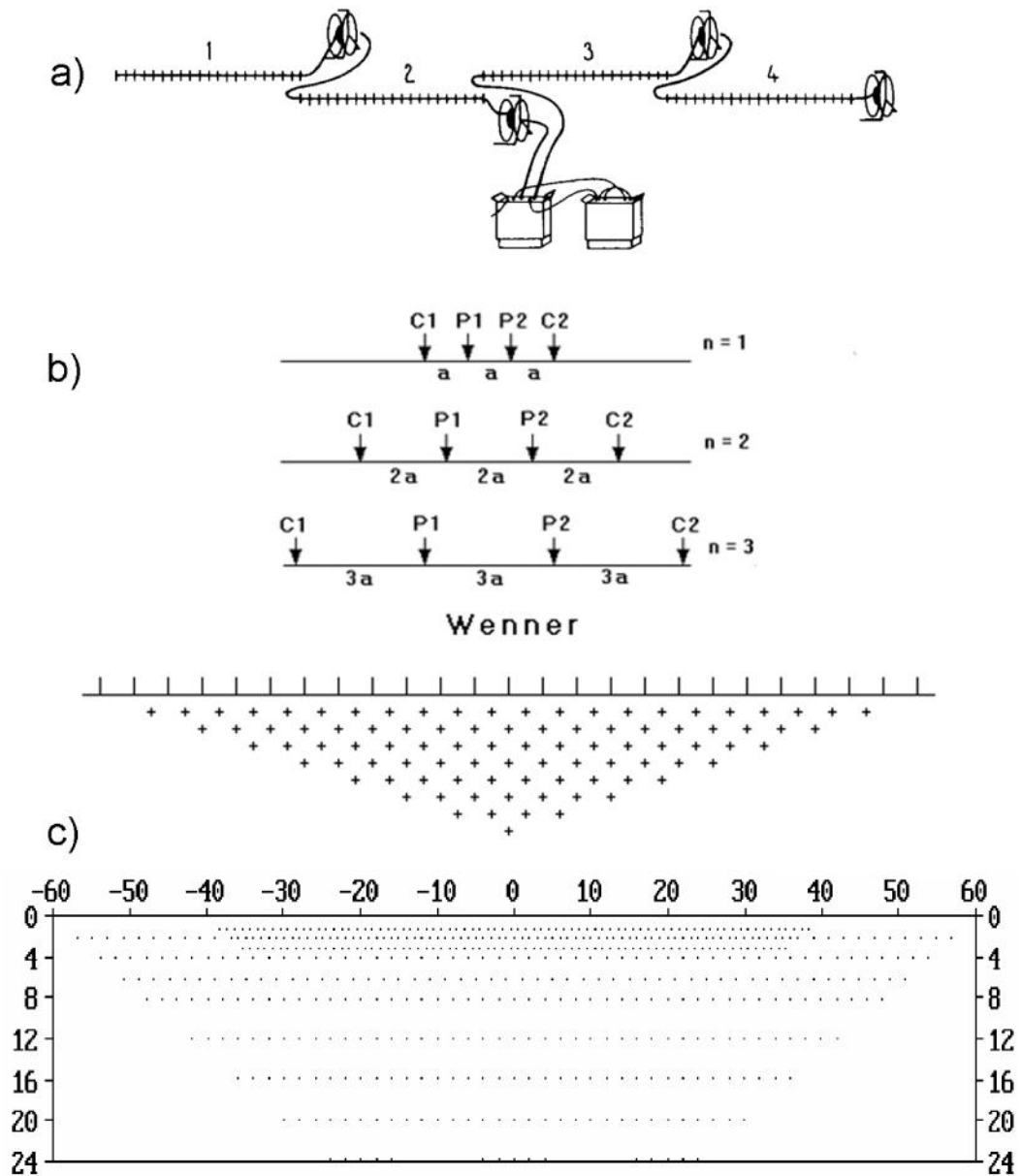


Figure 3.7: (a) Equipment set up for four wheel cables; (b) A typical field arrangement for 2D electrical imaging surveys; and (c) Data cover of standard Wenner using WENNER_L and WENNER_S for roll-along with three stations (ABEM, 2010).

3.2.7 Two Dimensional Resistivity Modelling

The modelling illustrates how one of the array parameters, the electrode spacing has influenced on the resolution and the depth of investigation. In two-dimensional surveys, multi-electrode arrays supply a 2D cross-section of the subsurface structure. A

forward modelling subroutine is planned for computation of the apparent resistivity over an underground structure with particular parameters that can be calculated by a real field data. The main techniques to compute the apparent resistivity values for a determined model consist of analytical approaches, boundary element approaches and the finite difference and finite-element approaches (Loke, 2013). Analytical techniques are the most precise, nevertheless limited to basic models for instance a sphere or cylinder. Boundary element approaches are more flexible; however the small number of area with various resistivity values restricts its application for resistivity inversions. Therefore, the only sensible solution are generally the finite difference and finite-element techniques where in environmental and engineering research the underground can accompany with an arbitrary resistivity distribution. In this study, the applied inversion approach is the finite-difference technique that is based on a method interpreted by Dey and Morrison (Dey and Morrison, 1979) with a modification by Loke (2013). The inverse simulation was carried out in the RES2DINV running and interpretation computer software.

3.2.8 Data Acquisition

Multiple sets of coupled four-electrode configurations which are automated electrical resistivity systems are extended for resistivity surveys. One couple of electrodes employed to induce the current into the ground, at any instant in time and to measure potential differences; a second couple of electrodes are employed. For automatic measurements, the multi- electrode resistivity meters are provided across the traverses which its length can be around hundred feet. The electrode space depends on the target feature. Four wheel cables are required for deep target when the maximum

electrode spacing is around 5 m. When the target requires higher resolution (water level), electrode spacing is to be smaller (1 m).

3.2.9 Depth of Investigation

The depth of investigation is determined by electrode spacing and array assesses the resistivity of the deeper target with increasing the electrode space. The sensitivity function (Edwards, 1977) of the array is employed to obtain a statistical value for the depth of investigation (Loke, 2013). In resistivity sounding surveys, when the resistivity of a thin horizontal layer is changed, a variation in the potential can be estimated. The layer extends from - to + in direction of x and y, for a horizontal layer. Consequently, the integration of 3D Frechet derivative (McGillivray and Oldenburg, 1990) results in the sensitivity function in the x and y direction, i.e

$$F_{1D}(z) = \frac{1}{4\pi^2} \iint_{-\infty}^{+\infty} \frac{x(x-a)+y^2+z^2}{[(x^2+y^2+z^2)((x-a)^2+y^2+z^2)]^{1.5}} dx dy \quad \text{Equation (3.15)}$$

Roy and Apparao (Roy and Apparao, 1971) formulate the depth investigation characteristic (equation (3.16)) as a simple analytical solution for equation (3.15)

$$F_{1D}(z) = \frac{2z}{\pi(a^2+4z^2)^{1.5}} \quad \text{Equation (3.16)}$$

Equation (3.16) is employed by many authors to determine the properties of different arrays in resistivity sounding surveys (Barker, 1989; Edwards, 1977; Loke, 2013; Merrick, 1997).

A plot of this function for Wenner array shown in Figure 3.8 and some authors have applied the maximum point as the depth of investigation of the array based on this figure. Edwards (Edwards, 1977) and Barker (Barker, 1989) have presented that a more robust calculate is the "median depth of investigation". That means, the depth of investigation is 0.519 time electrode spacing for Wenner array. However the depth is strictly valid for a homogeneous earth model, they are possibly sufficient for organizing field surveys (Loke, 2013).

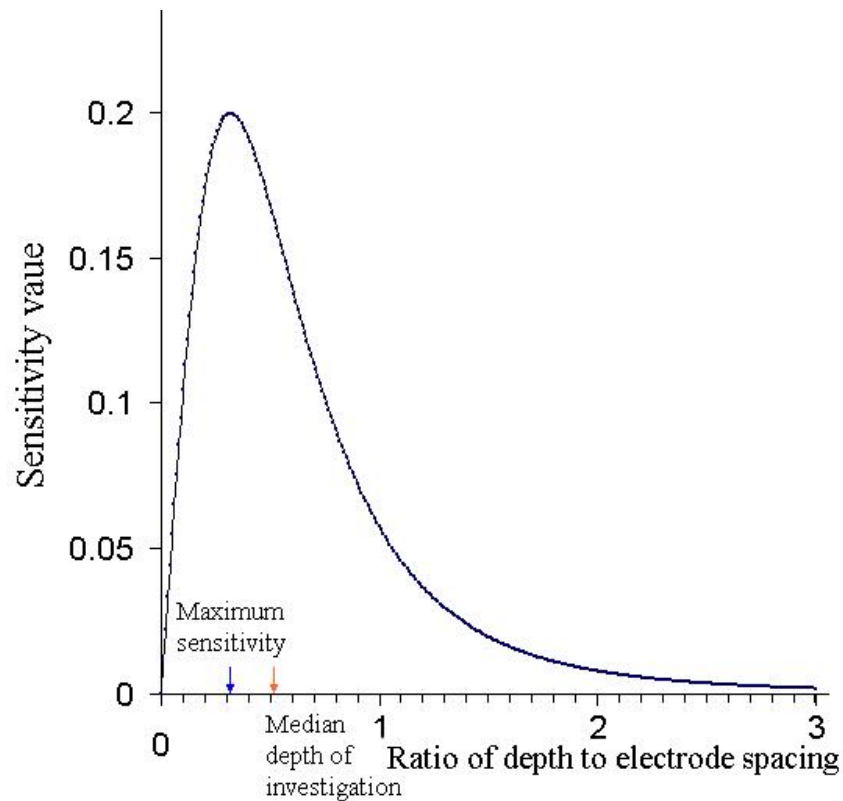


Figure 3.8: The sensitivity function and median depth of investigation for the Wenner array (Loke, 2013)

3.2.10 Considerations and Limitations of the Electrical Resistivity Imaging Method

The electrical resistivity Imaging (ERI) method is an efficient tool; although the choice of array, field conditions, team knowledge, data processing as well as interpretation time are the factors that control the applicability of this method. The ERI method has advantages and restrictions. The ERI technique is capable to be applied to effectively map variable depth to top of rock, prominent solution widened joint sets, fracture zones, and water filled openings. The criteria and restrictions with applying the Wenner electrode array are summarized as following:

ERI have some advantages including the point that ERI method has lots of applications such as lithology, porosity, ground water mineralization, depth to bedrock, detection of air- and water filled openings in rocks, joints orientation, mapping variable depth to top of rock, terraces mapping, mapping aquifers and so on. Profiling is capable to determine the underground layers' depth and thickness, depth to the water table, and bedrock. Fields are cost effective and the equipment is light, portable and inexpensive. Moreover, qualitative data interpretation is straightforward and fast and finally shallow investigations are quick.

On the other hand, there are some limitations for this method. For example, similar to all geophysical approaches resistivity data are ambiguous, which means many various "models" are able to produce similar data. So to define a number of feasible models, other geological data such as borehole and/or monitoring well data are required. The data depends on near surface conductive layers. The current will always moves along with highly conductive layers, easily. In the case of highly conductive surface, data collection below the top layer may not be possible. Deep investigations need long cables and it is time consuming and accuracy reduces with increasing depth.

3.2.11 Electrical Resistivity and Porosity

Many researchers have identified a correlation between the properties of materials measured using geophysical methods and their hydrogeological properties

A fraction of water present controls rock resistivity and it will decrease as water salinity increases as Archie (Archie, 1942) described. The first relationship of the electrical resistivity of saturated soil (ρ_{Sat}) to the electrical resistivity of its pore fluid (ρ_w) was derived by Archie using the formula

$$F = \frac{\rho_{\text{Sat}}}{\rho_w} = \frac{a}{\phi^m} \quad \text{Equation (3.17)}$$

where

F is the formation factor, ϕ is porosity while 'a' and 'm' are constants related to the type of soil and rock. Equation 1 presents the electrical resistivity sensitivity of saturated soil to porosity, pore fluid, and type of soil. a =1 is for unconsolidated clay free soil and "m" usually refers to the cementation factor that ranges from 1.4 to 2.2 for clean sand and gravel encountered in groundwater aquifers (Abu-Hassanein et al., 1996). The electrical resistivity of soil and rock depends on ionic content in the pore fluid. The Archie's Law is not valid (Keller and Frischknecht, 1966) for rocks containing a significant percentage of clay. A reduction in resistivity can be observed when the ionic content of pore fluid increases. The presence of insulating material in pore water can result in a large increase in resistivity (Campanella and Weemees, 1990). Regarding equation 3.17; electrical resistivity of sediment without clay is expressed by

$$\rho = \frac{a\phi^{-m}\rho_w}{S_w^n} \quad \text{Equation (3.18)}$$

where

ρ_w is the resistivity of pore fluid, ϕ the is porosity, S_w^n the degree of saturation and a, m and n are constants related to the soil or rock type. Clayey sands show a complex behaviour. In the case of clayey sand, the formation conductivity is generally larger due to added surface conditions because of the presence of ions in the double layer at the surface of clay particles (Patnode and Wyllie, 1950; Waxman and Smits, 2003).

$$\rho = \frac{\rho_w F_t^*}{1 + \rho_w B Q_v} \quad \text{Equation (3.19)}$$

where;

F_t^* is the formation factor, B the average mobility of the cations close to the grain surface and Q_v is cation concentration per unit pore volume. It reflects Archie's first equation for $Q_v=0$ when there is no clay in the sediment.

3.3 Conceptual Model

3.3.1 Borehole Data

The stratigraphic sequence has described according to 17 boreholes drilled in the area by the Mineral and Geosciences Department. These wells were installed for surveying physical and chemical properties of groundwater and drilled until reaching the basement. The drilled boreholes explain the thickness of different type rocks' changes vertically with depth. Table 3.2 is the list of surveyed wells in the study area. Figure 3.9 shows a typical lithologic description of wells No 2, 7, 16 and 17 in the study area.

Table 3.2 Table of existing well information

| No | Elevation (masl) | Depth to water (m) (October 2013) | Depth of Bedrock (m) | Depth of Borehole (m) |
|----|------------------|-----------------------------------|----------------------|-----------------------|
| 1 | 9 | 5.04 | 58 | 60.5 |
| 2 | 12 | 3.85 | 67 | 70 |
| 3 | 7 | 1.55 | 58.5 | 61 |
| 4 | 8 | 2.11 | 38 | 39.5 |
| 5 | 8 | 3.84 | 65 | 67 |
| 6 | 9 | 0.93 | 36.25 | 38.5 |
| 7 | 10 | 1.4 | 25 | 27 |
| 8 | 11 | 0.68 | 27 | 33 |
| 9 | 11 | 0.49 | 31 | 36 |
| 10 | 12 | 1.14 | 27 | 34 |
| 11 | 11 | 1.2 | 36 | 38 |
| 12 | 10 | 1.16 | 40 | 42 |
| 13 | 8 | 2.58 | 32 | 35 |
| 14 | 8 | 2.31 | 30 | 32 |
| 15 | 11 | 1.05 | 27 | 33 |
| 16 | 12 | 0.13 | 31.8 | 34 |
| 17 | 10 | 0.85 | 36.5 | 37 |
| 18 | 8 | ----- | 43 | 45 |
| 19 | 10 | ----- | 48.2 | 52 |

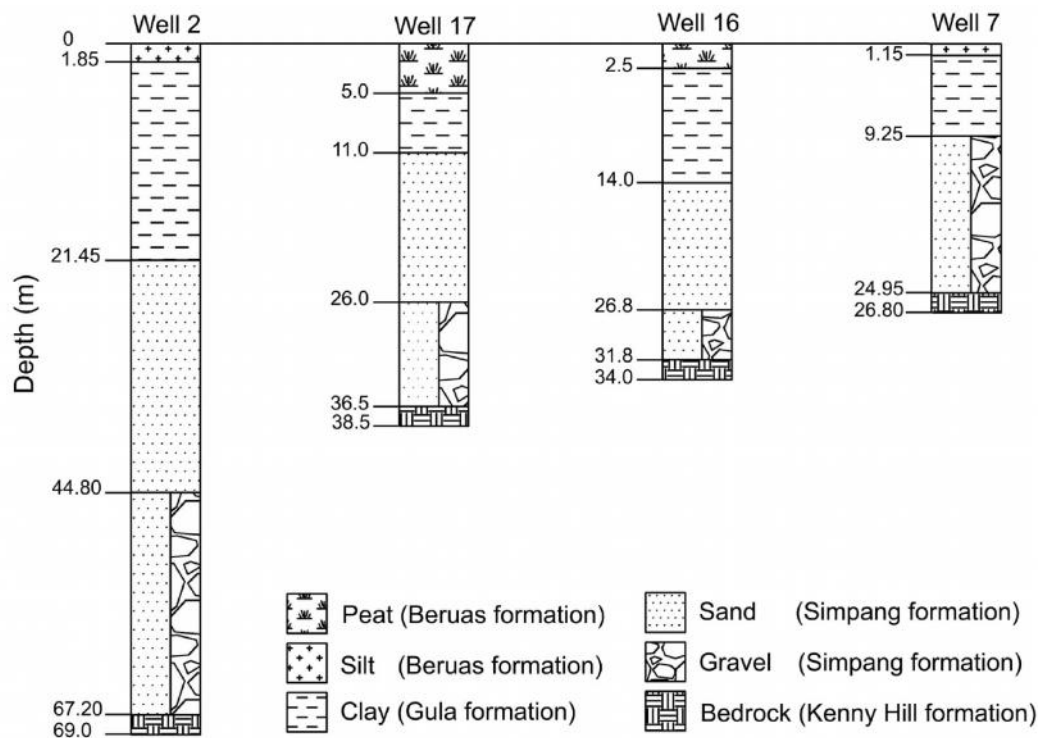


Figure 3.9: Typical stratigraphic log of the geological environment in the study area

Borehole data indicates that the thickness of sand and gravel layers vary between 10 and a maximum of 40 m. According to data obtained from boreholes, the study area consists principally of interbedded sand and clay layers, which overlay the metasediment bedrock. In general, the thickness of the clay, sand and gravel layers become thicker towards the coast.

3.3.2 Geoelectrical Resistivity and Lithological Correlation in the Study Area

The correlation of resistivity profiling with lithological borehole logs or well logging data is the best way to gain a realistic interpretation of the profiling. The clay content and the total dissolved solids concentration in the pore water mainly control the electrical properties of the subsurface (Scanlon et al., 1999). Uniquely low resistivity values are due to high levels of total dissolved solids at the west of the study area as it is expected in saline water. Lithological information from 20 boreholes was applied to interpret the resistivity results to limit the ambiguity in the interpretation of resistivity profiles. Two typical calibration ERI surveys were performed for lithologic correlation purposes within existing wells with high quality lithologic logs. The comparison of the inverted resistivity-depth curves with the lithologic logs demonstrate resistivity ranges characteristic of the main sedimentological facies that were encountered in the boreholes as shown in Figure 3.10. Figure 3.10a is related to northern region (upside) and Figure 3.10b corresponds to southern region (down side) of the Langat River which southern region was affected by sea water intrusion. The correlation assumed no lateral lithology variation happened between well locations and the ERI survey points. The direct comparison of lithology versus apparent resistivity information can be attempted from these correlation ERI survey result.

Regarding to these local correlation values of resistivity of less than 22 ohm.m demonstrate peat and clay zones while, values of resistivity ranging between the 22-132 ohm.m represent sand with gravel zones and values more than 132 ohm.m demonstrate bedrock for the upper part of Langat River. For the down part of Langat River, values of less than 4 ohm.m represent peat and clay zones, values of resistivity between 4-28 ohm.m were taken to represent sand with gravel zones, values more than 28 ohm.m represent bedrock for the down part of Langat River.

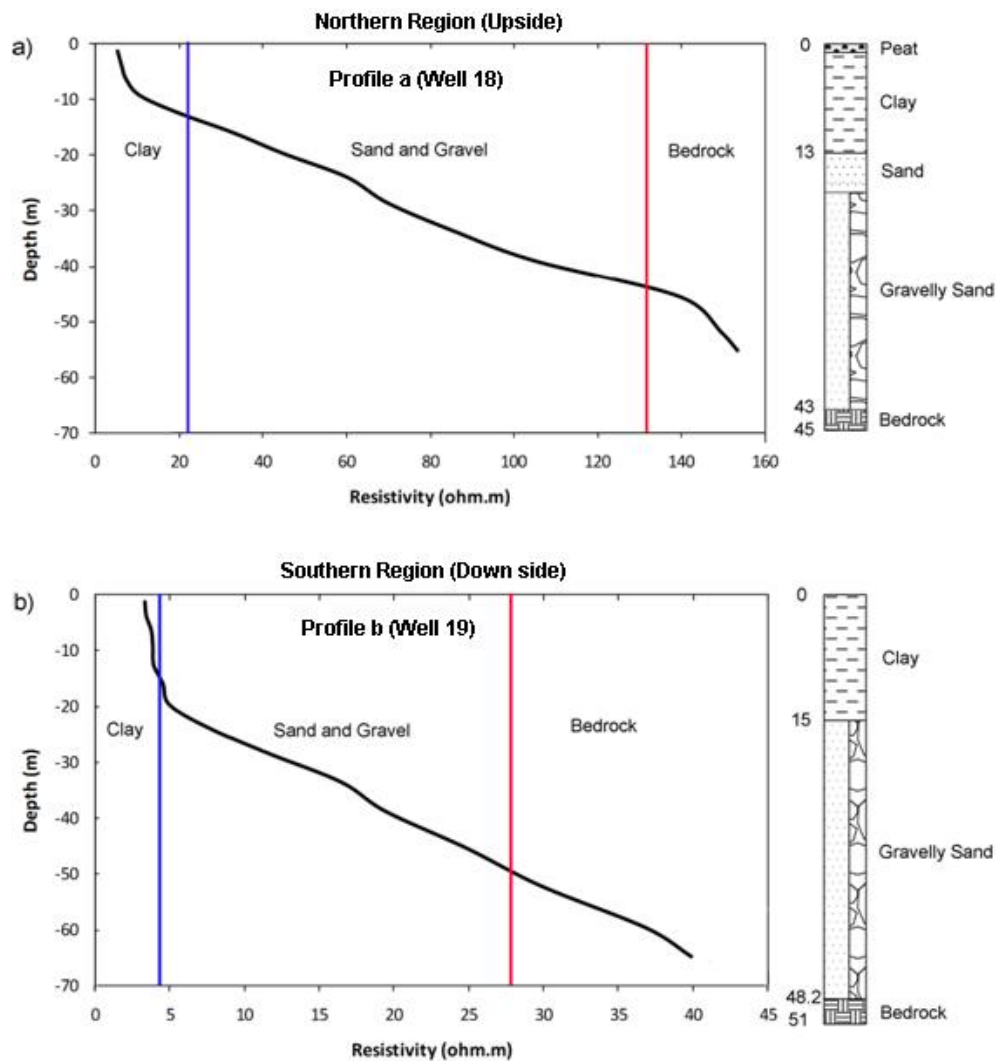


Figure 3.10: Correlation of ERI results based on direct comparison to lithologic logs at two wells.

3.3.3 Fence Diagram

JICA (2002) used exploration wells to show that alluvial sediments, mainly consisting of sand and gravel, are distributed widely in the lowlands of the Langat Basin. These are Terrestrial Pleistocene sediments (Lower Sand Member of Simpang Formation) and considered as an important aquifer in the Basin. The Alluvial sediments are covered by younger clayey layers and then, at some locations, by a peat layer. Figure 3.11 shows a fence diagram of the study area based on the borehole information. The aquifer thickness of the area gradually increases from east to west of the area. The aquifer consists of fine to sand or gravel and some aquifer is mixture of sand and gravel. The depth of aquifer is less than 15 m and around 60 m at the east and west of the study area respectively.

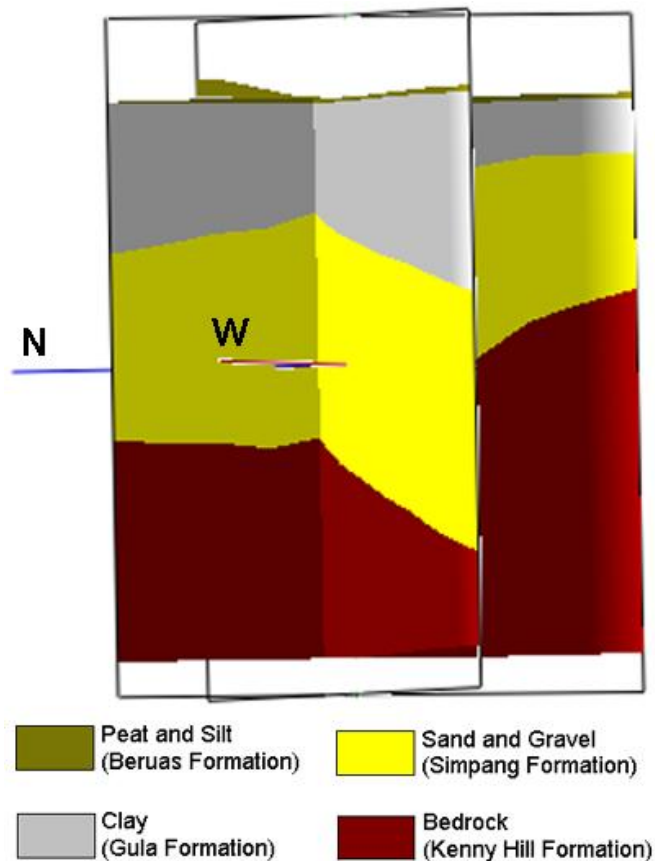


Figure 3.11: Fence Diagram of the study area based on boreholes information

3.4 Hydrogeological Method

3.4.1 Water Level Monitoring

The physical condition of the well is described by the measurement of the water fluctuation level with respect to time, ground level where the water sample is collected and depth of well. For determination of the position, the coordinates of the wells are collected while the groundwater level data is collected for determination of the location of the well which is installed relatively to the mean above sea level. Groundwater level is influenced by properties of aquifer such as local geological conditions and a complex balance between discharges to rivers or pumped extraction and recharge from rainfall. The data can be employed to quantify resources, to measure the impacts of extraction and to understand the groundwater flow.

3.4.2 Hydrogeochemical Method

Geoelectrical resistivity survey is not capable for accurate interpretation of groundwater condition such as total dissolved solid (TDS), electrical conductivity (EC), pH and other cation and anion parameters. To assess groundwater quality, other methods such as the hydrogeochemical method were used in the study area.

The study on the chemical character of groundwater was carried out because of following reasons: the compatibility of the water for drinking water supply, irrigation, and industrial use can be determined; the background conditions to identify the occurrence of contaminants and to monitor the effectiveness of remediation can be established. To obtain hydrogeochemical data, two conditions of measurement are required. First, in-situ parameters at the site and second, the concentration of cations and anions in the water samples are measured.

3.4.2.1 In-situ Parameters

Immediate water parameters measurement consist of any measurements which are obtained through field equipment and the obtained data are recorded on a field sheet In-situ water parameters which are considered the standards for basic water quality testing consisting of pH, temperature, electrical conductivity, total dissolved solid, and salinity. The physical well parameters which are well depth, water level, well coordinate and ground level are measured at the site

3.4.2.1.1 Temperature

Temperature as a significant water quality parameter has a direct effect on the organisms living in the water. Temperature is relatively easy to measure and generally in most groundwater conditions including in situ contamination, temperature is presumed constant at a value of 1 to 2°C more than the annual average ambient air temperature for the area (Freeze and Cherry, 1977).

3.4.2.1.2 pH

Essential information concerning many chemical and biological processes and indirect correlations to a number of various impairments can be provided by pH of natural water. The reaction that creates or consumes hydrogen control pH in aqueous solutions and it is strongly affected by the groundwater temperature. These reactions consist of practically all dissociation of acids, bases, and hydrolysis (Boghici, 2003). pH values vary from 0 to 14 where, values below 7 pH exhibit acidic properties and values above 7 pH belongs to basic (also known as alkaline) properties. 7 pH is the middle of the measurement level and is known as "neutral" which is neither acidic nor basic. pH is

determined mathematically as the negative logarithm of the hydrogen ion concentration which is shown as

$$\text{pH} = -\log \text{H}^+ \quad \text{Equation (3.20)}$$

where (H^+) is hydrogen ion concentration in mol/L.

The ideal pH level of drinking water is between 6-8.5. pH equilibrium in the human body remains on a constant basis and it is not dependent on water consumption. Measurements of pH is significant in many applications such as medicine, biology, agriculture, chemistry, food science, forestry, environmental science, civil engineering, oceanography, water treatment and water purification.

3.4.2.1.3 Total Dissolved Solids

The phrase applied to explain the inorganic salts and small amounts of organic matter contained in solution of water is total dissolved solids (TDS) which TDS concentrations are generally employed for evaluation of the water suitability for domestic provides and industrial purposes.

In general, the TDS concentration is considered as the sum of the cations and anions (positively and negatively charged) ions in the water. The TDS concentration has impacts on the water equilibrium in the cells of aquatic organisms. TDS is not considered mainly as pollutant. TDS has generally high ranges due to hard water and it may cause to scale buildup in pipes, reduced water filters efficiency, hot water heaters, etc., and causes aesthetic problems such as salty or bitter taste.

Primary resources for TDS consisting of waters are agricultural run-off, urban run-off, sewage, industrial waste water, and natural resources such as silt, leaves, and rocks. Additional exotic and dangerous elements of TDS are pesticides arising from surface runoff (Grasby et al., 1997). In general, finished drinking waters consist of TDS over 1000 mg/L are considered as improper. These waters also are not acceptable in agricultural uses because they have harmful effects of excessive ionic concentrations on plants.

3.4.2.1.4 Electrical Conductivity

The normalized measure of the water's capability to conduct electric current called electrical conductivity (EC) in natural waters and is the inverse of electrical resistivity. Dissolved salts such as sodium chloride and potassium chloride mostly affects on electrical conductivity. Siemens per meter (S/m) is the common unit for electrical conductivity. Electrical conductivity variations because of other factors that impact electrical conductivity, for instance lithology, are smaller than or separable from electrical conductivity variations which happen because of contaminant. Generally, organic contaminants decrease electrical conductivity and inorganic contaminants commonly increase the electrical conductivity (Delleur, 2010). An abundance of dissolved salts can be known as the source of EC, because of poor irrigation management, minerals from rain water runoff, or other discharges.

3.4.2.1.5 Salinity

Investigation on salinity and chemistry of groundwater have been popular and noticed by many researchers. The evaluation in salinity of water can be done in several ways. The sum of concentration of all individual, inorganic ions in water corresponds to

salinity (Wetzel, 2001) which usually is measured with salinity refract meter. In addition, Analyses to calculate salinity obtained from the sum of ions are prohibitively expensive. The specific conductance of water increases with enhancement of ions concentration. Therefore, specific conductance is known as a good indicator of total dissolved solids and salinity (Hem, 1985). Meters that read total dissolved solids or salinity directly are modified conductivity meters. There is no great common agreement concerning the degree of salinity, total dissolved solids, particular conductance, or chloride that determines a freshwater form a saline one. Applying 1 ppt salinity or a specific conductance of 1.5 mS/cm is suggested by Boyd (Boyd, 2002) who creates this separation. Three main types of salinity consist of Dry land salinity which takes place when saline seeps scale the surface of non-irrigated lands, which in turn affects plant growth and degrades soil structure. The second type is irrigation salinity which results from overuse of water for irrigation causes rising water tables and the last type is the intrusion of sea water which happens in coastal aquifer systems where sea water replaces groundwater that has been over-exploited (Islami, 2011).

3.4.2.2 Cation and Anion

Calculating the cation–anion balance is used to examine the precision of chemical analyses of water samples. The sum of the major anions the sum of major cations should be equal. This examination is usually performed by calculating the ionic balance error of the major ions where (Hiscock, 2009):

$$\text{Ion balance error (\%)} = \frac{\sum \text{cations} - \sum \text{anions}}{\sum \text{cations} + \sum \text{anions}} \times 100 \quad \text{Equation (3.21)}$$

An ionic balance error of less than 5% and certainly less than 10% is achievable with modern analytical equipment. Larger errors suggest that one or more analyses are in error and consequently, they are not acceptable.

3.4.2.2.1 Major Cations

1 Calcium

Calcium (Ca) is available in all water sources due to its abundance on the earth which is around 4.9% of the Earth's crust. The Ca concentrations are commonly less than 20 mg/L in groundwater but it can reach as high as 200 mg/L and ~400 mg/L in seawater. Concentrations of Ca in surface water change greatly. Calcium is known as one of the essential elements for humans and plants. Direct toxic impacts related to high concentration of Ca in water are not observed except in drinking water, high concentration of Ca may cause to increase kidney stones' risk (Li et al., 2010).

2 Iron

One of the most abundant metals on the crust of the Earth is Iron (Fe) that is found in natural fresh waters at levels varying from 0.5 to 50 mg/L (World Health Organization, 2004). Fe can be found in drinking water due to using iron coagulants or the corrosion of steel and cast iron pipes through water distribution. Dissolution of iron-containing minerals, and human activities and the decay of organic matters can be caused to the existence of Fe in water. Fe is known as an essential element in human nutrition, which particularly is in the iron oxidation state.

Dissolved Fe in groundwater is determined by the amount of oxygen in the water level of acidity. In deep aquifers, especially when the aquifer consists of organic matter,

the dissolved oxygen amount is commonly low. The decomposition of the organic matter depletes the oxygen in the water and the iron is dissolved as Fe^{2+} . The dissolved iron is mostly accompanied by dissolved hydrogen sulphide or manganese, under this condition. If this water is excavated to the surface, the dissolved iron and oxygen reaction occurs in the environment, varies to Fe^{3+} and rust- colour iron minerals is formed (Drever, 1997).

3 Magnesium

The eighth most abundant natural element is Magnesium (Mg) that exists in all natural waters as one of the main cations in water. Mg is known as a one of main contributors to hardness of water. 2.5% of the crust of Earth contains Mg and is generally obtained in such minerals

The erosion of rocks for instance limestone and dolomite, and minerals, such as calcite and magnetite is the most typical source of magnesium in groundwater.

4 Sodium

Sodium (Na) is a highly soluble chemical element which is naturally found in groundwater. In water, Na has no smell although it may be tasted by most people at contents of 200 milligrams per liter (mg/L) or more. Pollution from point or non-point sources or salt water intrusion may occur with increasing of sodium in groundwater above ambient or natural degrees. Some sodium exists in all groundwater because most rocks and soils consist of sodium compounds because sodium is easily dissolved. There are common sources of elevated sodium levels in groundwater which are a) Erosion of salt deposits and sodium bearing rock minerals, b) Natural occurrence of brackish water of some aquifers naturally, c) Intrusion of salt water into wells in coastal areas, d)

Infiltration of surface water which is contaminated by road salt, e) Irrigation and precipitation leaching through soils which are high in sodium, f) Pollution of groundwater by sewage effluent, g) Infiltration of leachate from industrial.

5 Potassium

Potassium (K) is an important element for humans and is rarely, if ever, present in drinking water at degrees that is significant for humans health. Potassium exists broadly in the environment which consists of all natural waters. It considers also as a result of the use of potassium permanganate as an oxidant in water treatment, in drinking water (World Health Organization, 2004). Potassium exists usually in rocks and soils which includes appreciable content of clay. Dissolution of rocks release sodium and potassium slowly upon. Moreover, potassium is a significant element in chemical fertilizer (Yang et al., 2006). Though potassium can cause some health impacts in sensitive individuals, its consumption through drinking water is approved below the level at which adverse health effects may occur. Health issues would be dependent on the intake of drinking water that is potassium-based water treatment (World Health Organization, 2004).

3.4.2.2.2 Major Anions

1 Chloride

One of important anions in water is Chloride (Cl) and the salts of hydrochloric acid HCl include chloride ions and can also be known as chlorides. The concentrations of Cl are generally less than 100 mg/L in groundwater and its concentration can be as high as 19,000 mg/L in seawater (Boyd, 2000). Concentrations of Cl > of 250 mg/L are progressively likely and detectable by taste.

High Cl concentrations in the groundwater and surface water happen frequently due to the intrusion of saltwater, industrial and homemade wastes and dissolution of mineral. When Cl is combined with Na, chloride causes to increasing a salty taste to drinking water and it results in increasing the corrosiveness of water. Human health is not influenced by normal concentrations of chloride, although increasing salinity can have effects on freshwater aquatic organisms (Li et al., 2010).

In addition, Cl is applied as a reliable chemical indicator of river / groundwater faecal contamination because it is a ubiquitous and non-reactive solute to sewage and potable water.

2 Sulphate

Sulphur (S) and oxygen (O) build up Sulphate (SO_4^{2-}) composition. Sulphate is naturally in many formations of soil and rock. Moreover, human activities may be caused existing of sulphate in well water. Most sulphates in groundwater are produced from the minerals dissolution, such as gypsum and anhydrite. The intrusion of saltwater and drainage of acid rock are also found as sulphates sources in drinking water. Taste of water may change with presence of above 500 mg/L of sulphate in water. Sulphate in drinking water can have a laxative effect when it exceeds 1000 mg/L, though these degrees are not usually available in drinking water.

3 Nitrate

Nitrate (NO_3^-) as an important plant nutrient is available naturally in the environment. It is a part of the nitrogen cycle and exists at changing contents in all the plants. Nitrate concentrations are generally less than 2 mg/L in natural groundwater and its contents can be reached over 100 mg/L in contaminated water. The oxidation of nitrogen comprises around 78% of the atmosphere, produces nitrate and nitrite (NO_2).

It is carried out by micro-organisms in plants, soil or water and, to a lower degree, by electrical discharges such as lightning. Nitrite does not typically exist in remarkable concentrations except for in a reducing environment, due to the fact that nitrate is found in a steadier oxidation state (World Health Organization, 2004). Nitrate can reach both surface water and groundwater as a result of different procedures which is initially, agricultural activities such as excessive application of inorganic nitrogenous fertilizers and manures from wastewater disposal and then, from oxidation of nitrogenous waste products in human and animal excreta, such as septic tanks.

3.4.2.3 Field Procedures

Collecting of a raw water sample as a representative of the hydrochemical conditions in the aquifer is the purpose of field sampling. Attention should be drawn to initial flush of the water standing in the well through three or four times remove up of the well volume prior to sampling.

Two various kinds of samples can be obtained from aquifer. Rinsing the bottles of water sample with deionized water prior to sampling had been done and a 0.45 μm membrane filter was used for samples containing suspended sediments or particles. Prior to analyses, the standard approach of preservation was done to add a few drops of concentrated nitric or acetic acid done after the sample has been obtained and filtered to reduce the pH to 2 (Hiscock, 2009). It is necessary to maintain samples away from other reactions before the processing step in the laboratory. Once collected, samples must be kept in a cool box and shipped immediately, ideally the same day, to the laboratory for storage at 4°C or fewer, without freezing, until analysed.

The inversed auger method is employed to measure the groundwater level near the electrical resistivity profiling. Four boreholes (section 4.2.5) with the maximum

depth of 3 m along the four ERI profiles are created using hand auger with 11 cm of diameter and 60 cm depth (Figure 3.12) to obtain groundwater level. After the digging finished, we put the piezometer (Figure 3.13) in all of the wells. After this stage, make sure that the surface in borehole bottom is clear and flat.



Figure 3.12: Hand auger with diameter of 11 centimetres

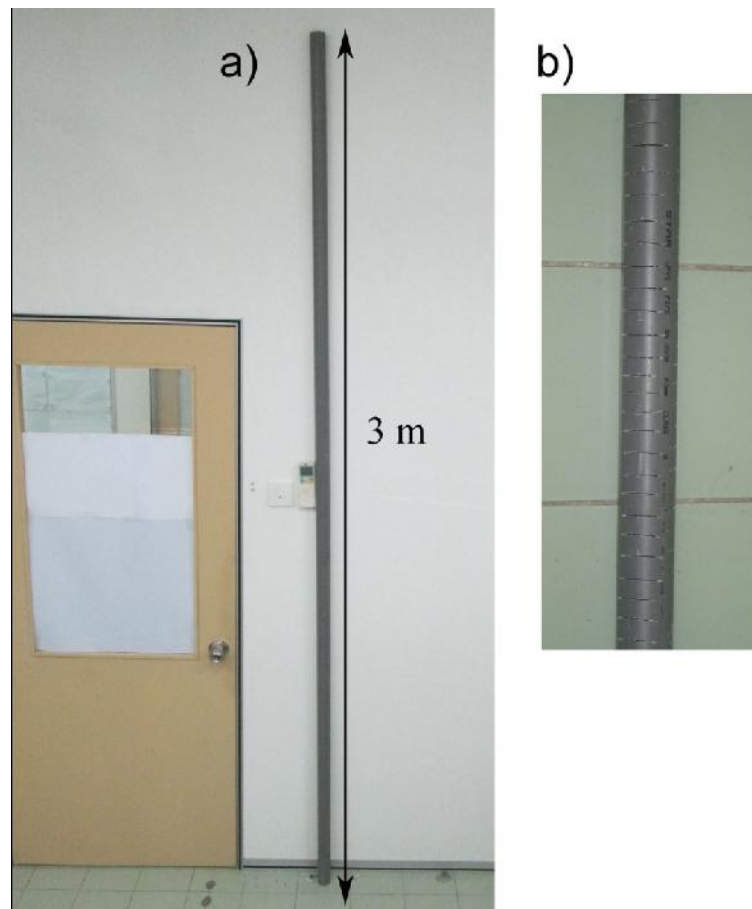


Figure 3.13: Typical piezometer with length of 3 meters

3.5 Artificial Intelligence Techniques

The purpose of artificial intelligence (AI) methods and knowledge engineering is to model human intelligence and cognitive abilities from various aspects. This section is targeted to present the AI techniques and concepts consisting of artificial neural networks (ANNs) and the adaptive neuro-fuzzy inference system (ANFIS). The significant feature of this technology that makes it more interesting is described in hydrogeology research. In addition, structure, components and architecture models of ANN and ANFIS are presented in details. In particular, the useful tools and the supervising learning problems as well as the design and training of layers are discussed.

3.5.1 Introduction

Intensive computational needs are considered as common problems in both model calibration and application. Creating an approximator of the original model which is not as computationally demanding as the original model and is capable to meet the requirements of precision is a way to break the computational bottleneck. The principle of modelling is predicting that is found by mapping parameters set in input space response parameters in output space via a model as shown in Figure 3.14. Figure 3.14 illustrates, conventionally a mathematical model which is placed in the box representing a model. Consequently, the conventional modelling of the underlying systems almost has a tendency to be a very difficult and intractable system.

The more typical field of AI develops expert systems to a technology which is commonly determined as methodologies and computer programs planned for modelling and emulating the thought processes a human expert(s) would employ to perform a usually repetitive and specialized task. AI is aimed to produce algorithms to do cognitive tasks, when that human are currently better (Haykin, 1999). Three things must be done by an AI system. Storing knowledge, applying the stored knowledge for problem solutions, and acquiring new knowledge through experience. The growing AI approaches have the ability to fill the gaps in the measurements and to forecast future values without long observational data (Karimi et al., 2013).

In previous decades, AI methods had been employed as robust tools and precise solutions to most of the extremely complicated problems faced in water sciences and hydrology, and this application has recently draw more attention. Recently, AI has supplied tools to optimize the operations and equipment selection as well as to solve issues including large amounts of information which humans are not capable of coping within the decision making processes. The AI methods are generally employed by ANN

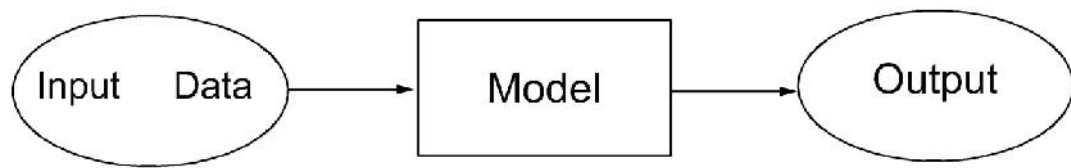


Figure 3.14: An input – output mapping

and ANFIS that is complementary rather than competitive. These techniques are known as appropriate AI tools for simulating complex non-linear systems and they have been widely used to classify and consider water resources.

3.5.2 A Framework for an Artificial Network

ANNs are determined as computational models and inspired by human brain biological neural network (Dayhoff, 1990). Though these models are considered as a subset of AI, artificial neural networks do not obtain the humans learning capacity and a neural network may achieve “learning” and refers to the ability of model to learn patterns by examples (Openshaw and Openshaw, 1997). ANNs is compared to biological neural networks which are dependent on the concept of inter connectedness between neurons in the learning process. The data processing paradigm consists of extremely interconnected nodes (neurons) in which a complicated input structure is mapped with a related output structure. ANNs as a machine learning tool are permitted to fit very general nonlinear functions to experimental data sets; the accessibility of suitable data is mandatory for any statistical approach, (Dreyfus, 2005) which is the actual target of a trained back-propagation network. If correctly trained, reasonable answers are given when ANNs are presented with inputs that they have never been seen. Generally, when presented with a new input, similar to one of the training inputs, the network will give an output similar to that training output. This generalization property causes to train a network on a representative set of input/target pairs and give great

outputs without training the network on all feasible input/target pairs. A structure for ANN model can be known as processing elements, activation functions, connections between units, external input and learning rule (Rumelhart et al., 1986).

3.5.2.1 Proceeding Elements

Processing elements similar to biological neurons involve input and output paths which emulate the axon and dendrite functions. In an ANN, every processing element receives multiple input signals sent from one set of artificial neurons and then it transmits an output signal to another set of neurons. In ANNs, the signals that are being transferred from one neuron to another are numeric values (Dayhoff, 1990). Processing elements assign connection strength to every input when connected neurons have sent input signals. A simple numerical model of an artificial neuron proposed by McCulloch and Pitts (McCulloch and Pitts, 1943) displays in Figure 3.15. The processing element calculates a weighted sum of its inputs and outputs, in this model that is based on whether this weighted input sum is above or below a certain threshold θ_i . The equation 3.22 indicates the calculation in the i th neuron:

$$y_i(t + 1) = f \left[\sum_{j=1}^m w_{ij} x_j(t) - \theta_i \right] \quad \text{Equation (3.22)}$$

where, t is the time instants, m is the number of the inputs to the i th neuron and y is output from the i th neuron.

The activation function $f(z)$ is a binary step function and given as:

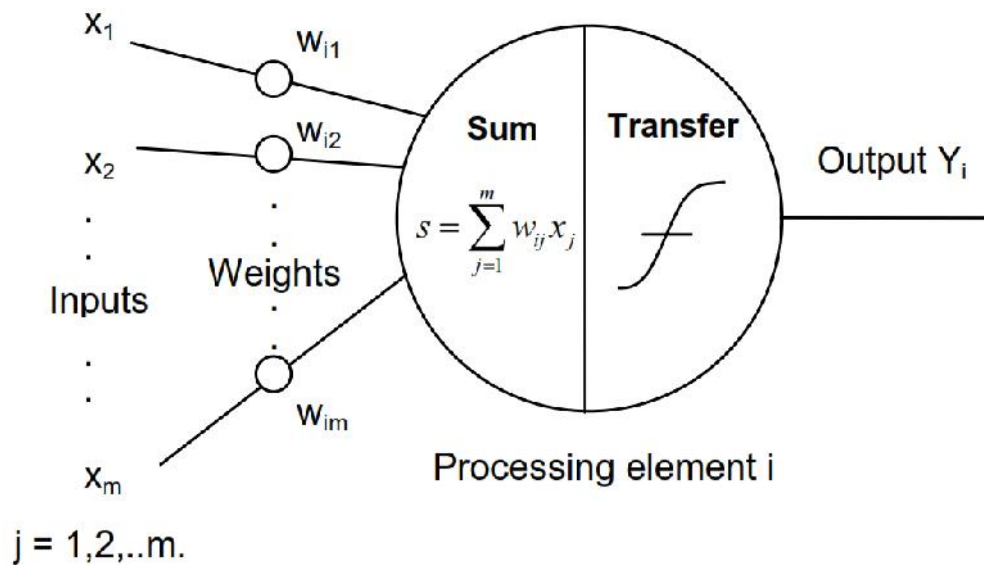


Figure 3.15: A typical artificial neuron

$$f(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{Equation (3.23)}$$

The weight , w_{ij} , represents the strength of the synapse connecting input variable j to neuron i . A positive and negative weight is corresponded to an excitatory synapse and an inhibitory synapse, respectively. If $w_{ij} = 0$, then there is no connection between the two neurons.

3.5.2.2 Activation Functions

Processing elements may consist of broadly various kinds of numeric data, in ANNs. As processing elements receive input data through multiple neurons, individual processing elements should process every input value and hence they produce an output value. Processing elements transform the summed input data to an output through an activation function, to complete the process. The activation function aimed to turn the

summed input values ranging from plus to minus infinity, into a scaled output value (Hagan et al., 1996). In order to scale output values in an ANN, a wide variety of activation functions can be applied which include both linear and non-linear functions. The most typical functions in ANNs consist of sigmoid that is a family of curves including logistic and hyperbolic tangent functions (Samarasinghe, 2006). The sigmoid function is expressed as:

$$f(z) = (1.0 + e^{-z})^{-1} \quad \text{Equation (3.24)}$$

Figure 3.16 shows this function as a scaling function and converts inputs values to an output value in the range of 0.0 to 1.0. Sigmoid functions are differentiable, so they are generally applied to many artificial neural network architectures (Hagan et al., 1996). Hyperbolic tangent functions such as sigmoid functions are known as a family of S-shaped functions (Samarasinghe, 2006). Consequently, this function scales the output data between the values of - 1.0 and 1.0. The hyperbolic tangent function is described by following equation:

$$\tanh(x) = \frac{x - e^{-x}}{x + e^{-x}} \quad \text{Equation (3.25)}$$

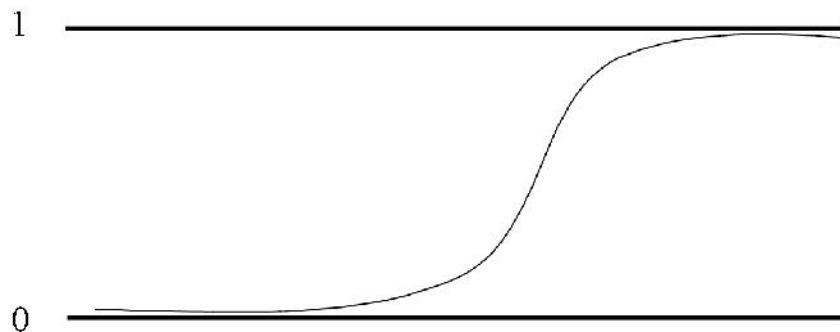


Figure 3.16: Sigmoid function

3.5.2.3 Learning Rules

ANNs are often applied as predictor models that it involves the capability of models to calculate the value of a specific variable according to the known values of other variables. ANNs have been considered as efficient predictors because the networks can learn by examples. In order to learn, most ANNs should initially display an input data structure besides a set of known or “desired” results. This kind of learning procedure is named supervised learning or associative learning (Hassoun, 1995). A supervised learning does not involve competition; however it employs an identified output for every input structure that leads the learning process. The network will compare the results obtained from the network against the desired results and define an error value (Samarasinghe, 2006), in order to learn how to compute the desired results. Reducing the error occurs with updating the connection weights between the layers regularly by network and recalculating the results before the network can no longer modify the error between the two values and eventually it reaches the convergence.

After the network has attained convergence, the testing step as the second step in the learning process is completed. During this step, the network presents input datasets again. However it is not similar to the training step, no desired output is announced. The connection weights are not adaptive, but rather fixed, while the testing step is implemented. The testing step attempts to define the accuracy of neural networks based on the connection weights values at the time of network convergence.

A number of neural network architectures also use an unsupervised learning rule besides supervised learning. Unsupervised learning rules do not need a desired output to be announced which is in contrast with the supervised learning, where a desired result is presented for the network from which to learn. Unsupervised learning relies only on the

ability of network to define patterns within the input data for the purpose of producing an output, instead of presenting a desired output (Hassoun, 1995).

3.5.2.4 The Effects of the Number of Hidden Layers

A rough principle for selecting the number of neurons within the hidden layer has been provided by Kolmogorov's Theorem (Kolmogorov, 1957). Regarding this theorem, an exact mapping from m inputs to n outputs with $2m+1$ neurons in hidden layers are resulted from a network with one hidden layer. Other rough criteria for selecting the number of hidden neurons is the geometric pyramid rule (Masters, 1993). For many practical networks, it states that the number of neurons employs a pyramid shape which the number reducing from the input to the output direction. Consequently, this criterion does not guarantee that this network is the ideal one. The trial-and-error technique is often the most effective technique to obtain a suitable hidden neurons number.

3.5.2.5 Neural Network Architectures

The system contains a number of simple, highly interconnected processing algorithms (neurons) which makes ANNs and processes data by their dynamic state response to external inputs (Fausett, 1994). Net architecture is the design of neurons into layers and the connection styles within and between layers. Many neural nets consist of an input layer in which every unit is activated equally to the external input signal. Regarding to the network architecture which is structured and on the kind of connections, various types can be found. Three common architectures are applied in current study which will be discussed in the next section.

3.5.2.5.1 Feed-Forward Neural Network (FNN)

One of the most well-known neural networks is the Feed-Forward Neural Network (FNN). FFNs have been utilized effectively in various issues since the emergence of the error back-propagation learning algorithm. A multilayer perceptron (MLP) network is made of an input layer, one or more hidden layers of calculation nodes, and an output layer. Figure 3.17 demonstrates a typical feed-forward network made up of layers i , j and k , with the interconnection weights W_{ij} and W_{jk} between the layers of neurons. The main advantage of feed-forward neural networks is that they are simple to manage, which enables them to approximate any kind of input and output map, as established by Hornik et al. (Hornik et al., 1989). The node is the key to the training of the neural network. The difference between the neural network output and the preferred result is determined at every iteration. The training process becomes manageable, through differentiating the neural network as a result of the nodes.

3.5.2.5.2 Elman Neural Network (RNN)

The Elman neural network (recurrent neural network) is acquired through an additional feedback cycle through the output of a hidden layer to the input of this layer, which makes up the context layer that maintains information between observations (Elman, 1990). The effect of processing in a previous period stage can be employed in the present period stage. This quality of the Elman network provides an extremely important benefit, particularly in real-time purposes to follow the dynamic change of water variables in practice. Figure 3.18 displays a typical Elman network composed of four input nodes, a hidden layer with three nodes and one output. These recurrent networks may have an unlimited memory level and so are able to discover relationships the through time as well as through the immediate input space (Haykin, 1999).

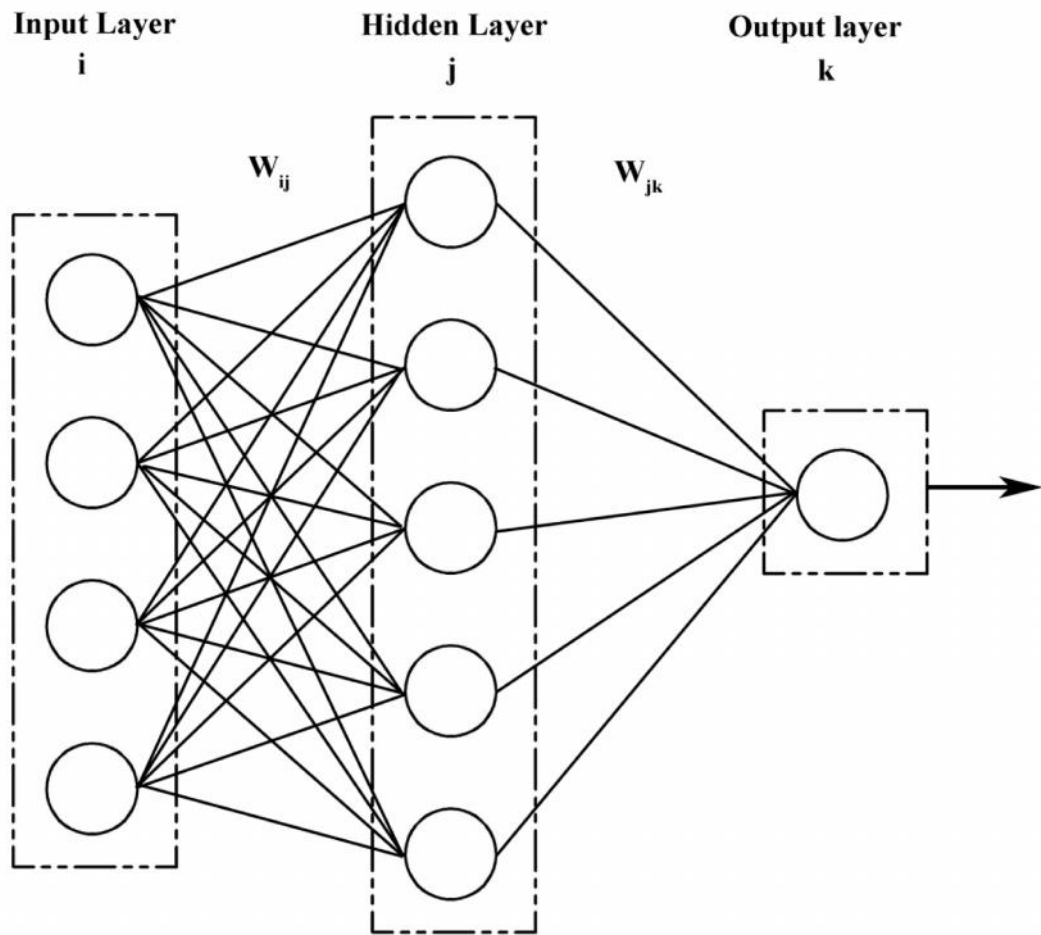


Figure 3.17: Structure of typical Feed-Forward Neural Network (FNN)

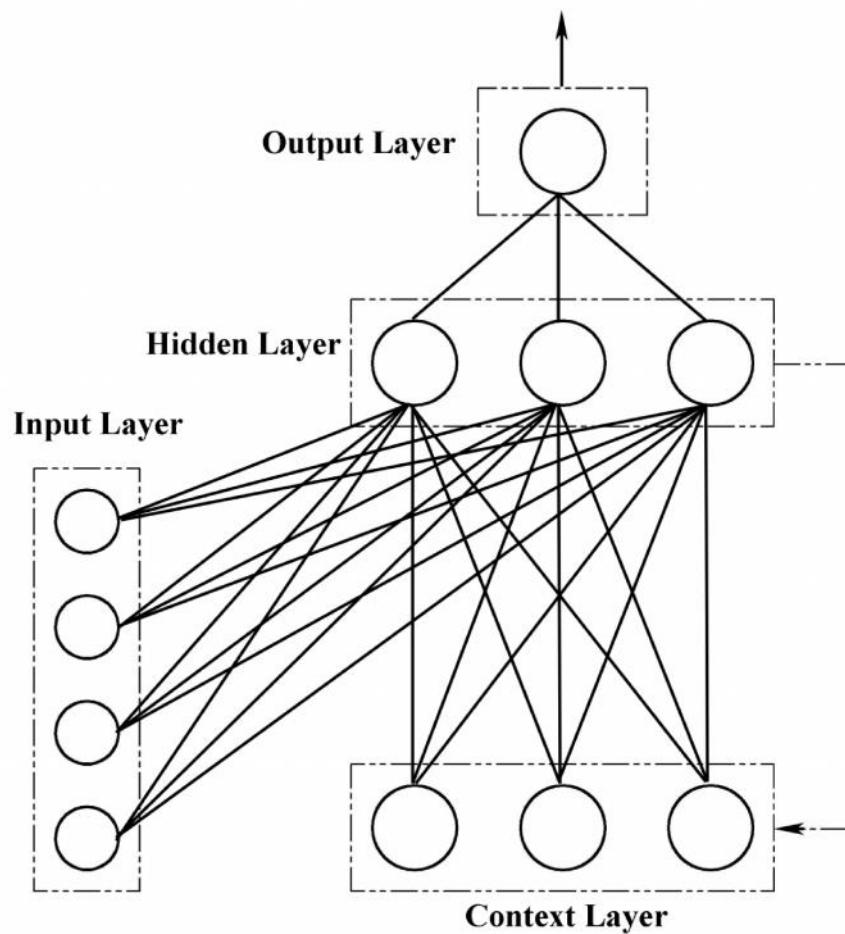


Figure 3.18: Structure of typical Elman Neural Network (RNN)

3.5.2.5.3 Cascade Forward Network (CFN)

The cascade back-propagation algorithm is the basis of a conceptual pattern intended for accelerating learning in artificial neural networks. The CFN neural networks are exactly like feed-forward networks, but they consist of a weight relationship through the input to each layer and also through each layer to the effective layers (Lashkarbolooki et al., 2013). Filik and Kurban (Filik and Kurban, 2007) identified that which cascade forward back-propagation technique can be more effective than feed-forward back-propagation technique within the several cases. Figure 3.19 demonstrates typical cascade-forward back propagation networks.

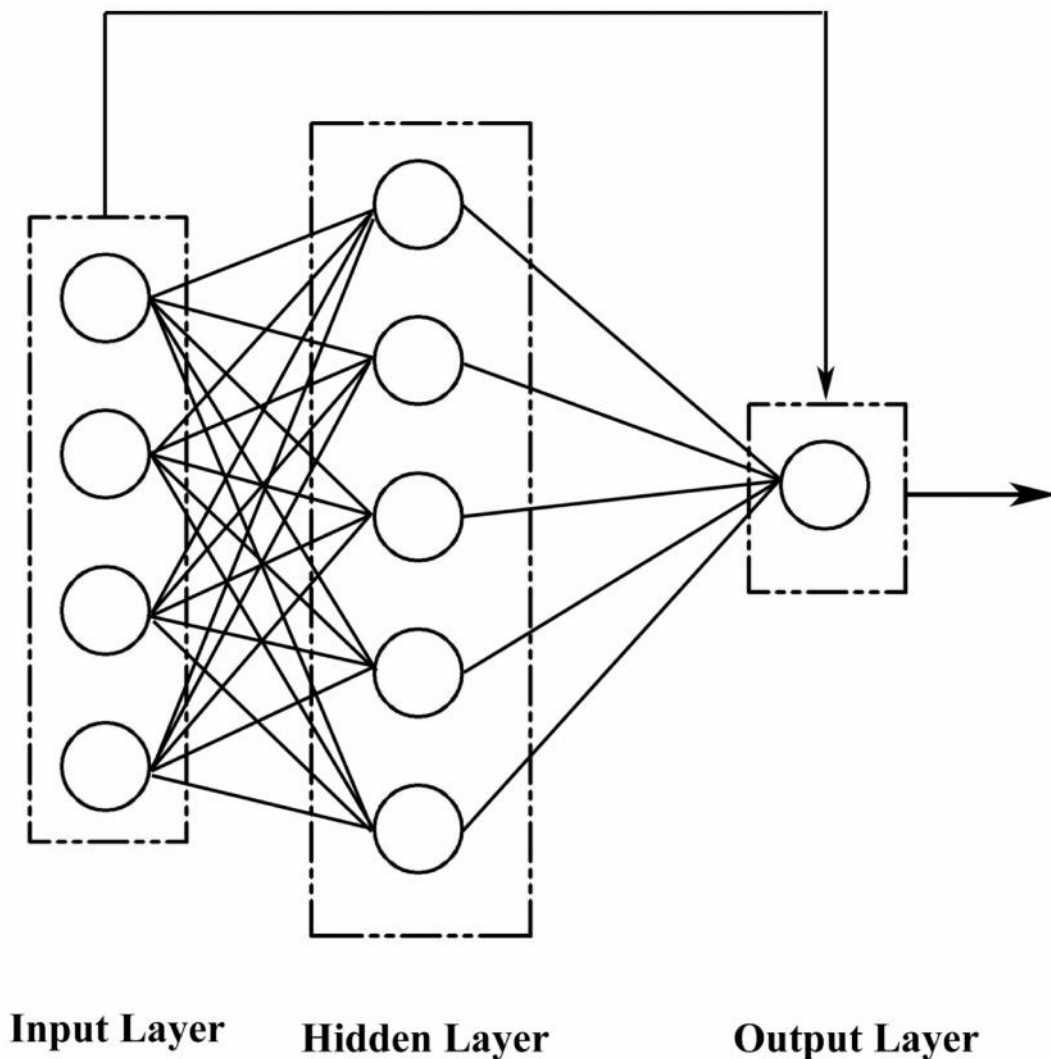


Figure 3.19: Structure of typical Cascade Forward Network (CFN)

3.5.2.6 Back Propagation Algorithm

The back propagation (BP) approach is broadly employed for training algorithm result due to its simplicity and being cost effective program (Bishop, 1995; Gallant, 1993; Ripley, 1996; Rumelhart et al., 1986; Werbos, 1990). It is known as an efficient systematic method used for training multilayer ANNs (Wasserman, 1989). The learning is aimed to minimize the error between the desired outputs and the computed outputs of the network, with the BP learning rules.

The BP algorithm propagates variations to the weights the neural network through modifying the weights of one individual node at the same time. The optimization approach known as gradient descent is the numerical foundation for the back propagation algorithm (Rumelhart et al., 1986). The computing the partial derivations of the error function, with respect to each weight results in the essential corrections to the network weights for each moment. The resulting weight update has been computed. The BP may consist of several types of paradigms (El-Qady and Ushijima, 2001; Werbos, 1994).

3.5.2.7 Artificial Neural Network Algorithms

Finding the quickest ANN training algorithm for a given problem is very difficult because it is determined by several factors such as the complexity of the problem, the number of weights and biases in the network, the number of data points in the training set, the error objective, and possibility of applying the network for design recognition for function approximation. This section presents four different algorithms that are employed in order to determine the one that trains a confirmed network effectively.

3.5.2.7.1 Levenberg-Marquardt (LM)

The Levenberg-Marquardt (LM) algorithm is essentially the algorithm that is mostly used for updating MLP weights and biases (Hagan and Menhaj, 1994) in neural network. The LM approach is a modification of the classic Newton algorithm which is a simplified form of the Hessian matrix to minimize a problem with finding an optimum solution. The Hessian matrix can be estimated as:

$$H = J^T J \quad \text{Equation (3.26)}$$

and the gradient is calculated as it is shown in following equation

$$g = J^T e \quad \text{Equation (3.27)}$$

in which J is the Jacobian matrix which contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network error. An iteration of this algorithm can be written as

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad \text{Equation (3.28)}$$

Where J is the Jacobian matrix of the performance criteria and x is the weight of the neural network to be minimized, μ is the scalar which controls the learning process, I is the identity matrix and e the residual error vector (Dedecker et al., 2004). When the scalar μ is definitely zero, it is actually Newton's method while using the estimated Hessian matrix. When μ is colossal, the equation becomes a gradient descent with a tiny step size. The mathematical basis of LM is considered to be the most successful paradigm in this study.

3.5.2.7.2 Gradient Descent with Momentum and Adaptive Learning Rate Back Propagation (GDX)

This approach utilizes a common back-propagation algorithm in order to compute derivatives of the performance cost function according to the changeable weight and bias of the network. This technique utilizes gradient descent with momentum for adjusting each variable. For each stage of the modification, if performance reduces, the learning rate is enhanced. This is probably the simplest and most common way to train a network (Haykin, 1999).

3.5.2.7.3 Resilient Back Propagation (RBP)

As a local adaptive learning scheme, the algorithm resilient back propagation (RBP) fulfils interesting batch learning in feed-forward neural networks. The elimination of the detrimental effect of the size of the partial derivative on the weight step is a main principle of RBP. As a result, the sign of the derivative identifies the direction of the weight update. In this method, for each weight individual update value which only indicates the size of the weight update (Riedmiller and Braun, 1993).

3.5.2.7.4 Scaled Conjugate Gradient (SCG)

The quadratic approximation to the error in the neighbourhood of a point is defined by the scaled conjugate gradient (SCG) algorithm (Møller, 1993). SCG is a second-order conjugate gradient algorithm and it helps minimize the target function of several variables. Møller (Møller, 1993) proved this theoretical foundation, which remains a first-order technique in first derivative such as standard back-propagation and discovers the premier way to local minimum in second-order techniques in second derivatives. SCG creates the fast algorithm and employs the step size scaling mechanism which avoids a time-consuming line search per learning iteration (Karmokar et al., 2012). The SCG method presents super-linear convergence in most problems as confirmed by Møller (Møller, 1993).

3.5.2.8 Training and Testing Neural Networks

The training step attempts to provide each neural network with an input data pattern in accordance with a desired output. A numerous examples should be produced

to obtain a best training process. Therefore, as instance problems, more examples are needed. Consequently, poor training process gives rise to an invalid network. The ANN is over trained due to the fact that the output error falls below a particular error threshold, in some instances. Overtraining makes the network modified to learn the training data and may not be capable to forecast the precise new data. Updating the connection weights involving layers which are resulted from the neural network has “learned” the input-output patterns. The best configuration of connection weights is checked in the testing step with another group of independent input data and desired output data is not explicitly defined. It is in the testing step where the output of the neural network models is produced.

The neural networks are able to identify the data sets’ patterns during the training process. Figure 3.20 displays a schematic flowchart of the ANN structure. This is the most reliable and common approach employed for neural networks assessment. Neural network elements to each model were arranged through the preliminary analyses which are examining optimal weight matrix and bias vector consisting of numbers of hidden nodes and value of MSE that were decided at the satisfactory level. The computed output was compared with the actual output, and the MSE was computed. For each of the example case, the whole process was repeated and then back to the first case again, and so on. The cycle is frequentative till the overall error value falls below the predetermined threshold.

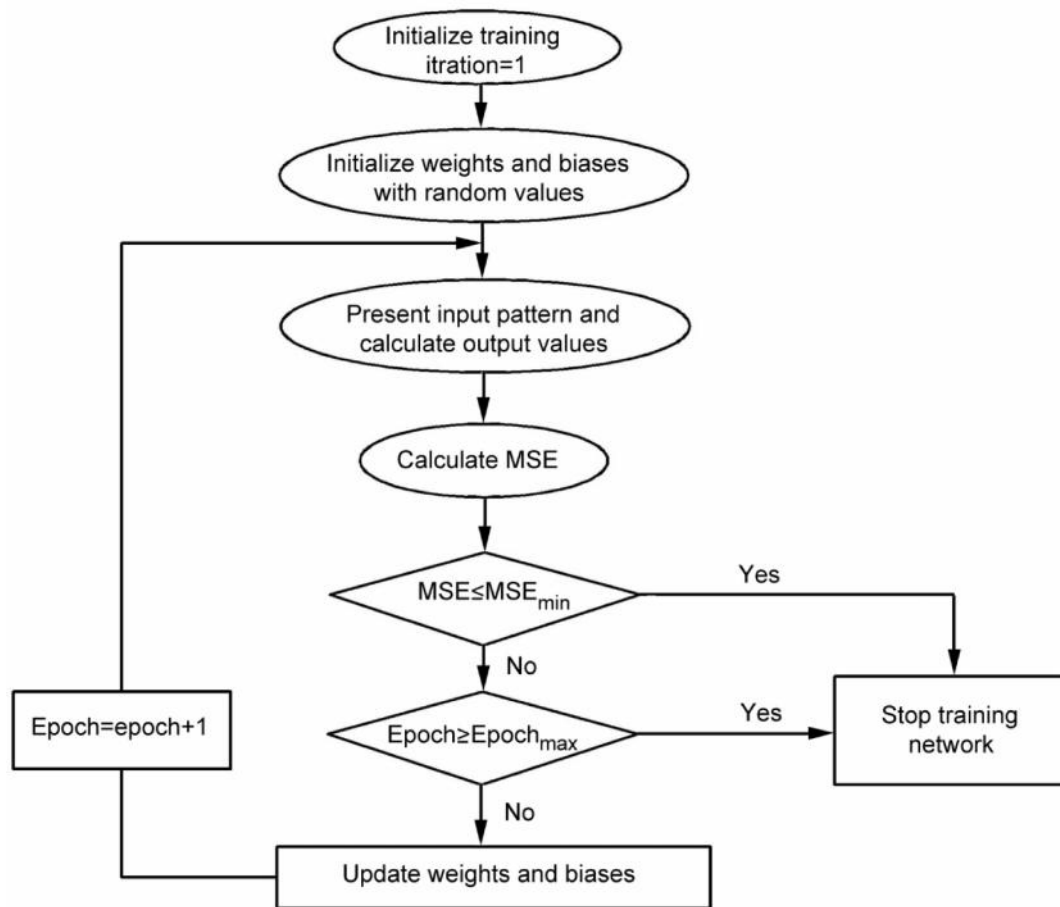


Figure 3.20: A training process flowchart

The summarized ANN development used in this study discussed as follow:

1. The monthly groundwater level, rainfall, evaporation, humidity, minimum and maximum of temperature data were chosen to employ for training and testing of the models, and also the input and output parameters were selected.
2. The selection network architecture as shown in section 3.5.2.5 was configured
3. The number of hidden layer neurons was defined for improvement of neural network performance.

4. By Applying selected performance evaluation statistics such as mean square error (MSE) and correlation coefficient (R), the trained neural network was tested.

5. Item four was repeated until a predetermined predictive performance criterion was satisfied.

3.5.3 Adaptive Neuro Fuzzy Inference System (ANFIS)

Neural network and fuzzy logic as soft computing methods are very complementary. The powerful and efficient solutions have been obtained by combination of these two methods. A neuro fuzzy system is to implement fuzzy inference system (FIS) by ANN construction. It includes the benefits of fuzzy systems, which deal with explicit knowledge that can be described and understood, and neural networks which deal with implied knowledge that can be obtained by learning. A great way to optimize fuzzy rules and membership functions is provided by neural network learning. On the other hand, fuzzy logic enhances the system robustness and the generalization ability to a neural network system by providing more dependable output with limited training data.

Jang (Jang, 1993) introduced a novel architecture and learning product for the FIS, employing a neural network learning algorithm to provide a set of fuzzy If-Then rules with proper membership functions which are obtained by the stipulated input-output pairs (Chang and Chang, 2006). ANFIS is capable to enhance the performance of the FIS effectively, for a provided set of training data and through adopting proper learning algorithms

3.5.3.1 Fuzzy Set and Membership Function

Zadeh (Zadeh, 1965) presented fuzzy sets as an extension of classical (crisp) set theory to represent and manipulate data and information in which there are different alternative uncertainties. It is planned to represent uncertainty and vagueness mathematically and provide official tools to deal with the imprecision intrinsic to several problems. Fuzzy logic that is based on fuzzy set concept and represents a component into a certain degree of membership function was also presented. It usually considered as a real number between 0 and 1 (Zadeh, 1975). Men height is considered as an example for understanding the fuzzy concept. To determine a person (x) is a member or non-member of the set 'tall' (A), the membership function $\mu_A(x)$ of x into A has its values in (ABEM, 2010) and is determined as below:

$\mu_A(x) = 1$ if x is a member of A ($x \in A$) and 0 otherwise ($x \notin A$). That is,

$$\mu_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad \text{Equation (3.29)}$$

Consequently, the boundary of set A is sharp and rigid and generates a two-class group. In classical terms the set is considered crisp and may be determined as a number greater than 180 cm that can be presented as:

$$\text{Tall} = \{x | x \geq 180\} \quad \text{Equation (3.30)}$$

How crisp sets impose a sharp boundary is shown in Figure 3.21(a). However the concept of membership degree is implemented, in the fuzzy set technique. A membership function characterize a fuzzy set A of a universe of discourse X (the range over which the variable spans) $\mu_A(x) \in X \rightarrow [0, 1]$ it associates with each element x of X a number $\mu_A(x)$ in the interval [0, 1], while the membership degree of x in A is

expressed by $\mu_A(x)$. Fuzzy sets with a membership degree that determines fuzzy or blurred boundaries are shown in figure 3.21(b).

The membership function which donates the designation of μ , is the principle of fuzzy sets. A membership function is a curve that determines how every point in the input space is mapped to a membership degree and often considered in the interval and as a real number $[0, 1]$. The experts generally determine membership function selection kind for fuzzy sets or its suitability such as simplicity, convenience, speed, and efficiency resulted in this selection (Mathworks, 2013). Figure 3.22 illustrates an example of membership functions that can be employed for the fuzzy sets. Consequently, creating of a fuzzy set depends on the identification of an appropriate universe of discourse and the specification of a proper membership function (Jang and Sun, 1995).

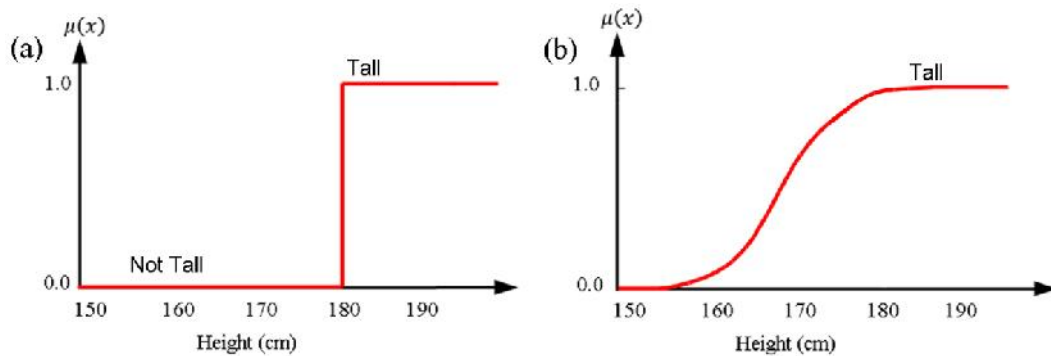


Figure 3.21: (a) Crisp sets; (b) and fuzzy sets of tall man

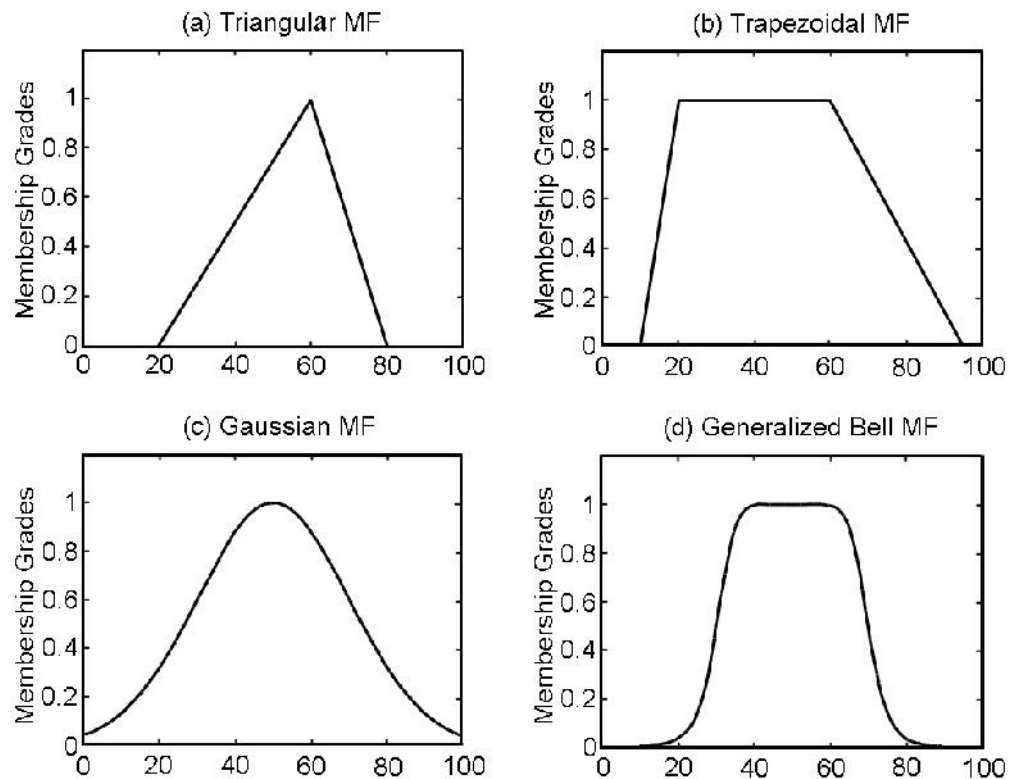


Figure 3.22: Example of membership functions for fuzzy sets

3.5.3.2 Fuzzy Inference System

Fuzzy Inference Systems (FIS) is known as fuzzy rule based systems due to the fact that the systems are made of if-then rules. A given input to output utilizing fuzzy logic is mapped by the FIS. When these systems are used in control systems, they are known as controllers. Figure 3.23 illustrates five functional blocks which are in FIS. In general, the main steps performed in the FIS consist of a) Fuzzification Inference, which transforms crisp inputs to fuzzy. b) Database, is to define membership function. c) Rule base, contains the selection of fuzzy if-then rules. d) Decision Making Unit performs operations on fuzzy rules. e) Defuzzification Interface transforms fuzzy results to crisp values.

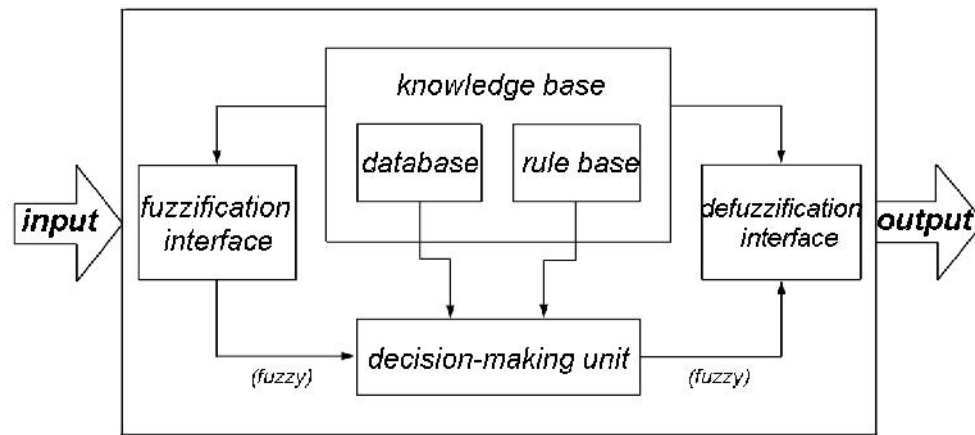


Figure 3.23: Functional block of a fuzzy inference system diagram (Jang, 1993)

The Mamdani approach and the Sugeno approach are two kinds of fuzzy inference that are most commonly employed (Negnevitsky, 2005). The specification of the consequent part is the difference involving these two fuzzy inference methodologies. In the Mamdani approach (Mamdani and Assilian, 1975), consequents are fuzzy sets, and the last crisp output of the Mamdani approach is dependent on defuzzification of the total fuzzy output employing different kinds of defuzzification methods. On the other hand, in the Sugeno approach (Sugeno, 1985), consequents are real numbers and they can be either constant or linear. The final output as a singleton output membership function is the weighted average of each rule's output. The fuzzy rule of the first-order Sugeno-style has the form.

$$\text{If } x \text{ is } A \text{ and } y \text{ is } B, \text{ Then } z = ax + by + c \quad \text{Equation (3.31)}$$

where A and B are fuzzy sets in the antecedent, while $ax + by + c$ is a mathematical function in the linear consequent. When $(a = b = 0)$, the output degree of z has a constant value, which is called zero-order Sugeno model.

3.5.3.3 ANFIS Architecture and Training

ANFIS combines FIS with the back-propagation neural network algorithm determined by some input and output data collection. Employing learning capability of neural network is the most significant basis for combining fuzzy systems with neural networks. These combinations must be capable to learn linguistic rules and/or membership functions, or optimize the existing ones. This study presents successful adaptive neural network fuzzy inference systems, along with training algorithms known as ANFIS to illustrate the application of neural networks for fuzzy inference. The fuzzy inference system under consideration is assumed to have two inputs, x and y , and one output, f , for a first-order Sugeno fuzzy model; a common rule set with two fuzzy If-Then rules may be expressed as

$$\text{Rule 1: If } x \text{ is } A_1 \text{ and } y \text{ is } B_1, \text{ Then } f_1 = p_1x + q_1y + r_1 \quad \text{Equation (3.32)}$$

$$\text{Rule 2: If } x \text{ is } A_2 \text{ and } y \text{ is } B_2, \text{ Then } f_2 = p_2x + q_2y + r_2 \quad \text{Equation (3.33)}$$

where for inputs x and y , membership functions (MFs) are A_1 , A_2 and B_1 , B_2 respectively; and p_1, q_1, r_1 and p_2, q_2, r_2 are the linear parameters of the output function. Figure 3.24 shows the resulting Sugeno fuzzy reasoning system and equivalent ANFIS architecture. Nodes at the same layer have the same function for this ANFIS structure. It is important to note that the structure of this adaptive network is not unique. For instance, Layers 3 and 4 can be combined to obtain an equivalent with only 4 layers. By the same notion, the weight normalization can be performed in the last layer. $O_{l,i}$ specifies the output of the i th node in layer l . The five layers comprising the ANFIS structure are described as follows:

Layer 1: Input nodes. Membership grades based on the appropriate fuzzy set have been produced by each node of the layer they belong to using membership functions

$$O_{1,i} = \mu A_i(x) \quad \text{for } i = 1, 2 \quad \text{Equation (3.34)}$$

$$O_{1,i} = \mu B_{i-2}(y) \quad \text{for } i = 3, 4 \quad \text{Equation (3.35)}$$

where x, y are the input to the i th node and A_i, B_{i-2} is a fuzzy label (low, high and etc.) characterized by the appropriate membership functions $\mu A_i, \mu B_i$, respectively, which can be triangular, trapezoidal, Gaussian functions or other shapes. The generalized bell function (Eq. 3.36) and Gaussian membership function (Eq. 3.37) generally explain the membership functions for A and B

$$\mu A_i(x) = \frac{1}{1 + [x - c_i/a_i]^{2b_i}} \quad \text{Equation (3.36)}$$

$$\mu A_i(x) = \exp[-(x - c_i/a_i)^2] \quad \text{Equation (3.37)}$$

where (a_i, b_i, c_i) is the parameter set of the membership functions in the premise part of the fuzzy If-Then rules then the shapes of the membership function are variable. Genuinely, the qualified candidates for node functions in this layer can be any continuous and piecewise differentiable functions, such as commonly used triangular-shaped membership ones (Jang, 1993). Parameters in this layer are referred to as premise parameters.

Layer 2: Rule nodes. In the second layer, every node is a fixed node labelled , where the incoming signals and output product are multiplied. For example

$$O_{2,i} = W_i = \mu A_i(x) \mu B_i(y), \quad i = 1, 2 \quad \text{Equation (3.38)}$$

the firing strength of a rule is shown by the output of each node.

Layer 3: Average nodes. Each node in the third layer denotes N which is a stable node. The main aim of this layer is to compute the ratio of each i th rules firing strength to the sum of all rules firing strength

$$O_{3,i} = \bar{W}_i = \frac{W_i}{v_1 + W_2}, \quad i = 1, 2 \quad \text{Equation (3.39)}$$

The results of this layer are named as normalized firing strengths.

Layer 4: Consequent nodes. In this layer each node presents an adaptive node, with the following node function

$$O_{4,i} = \bar{W}_i f_i = \bar{W}_i (p_i x + q_i y + r_i) \quad \text{Equation (3.40)}$$

where \bar{W}_i is the i th nodes output from the previous layer, and (p_i, q_i, r_i) is the parameter set. The parameters of this layer are referred to as consequence or output parameters.

Layer 5: Output nodes. The overall output summing of all the incoming signals is computed by the single node. Therefore, the defuzzification process transforms each rule's fuzzy results into a crisp output in this layer

$$\text{Overall output} = O_{5,i} = \sum_{i=1} \bar{W}_i f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \quad \text{Equation (3.41)}$$

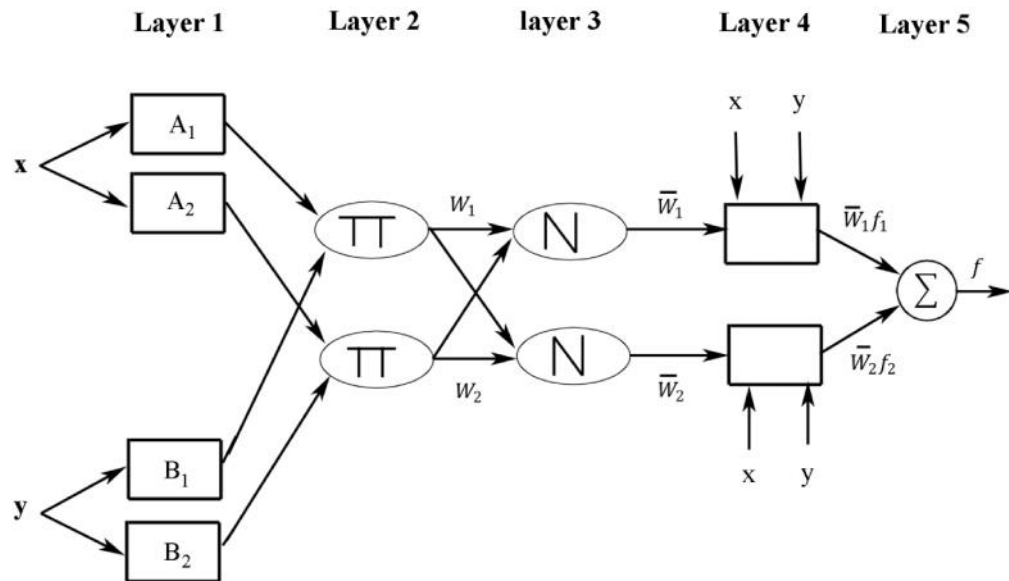
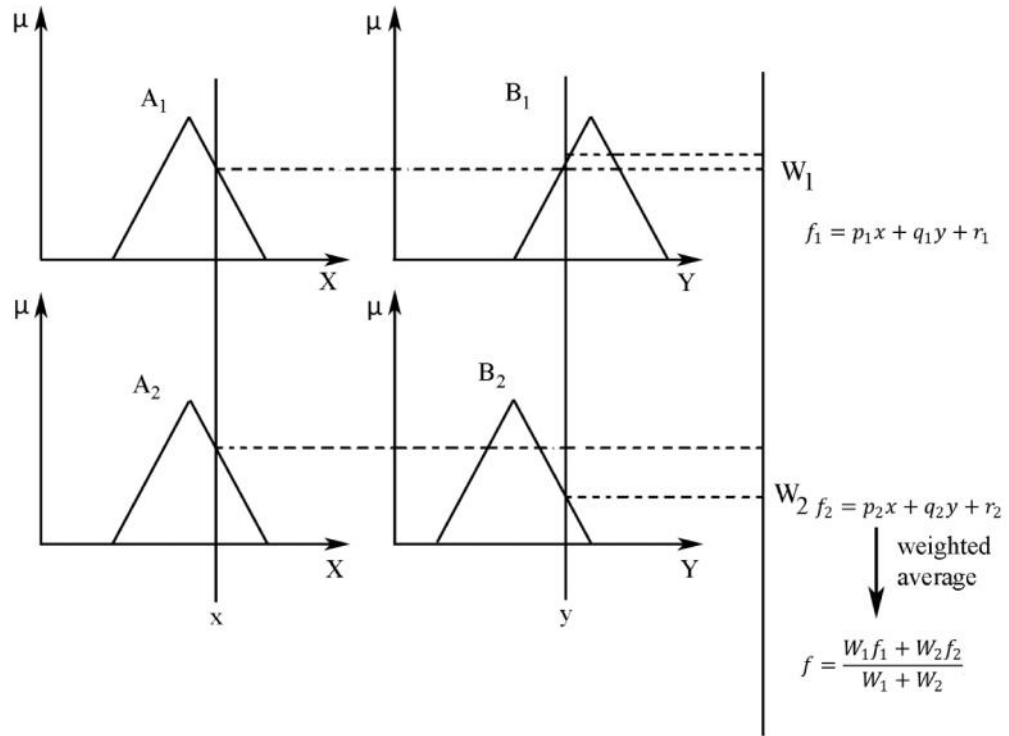


Figure 3.24: Two inputs first-order Sugeno fuzzy model with two rules and architecture of ANFIS.

3.5.3.4 Learning Algorithm of ANFIS

Utilizing an understanding algorithm, applying input and output information units, are resulted in adaptable parameters of the FIS which is the purpose of ANFIS. The duty of the understanding algorithm for this architecture would be to tune every adaptable variable, videlicet (a_i, b_i, c_i) and (p_i, q_i, r_i) , and to create the ANFIS result to match the training data. The result of the ANFIS type can be mentioned as Eq. (3.41) when the assumption variables a_i, b_i and c_i of the membership function are axed

$$\begin{aligned}
 f &= \frac{W_1}{W_1 + W_2} f_1 + \frac{W_2}{W_1 + W_2} f_2 = \overline{W}_1 f_1 + \overline{W}_2 f_2 \\
 &= \overline{W}_1 (p_1 x + q_1 y + r_1) + \overline{W}_2 (p_2 x + q_2 y + r_2) \\
 &= (\overline{W}_1 x) p_1 + (\overline{W}_1 y) q_1 + (\overline{W}_1) r_1 + (\overline{W}_2 x) p_2 + (\overline{W}_2 y) q_2 + (\overline{W}_2) r_2 \quad \text{Equation (3.42)}
 \end{aligned}$$

that is a linear mix of the adaptable consequent variables p_1, q_1, r_1, p_2, q_2 and r_2 . The determination of suitable values of these variables can be effortlessly carried out by the least squares technique. These supposition variables within the fuzzy level and the consequent variables within the defuzzification level were tuned during the ANFIS learning process until the preferred result of the FIS is obtained (Shahbudin et al., 2009). In this study a hybrid learning algorithm was utilized to achieve the best values of the FIS variables of the Sugeno-type. A combination of the least squares and the back-propagation gradient descent approach for training FIS membership function variables to imitate a given training data set is applicable. It has been proven that the hybrid algorithm within the training ANFIS is broadly effective (Jang, 1992, 1993). A schematic flowchart of the ANFIS structure is presented in Figure 3.25.

The structure of water level and quality variables predicting system determined by ANFIS and ANN are shown in figure 3.26. The model predicting process consists of

data acquisition, data standardization pre-processing, predicting, testing results, and application.

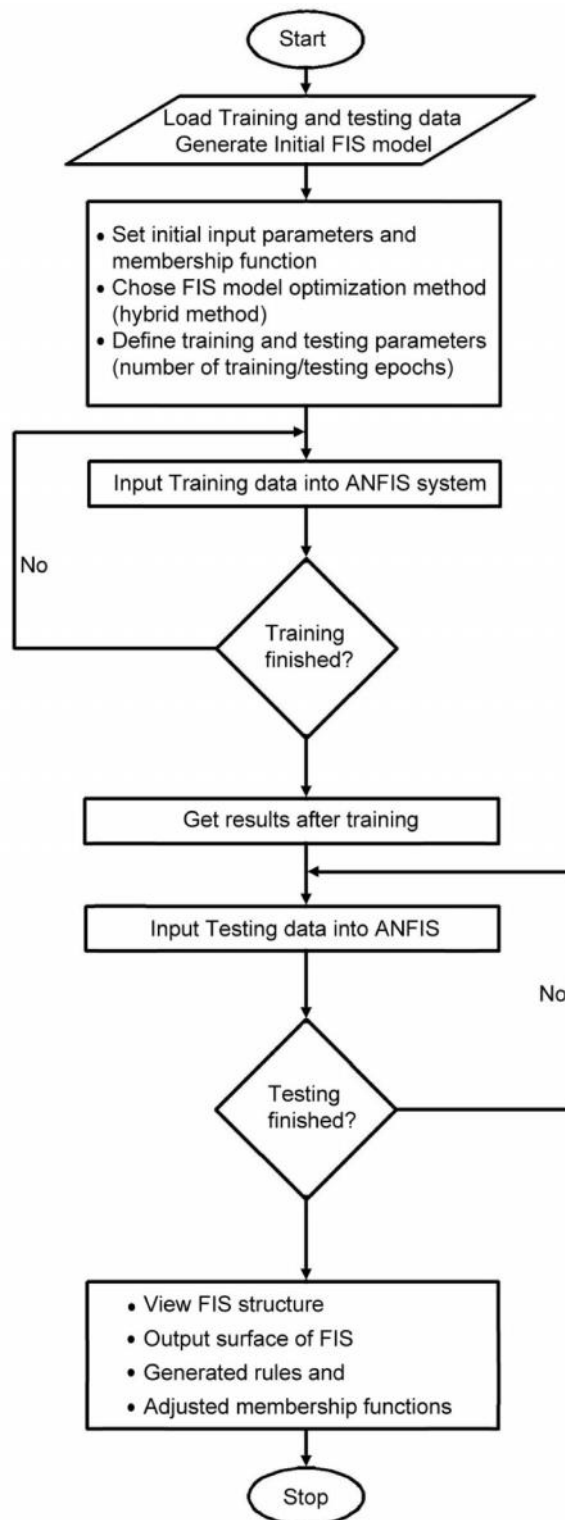


Figure 3.25: Flowchart of ANFIS system

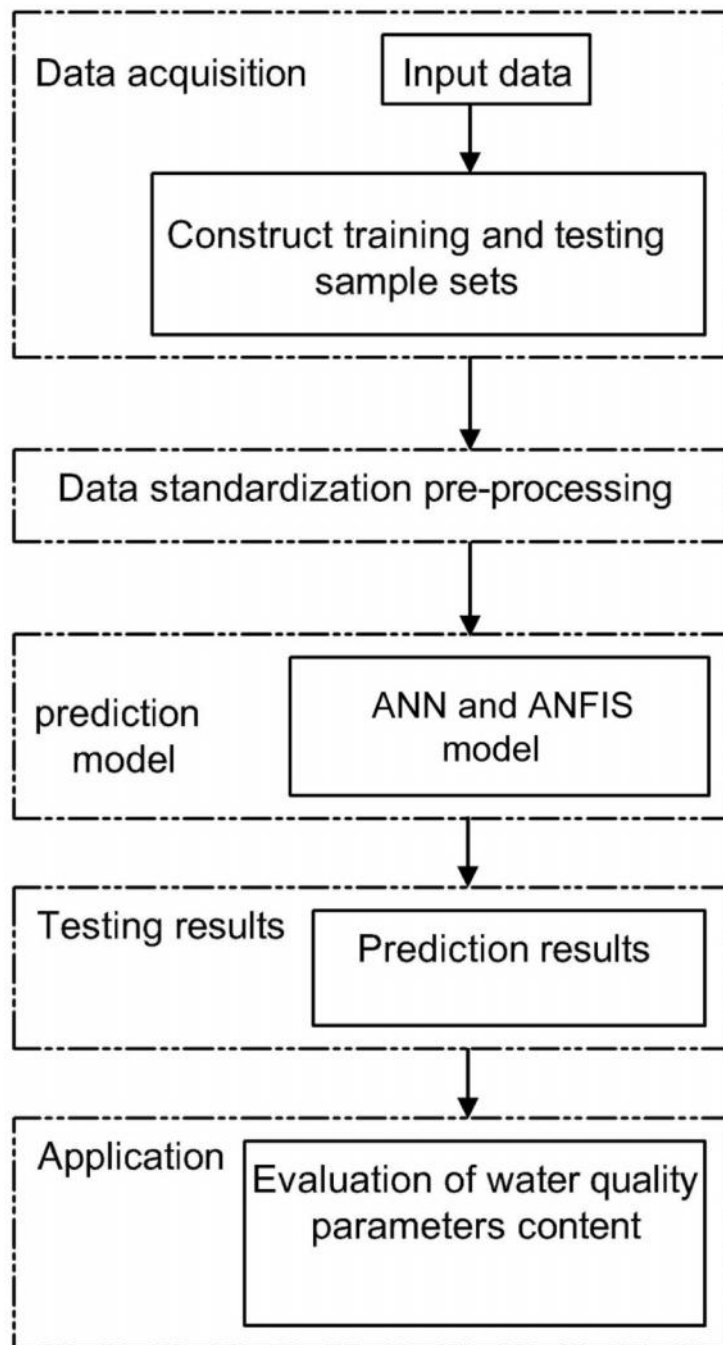


Figure 3.26: Methodology flowchart for prediction model

3.6 Evaluation of the Model

3.6.1 Data Pre-Processing

The input vector consisting of the raw data of numerous water quality factors has various units which should be normalized before being displayed to the network. Water quality parameters accompanying with various magnitudes may play unequal role for the weight update procedure, and hence it improves the performance of the model.

In order to remove dimension differences, the following equation was employed to standardize and normalize the data, and as a result, all input and output data were standardized and normalized to the range

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad \text{Equation (3.43)}$$

3.6.2 Performance Criteria

Two different criteria are applied to investigate the effectiveness and the precision of the simulation ability of each network. First, the mean square error (MSE) is calculated by

$$\text{MSE} = \frac{\sum_{i=1}^N (y_i - \bar{y}_i)^2}{N} \quad \text{Equation (3.44)}$$

Where N is the number of observation data, and y and \bar{y} are the observed and computed data respectively. The MSE reflects the difference between the observed and computed values; the lower the MSE results are, the more precise the simulation.

The correlation coefficient (R) between the network result and the network target outputs in three training, testing, and validation groups was the second to be employed and calculated as

$$R = \sqrt{1 - \frac{\sum (y_i - \bar{y}_i)^2}{\sum y_i^2 - \frac{(\sum y_i)^2}{N}}} \quad \text{Equation (3.45)}$$

Where R represents the percentage of the preliminary uncertainty explained by the model. The best fit between the observed and computed values, which is unlikely to happen, would have been MSE=0 and R=1.

Chapter 4: Geophysical and Hydrogeochemical Results and Discussion

4.1 Introduction

It was achievable to identify the geological formations and groundwater quality, through background knowledge of local geology, geophysical and hydrogeochemical surveys, in this study. Regarding to the aim and methodology in the area, the result mainly focused on two parts. The study of geoelectrical resistivity is the first part of the discussion. Electrical resistivity surveys were used to map the electrical properties of the subsurface structure particularly the detection of bedrock and the mapping of the aquifer. Hydrogeochemical investigation of groundwater quality and monitoring of the water level in the research area are discussed in the second part.

To investigate the groundwater characteristic a geoelectrical resistivity and hydrogeochemical analyses were combined and applied in this study. Groundwater samples were obtained from the existing wells. In-situ parameters such as well location, well depth, water level, pH, conductivity, salinity, total dissolved solids, and temperature were measured at each site. For determination of their major ion contents, the chemical analyses in the laboratory were conducted. Total dissolved solid and major ion concentration were the hydrogeochemical parameters applied in the study of the groundwater. Moreover, monitoring of groundwater level was carried out in the research area. For investigation of the characteristics of the subsurface and the quality of groundwater, a geoelectrical resistivity and hydrogeochemical analyses were combined and applied in this research. 2D geoelectrical resistivity surveys with Wenner array with 61 electrodes configuration were applied in research area.

The Wenner array is good at showing vertical changes. This array has strongest signal strength compared to the other arrays. This can be an important factor if the

survey is carried out in the areas with high background noise. The maximum profile length was 400 m and the minimum length was 40 m, while each profile spread location depended on the accessible space and goal in the field.

4.2 Geophysical Study

4.2.1 Geoelectrical Resistivity Result

In the groundwater prospecting qualification, two important parameters in the formation of geological rock are the depths of the bedrock and aquifer. The geophysical study was performed along the Langat River starting from northeast to southwest of the study area (Figure 4.1). The ERI method provided lithology away from existing boreholes and quick indication of aquifer geometry. The acquired, standardized and calibrated data in the previous chapter (3.3.2) are resulted in interpretation of the electrical resistivity. The significant precision of the correlation is necessary between the depths identified from the ERI results and lithological logs. For the ERI models, the investigation and discussion are commenced from close to the surface and to the deepest depth subsequently. A maximum depth of 70 m is found for the resistivity surveys. The eighty eight ERI surveys include measurements along several profiles placed along and in a perpendicular direction to the Langat River. Figure 4.1 illustrates the central point of the geoelectrical surveys.

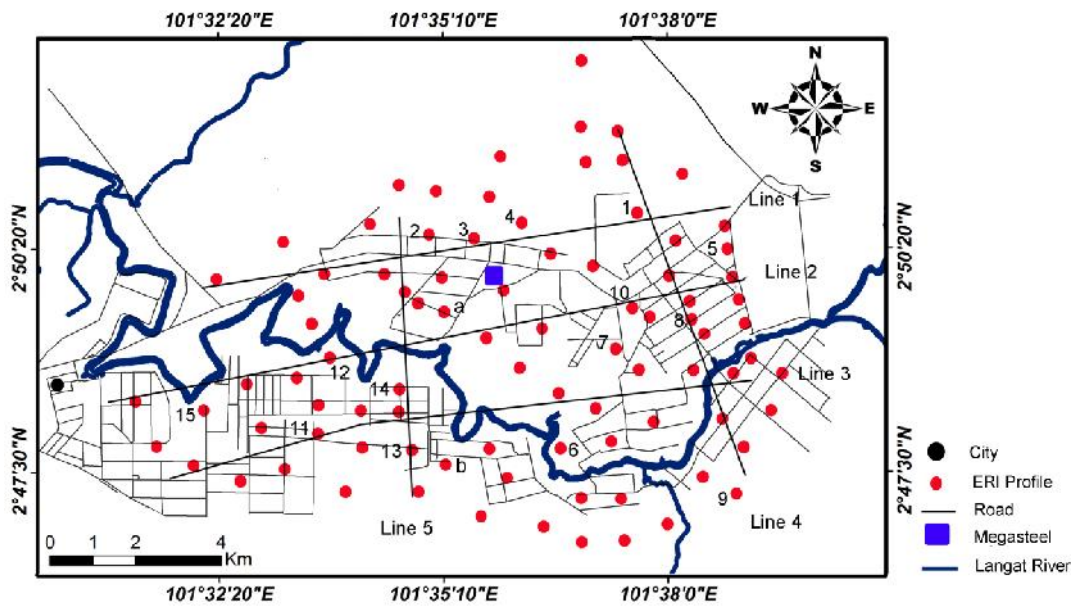


Figure 4.1: Location map of geoelectrical resistivity survey in the study area

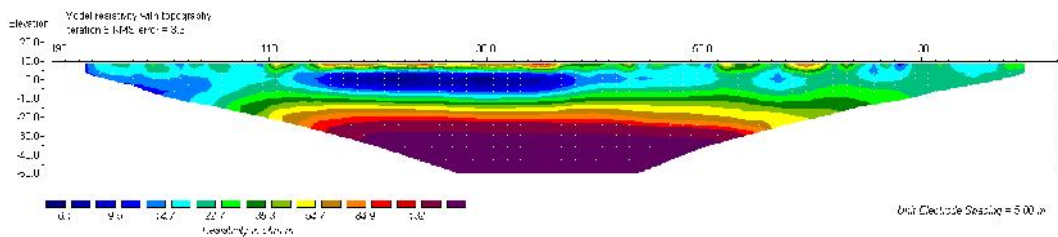
4.2.2 2D Electrical Resistivity Imaging Sections

The results of the interpretation of the ERI profiles determine the subsurface section, and include various geoelectrical units, which belong to Quaternary deposits comprising silt, clay, sand, and gravel. The main goal of these surveys was to define the aquifer in the site and to delineate the depth of the bedrock. The selected geoelectrical resistivity profile is interpreted in following section.

Profile 1

Profile 1 was situated close to the borehole no 17 with the elevation of 10 masl at northeast part of the study area. The profile generated three various resistivity of layered rock which is obviously determined by their various resistivity layers at the different depths. Based on the data, the clay layer thickness is approximately 15 meter with the resistivity of less than 20 ohm.m which is clearly illustrated by a resistivity

pseudosection in Figure 4.2. Comparatively high and low resistivity values take place in the top corresponding to more compacted material alternating with softer material. An intermediate resistivity layer at the intermediate depth zone illustrate weathered to moderately weathered rock material, is discovered afterwards. A comparatively moderate resistivity in the range of 22 to 132 ohm.m is related to this depth layer ranging from 15 to 30 m. The higher resistivity values of 132 ohm.m corresponds to the bedrock situated at the depth 30 to downwards.



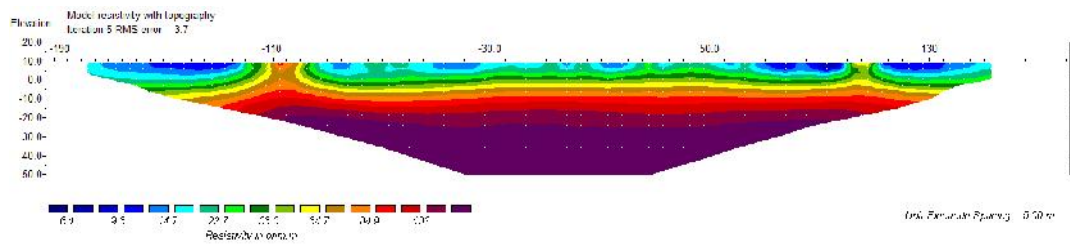


Figure 4.3: Selected 2D resistivity model of the apparent resistivity data profile 2

the existence of metasediment basement. Moreover, drilling information that informed metasediment bedrock in a depth of 30 m down supports this discussion.

Profile 3

The location of profile 3 is at 1 km east of the profile 2 with the elevation of 10 masl. Figure 4.4 displays this profile inversion image. The effective depth of investigation was found to be around 65 m that is based on the 5 electrode distance with total length of 400 m. Highly resistive materials with resistivity of higher than 132 ohm.m demonstrated as bedrock were detected to an approximate depth of 25 m. A well-defined boundary is evident between depth of 20 m and 25 m, with resistivity values decreasing from 132 ohm.m to less than 90 ohm.m in that interval.

The aquifer consisting of sand and gravel with the depth of 10 to 25 m and having the resistivity value less than 132 ohm.m corresponds to this layer. Figure 4.5 illustrated resistivity value of 6-22 ohm.m is observed closely to surface in the geoelectrical model. This layer with the depth of surface of around 10 m in the ground correspond the clay layer.

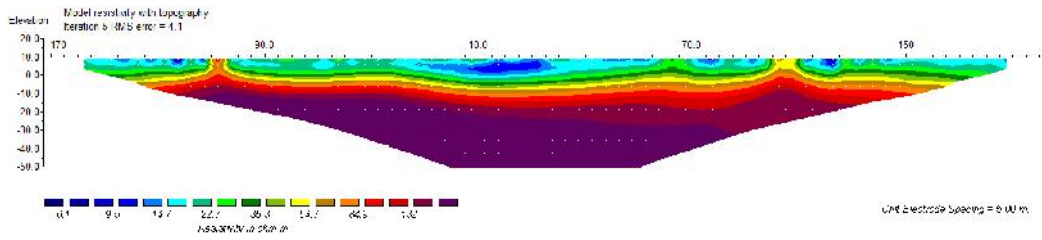


Figure 4.4: Selected 2D resistivity model of the apparent resistivity data profile 3

Profile 4

Profile 4 with the elevation of 9 masl was located at the northeast of the study area. The inversion model was resulted from this profile as illustrated in figure 4.5. The inversion process converged with a RMS misfit of 4.2 after five iterations for this profile. An electrode distance of 5 m with total of 400 m was employed with a Wenner array. Resistivity values of the model were set as done in the past model. The resistivity value of 20-35 ohm.m is observed close to the surface corresponding to more compact material alternating which is extended from the ground to depth of 10 m. This layer corresponding clay contains the fresh water. A resistivity value of 50-132 ohm.m is found from depth of 10 to 30 m and is related to the sand and gravel layer (aquifer) while, the higher resistivity values of 132 ohm.m with the depth of more than 30 m corresponds to the bedrock.

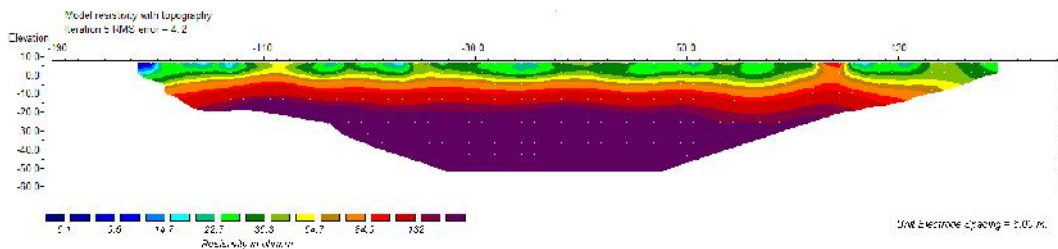


Figure 4.5: Selected 2D resistivity model of the apparent resistivity data profile 4

Profile 5

The geoelectrical survey for profile 5 with the elevation of 12 masl was carried out in the direction of north-south of the study area. Figure 4.6 displays the resistivity model of this profile. In general, the profile represents low values of electrical resistivity rock materials near the surface. The profile generated three various resistivity of the layered rock determined by their different resistivity layers at the various depths which are; at first, a thin subsurface layer with relatively low resistivity values that is lower 22 ohm.m corresponds to the depth of about 10 m for the wet clay material. An intermediate resistivity layer at the intermediate depth zone that might represent aquifer layer and contains sand and gravel was observed afterwards. The second layer with a relatively moderate resistivity in the range of 22 to 130 ohm.m was found for this depth layer between the depth ranges of 10 to 30 m. The nearest well to the survey profile is the well 6 and 17 in which fresh water characteristics are indicated by the chemical result (section 4.3.1). The third layer with a relatively high resistivity values greater than 132 ohm.m as a metasediment basement started from the depth of 30 m below the ground and is also found below the two layers of low to intermediate rock resistivity. In this profile the rock layers show a wavy pattern and are highly sheared.

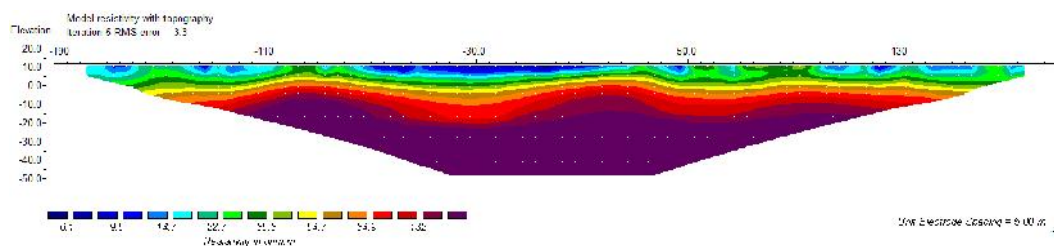


Figure 4.6: Selected 2D resistivity model of the apparent resistivity data profile 5

Profile 6

The resistivity survey of profile 6 presents the approximate trend in the direction of the north-south of the study area with the elevation of 10 masl. A lower resistivity zone (4 to 22 ohm.m) from the ground surface, extending to a depth of 15 m has been shown by the pseudosection model of this profile given in Figure 4.7. The variations of resistivity and thickness in this zone reflect the concentration of clay and water in the topsoil materials. The region with higher resistivity of greater than 22 ohm.m is situated under the clay layer and refers to the gravel and sand layer (aquifer) and zones of more than 132 ohm.m corresponded to the bedrock surface which is at the depth of 40 m. In the top relatively high resistivity values have happened that are related to compacted material alternating with softer materials. The obtained results from resistivity profiling have been confirmed by data from the closest drilled borehole besides the field and field observations.

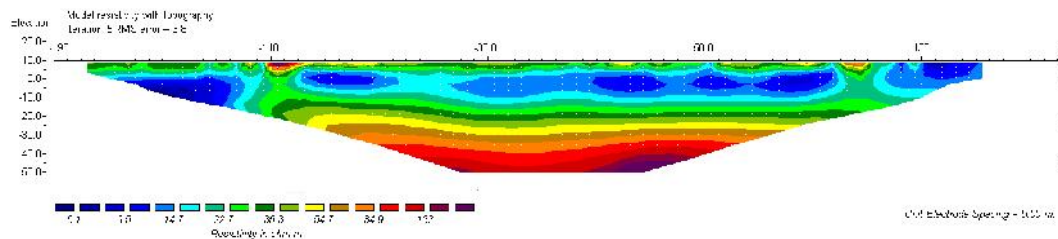


Figure 4.7: Selected 2D resistivity model of the apparent resistivity data profile 6

Profile 7

Profile 7 trend in direction of the north-south was situated in the central part of the study area and the elevation of 9 masl. The inverted resistivity section illustrates similar subsurface layers as explained in previous profiles, which have been shown in Figure

4.8. The resistivity ranges from 6 to 22.7 ohm.m corresponding to the topsoil layer, and might be maximized to the depth of 10 m which is related to the clay layer, throughout this section. Based on the inversion results from this profile, there is a fairly uniform sand and gravel layer with a thickness of 30 m that is located in the depth between 10-40 m and resistivity values of less than 132 ohm.m. The zones of more than 132 ohm.m correspond to the bedrock surface, which is situated at a depth of 40 m. Therefore the results of the 2D electrical resistivity imagine and the drilling results represented a good correlation.

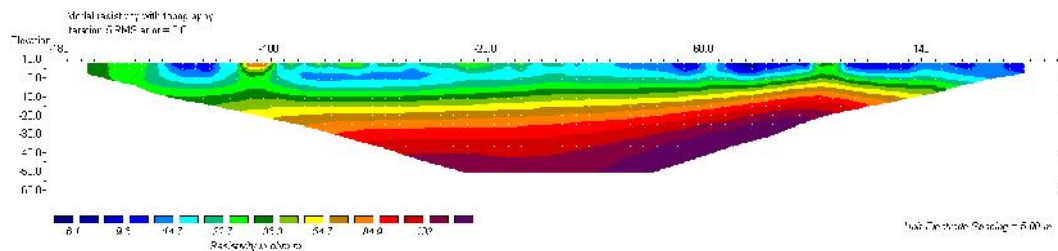


Figure 4.8: Selected 2D resistivity model of the apparent resistivity data profile 7

Profile 8

The survey line 8 was performed at the east of the study area which is located 400 m at the north of borehole no 6. A palm oil trees plantation with the elevation of 10 masl surrounded this site and there was a puddle of water in the drainage system in 1.5 m below the ground surface. Figure 4.9 illustrates the appearance of a lower resistivity value close to the surface up to 10 m which is related to the clay material with the moisture content. The zone between 22 and 132 ohm.m is considered as the aquifer which consists of sand and gravel with the depth of 10 to 30 m. Characteristics of fresh water close this profile have obtained by the chemical results of well no 6 (Table 4.3). A

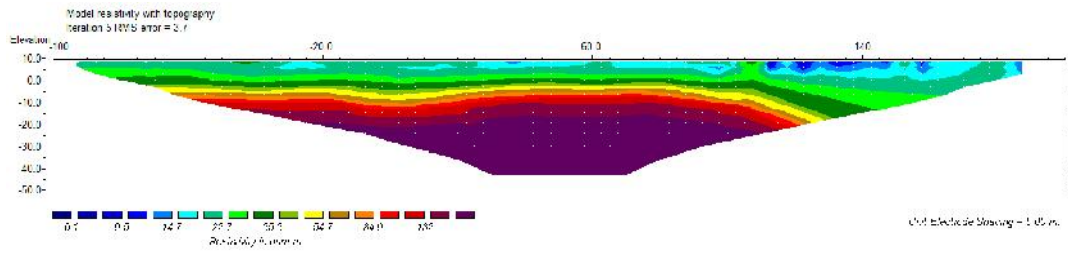


Figure 4.9: Selected 2D resistivity model of the apparent resistivity data profile 8

higher resistivity value of more than 132 ohm.m has been found at the depth of 30 downward which is related to the metasediment basement.

Profile 9

The next geoelectrical resistivity survey (profile 9) was located at the southeast of the study area. The site survey has the elevation of 10 masl. in the geoelectrical inversion model (Figure 4.10). The profile generally shows the electrical resistivity of the layers increasing with depth. The profile created three various resistivity of layered rock that is obviously described by their various resistivity layers at different depths.

Clay is the first layer with resistivity values less than 22 ohm.m, and it extends to the depth of 10 m. The sand with gravel layer represents the lower zone, as identified by its higher resistivity, and presents a depth of around 10 to 30 m. Based on the nearest borehole information (well no 7), this second layer corresponds to sand with gravel and is the water-bearing zone (aquifer).

The third layer with a relatively high resistivity is also observed below the two layers. This comparatively high resistivity rock layer (more than 132 ohm.m) corresponds to the bedrock in the depth of 30 m. The results of the 2D resistivity model and the drilling borehole results therefore showed a good correlation.

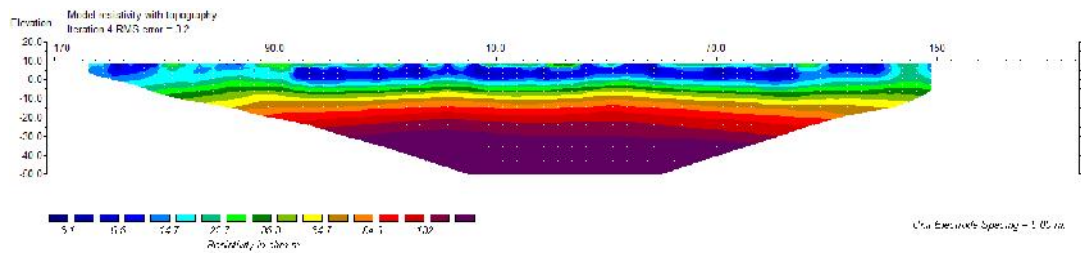


Figure 4.10: Selected 2D resistivity model of the apparent resistivity data profile 9

Profile 10

The profile 10 was located at a site with the elevation of 11 masl. The survey profile was laid in the direction of east to west. Figure 4.11 displays the geoelectrical model along profile 10. The investigation depth of the 2D images can be approximated into a three-layer model in which the model resistivity of the layers increases with depth. The resistivities lower than 22 ohm.m indicated by a blue region are investigated as the clay layer. In the section model, the region with the depth of 0 to 10 m corresponds to this layer. The aquifer (Simpang Formation) is the second layer and it is identified in the region by resistivities of 22 to 132 ohm.m. The inversion results obtained from this profile define fairly uniform sand and gravel (aquifer) is the second layer with the depth of approximately 10 to 40 m and the resistivity values of less than 132 ohm.m. Moreover, interbedded thin clay layers have been found in the sand and gravel layers. The zones of more than 132 ohm.m correspond to the bedrock surface is situated at a depth of 40 m.

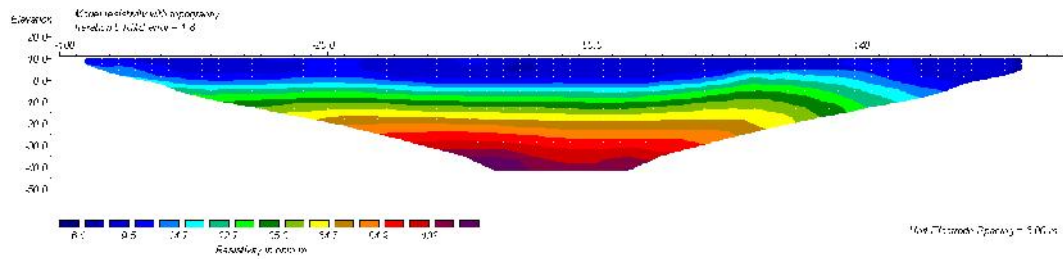


Figure 4.11: Selected 2D resistivity model of the apparent resistivity data profile 10

Profile 11

One of the geoelectrical surveys near to the coastal area is profile 11 which is around 13 km away from the beach and located at the west of the study area with the elevation of 11 masl. This profile is situated close to the borehole no 2 (200 m at the south). As a result of intrusion sea water, this site has low resistivity value. Inversion model of profile 11 was presented in Figure 4.12. Boundaries between the Kenny Hill Formation (bedrock), Simpang Formation (aquifer) and clay were defined based on the data obtained from borehole no 2. In addition, the depth of the bedrock is about 65 m. The bedrock resistivity value around 28 ohm.m is found to match resistivity values with approximate depths of 65 m. The blue region with the resistivity lower than 4.2 ohm.m is considered as the clay layer. The regions with the depth of 0 to 21.5 m are related to this layer, in the section model. The second layer was the aquifer (Simpang Formation) which was identified by resistivity in the region in the range of 4.2 to 28 ohm.m. Furthermore, thin embedded clay layers were found in sand and gravel layers. Based on the obtained information from chemical analyses of borehole no 2 (Table 4.3), the aquifer with brackish water was identified as an area with low resistivity values such as 10 ohm.m. Low resistivity values indicate excessive values of TDS and presence of marine clays of Gula Formation.

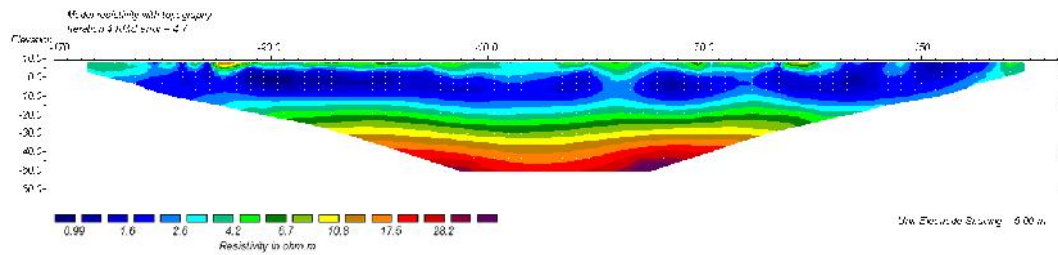


Figure 4.12: Selected 2D resistivity model of the apparent resistivity data profile 11

Profile 12

Profile 12 was another selected profile with the elevation of 11 masl which was influenced by the intrusion of sea water. The geoelectrical model of profile 12 was displayed in Figure 4.13. A lower resistivity value around less than 10 ohm.m occurred close to the surface from the surface to the depth of 10 m. This value is related to the clay layer with moisture content. After the clay layer, an average resistivity value of 15 ohm.m was appeared which was related to the sand with gravel (aquifer) layer that was obtained at the depth of 10 to 40 m. The nearest well was situated about 500 m from the survey profile to the west that confirmed the results. A resistivity value of more than 28 ohm.m reveals at the depth of 40 m downward which refers to the metasediment basement.

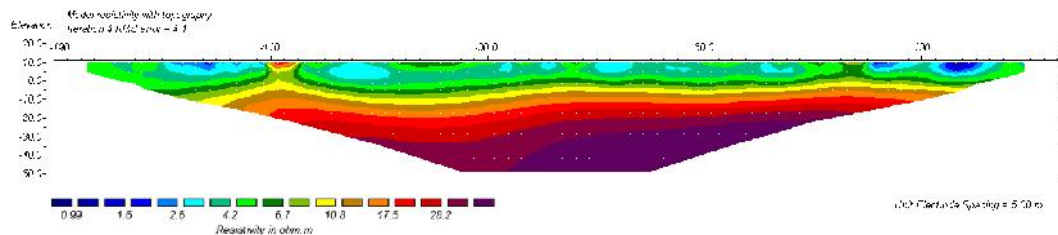


Figure 4.13: Selected 2D resistivity model of the apparent resistivity data profile 12

Profile 13

Profile 13 which is similar to profile 9 and 10 was influenced by sea water intrusion from the west side of the study area with the elevation of 10 masl. The resistivity inversion model of profile 13 has been shown in Figure 4.14. The investigation depth of the 2D model can be approximated into a three-layer model in which the model resistivity of the layers raises with depth. The resistivities lower than 4 ohm.m indicated by the blue region, are investigated as the clay layer. In the section model, the region with the depth of 0 to 15 m corresponds to this layer. The aquifer (Simpang Formation) as a second layer was identified in the region by resistivities ranging from 4 to 28 ohm.m at the depth from 15-45 m. In addition, interbedded thin clay layers have been found in the sand and gravel layer. Based on the data gained from borehole, the aquifer with brackish water was determined as areas with low resistivity values like 10 ohm.m; the low resistivity values reflect the high values of TDS and presence of marine clays of Gula Formation. The Simpang Formation consists of gravel and sand saturated with groundwater illustrates moderate resistivity values which determine the main aquifer within the study area. Relatively higher resistivity value more than 28 ohm.m is obtained at the depth of 45 m downward which corresponds to the metasediment basement.

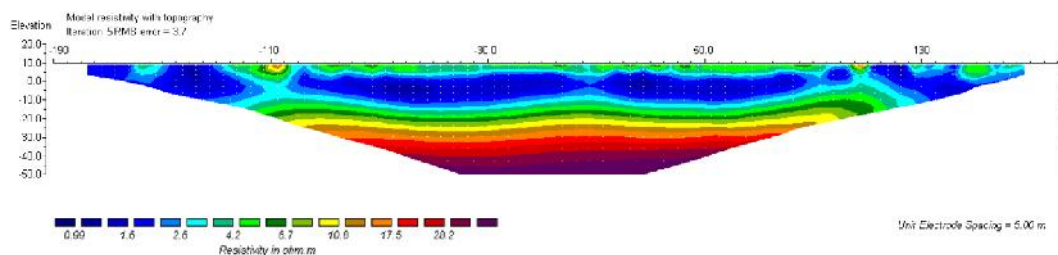


Figure 4.14: Selected 2D resistivity model of the apparent resistivity data profile 13

Profile 14

The profile 14 with the elevation of 9 masl was located around 1 km from well 1 station at the south of the Langat River with the survey line direction of the north to south. Figure 4.15 illustrates the inverted resistivity model. Using the data obtained from the borehole as a reference identify the boundary between the Kenny Hill Formation (bedrock), the Simpang Formation (aquifer), and the Gula Formation (clay). Furthermore, the bedrock depth is approximately 60 m. Matching the resistivity value with the approximate depth is resulted in determining a bedrock resistivity value of about 28 ohm.m. The resistivity values of lower than 4.5 ohm.m, indicated by the blue region, are considered as the indicative of the clay layer. This layer is found at the depth of about 21.5 m at the borehole position. The aquifer (Simpang Formation) is defined in the region of 4.5 to 28 ohm.m with the ranging depth of 21 to 60 m.

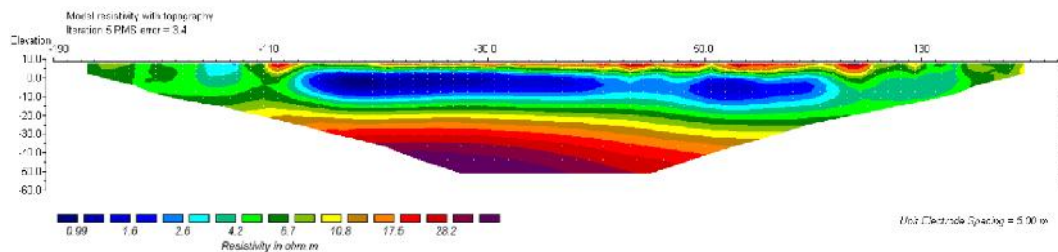


Figure 4.15: Selected 2D resistivity model of the apparent resistivity data profile 14

Profile 15

The profile 15, with the elevation of 10 masl was placed at the western part of the research area. Figure 4.16 illustrates the inverted resistivity model for this profile. The region is underlain by low resistivity rock materials which are layered as the profile illustrated. Moreover the sharp variation in electrical resistivity has clearly interpreted

the boundaries of the different resistivity material. The observed low resistivities are because of presence of marine clays layer. An overburden with low resistivity value less than 4 ohm.m was found at the depth of 20 m. That might be described as a clay layer. An intermediate depth zone shows comparatively high resistivity distribution when it was compared to first layer and it corresponds to the resistivity distribution on the profile ranging from 4 to 28 ohm.m which increases with depth. This intermediate depth layer shows sharp variation in resistivity values with the depth which corresponds to the sand and gravel (aquifer) layer. A resistivity value greater than 28 ohm.m was obtained from depth of greater than 60 m that corresponds to the bedrock.

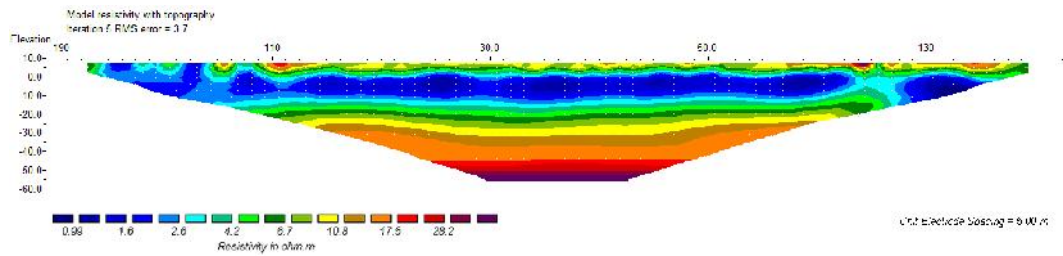


Figure 4.16: Selected 2D resistivity model of the apparent resistivity data profile 15

4.2.3 Geoelectrical Maps

True resistivities and thickness model of an aquifer layer was prepared based on the obtained results from geoelectrical measurements. Figure 4.1 shows the location of geoelectrical measurements in the study area. Figures 4.17 and 4.18 respectively illustrate the true resistivities and thicknesses defined by the 2D resistivity models which have been interpolated to map the heterogeneities of the expected sand and gravel (aquifer) layer. To smoothly interpolate the measured data, the powerful geostatistical kriging approach has been applied. The thickness distribution map in Figure 4.18

indicates that this area is characterized by a significant aquifer with the thickness of 40 m greater than the aquifer layer with shallower depths which was found in the western part of the study area. The low resistivity values (less than 10 ohm.m) in the western margin are attributed to the presence of saturated sand with saline groundwater. Therefore, the appearance of saline water due to the seawater intrusion has affected the western areas. A low topography and a deeper depth of the sand with gravel layer that is found in the western part of the area have characterized these areas.

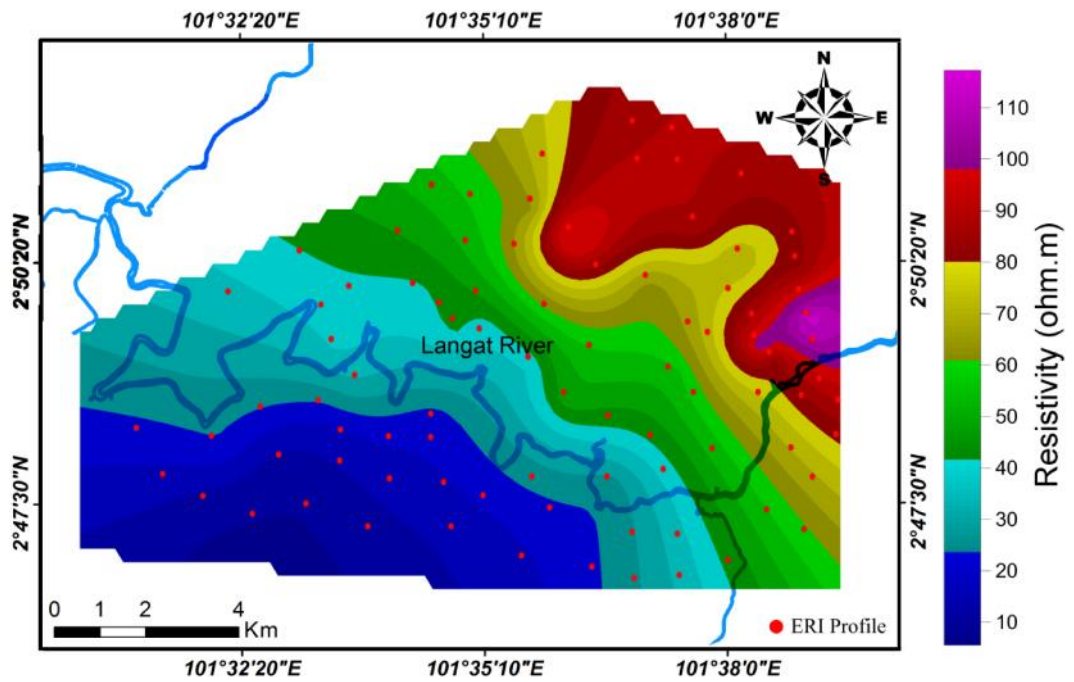


Figure 4.17: Contour distribution interpreted resistivity map of the aquifer layer

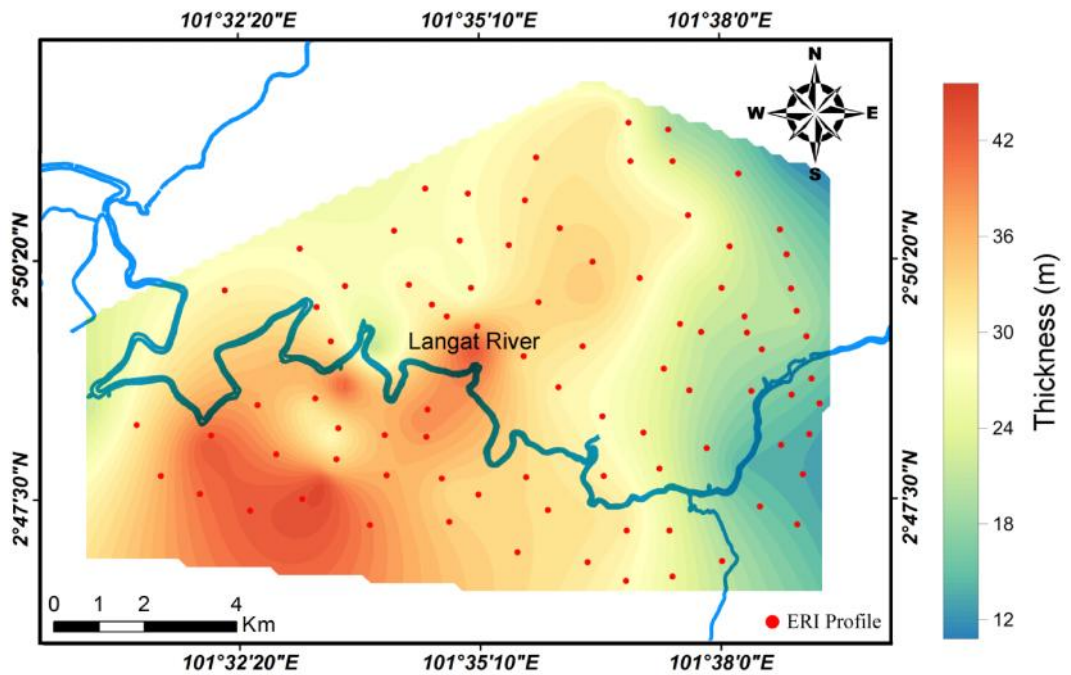


Figure 4.18: Contour distribution map corresponding to aquifer thickness based on the resistivity interpreted

4.2.4 Geoelectrical Cross-sections Model

From the results obtained from the resistivity data, five resistivity lines, as shown in Figure 4.1, were performed for detailed mapping of the shallow impermeable target over various lengths and orientations. Figures 4.19 and 4.20 are displayed the five geoelectrical cross sections in the study area. These sections provide the evidence for the distinct horizontal and vertical variations of resistivity and thickness of the subsurface layers. Four main subsurface layers have been identified following the analyses of these sections. The upper layer (thin level) indicates silty soil mixed in some places with clay and peat materials which the overburden develops to a depth ranging from 0 to 5 m. As indicated from the closest boreholes, the second geoelectrical layer corresponds to a clay layer. The geoelectrical cross sections have shown the depth of the clay layer increases gradually to the maximum 25 m towards the western part of the survey area (coastal area) which is presented in Figure 4.19. Characterization of the

third layer with the depth of 10 m to 60 m has been carried out by using high resistivity values, which are considered to reflect gravel and sand layer (aquifer). The fourth layer is bedrock, with the depth between 25 and 60 m in the east and west of the area respectively whose vertical extent increases gradually to the western part of the area. Figure 4.20 presented the variation of thickness and depth layers are approximately constant in the north to south direction.

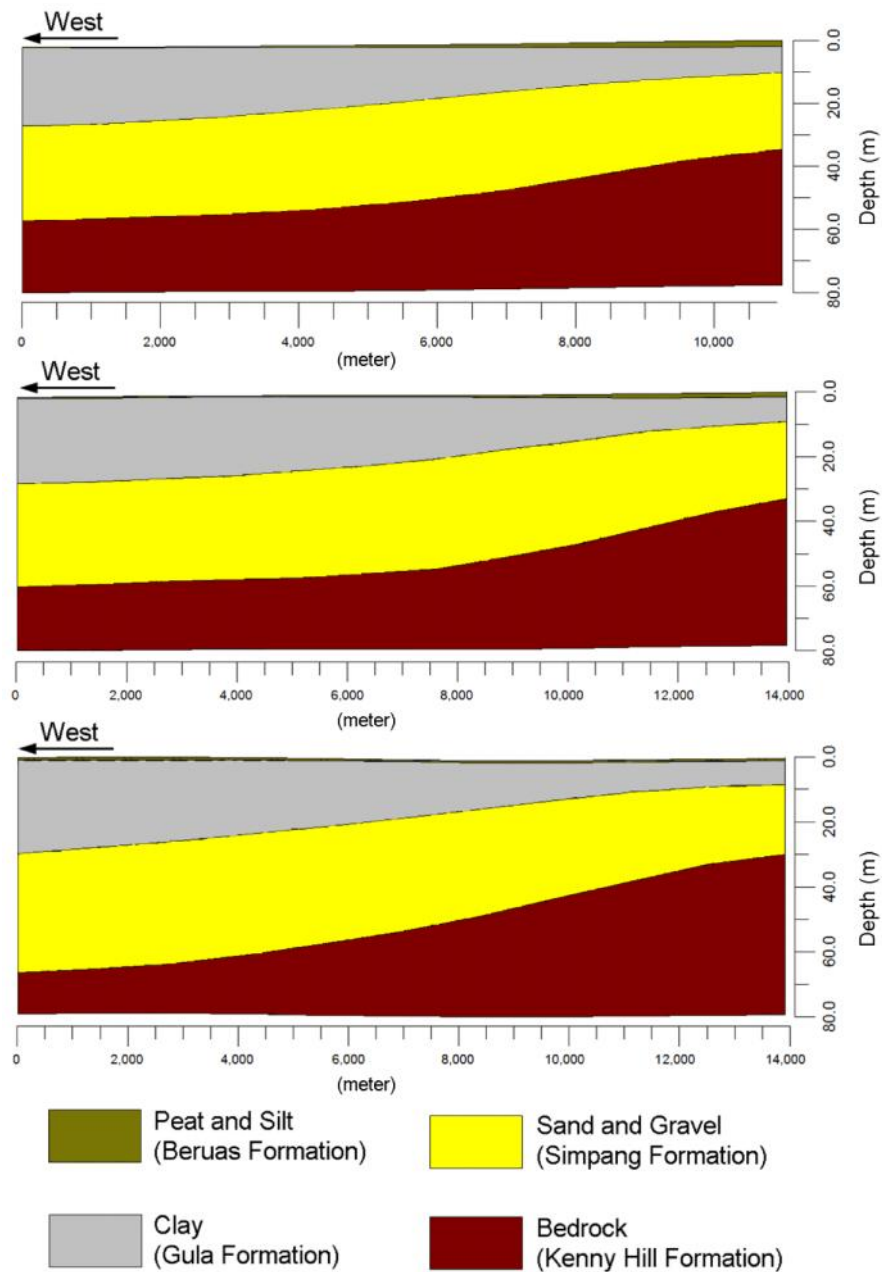


Figure 4.19: Geoelectrical cross sections of the study area along line 1 to 3

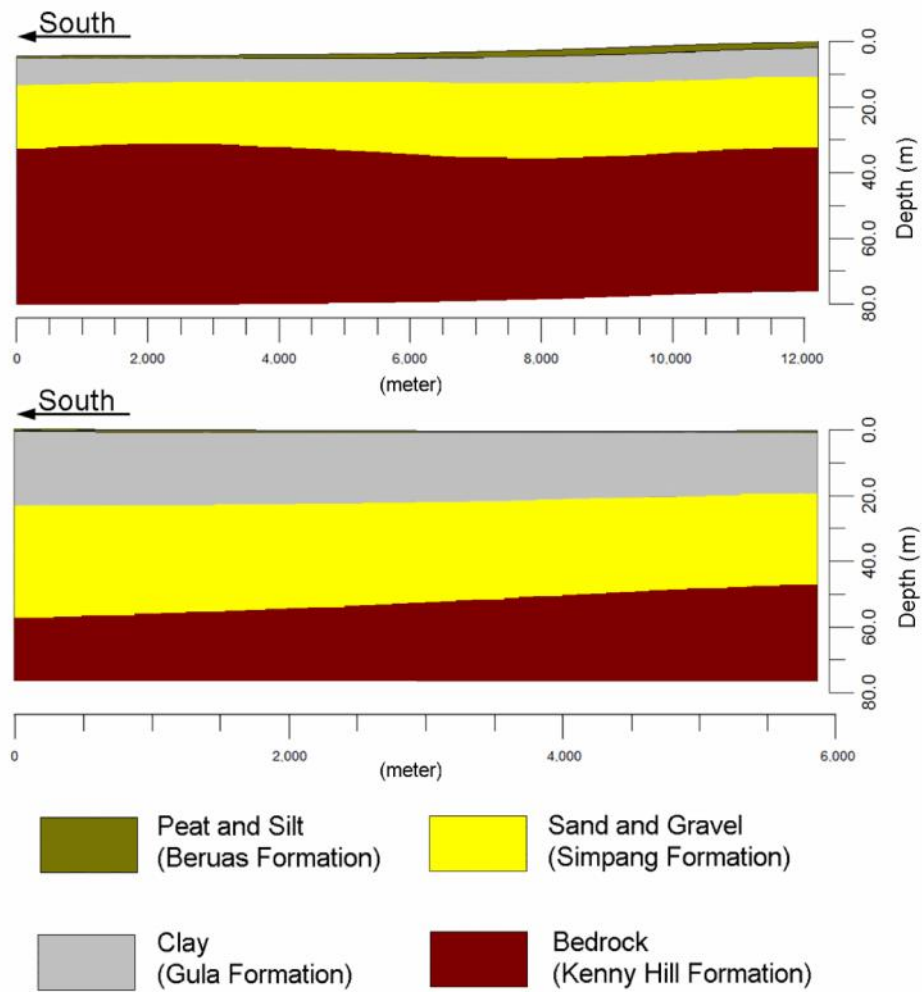


Figure 4.20: Geoelectrical cross sections of the study area along line 4 and 5

The 3D imaging of the horizontal and vertical variations in the study area has been shown in Figure 4.21. The depth distribution has been reconstructed in the form of a solid diagram. The output of the 3D resistivity visualization model of the interpreted true resistivities was sampled to a depth of 70 m from the ground surface. In the 3D resistivity model, the local which increases in the depths of the aquifer and bedrock layers can be observed from the east of to the west of the area. The aquifer layer is presented with a depth of 10 m in the east, increasing to a depth of more than 60 m in the west of the area.

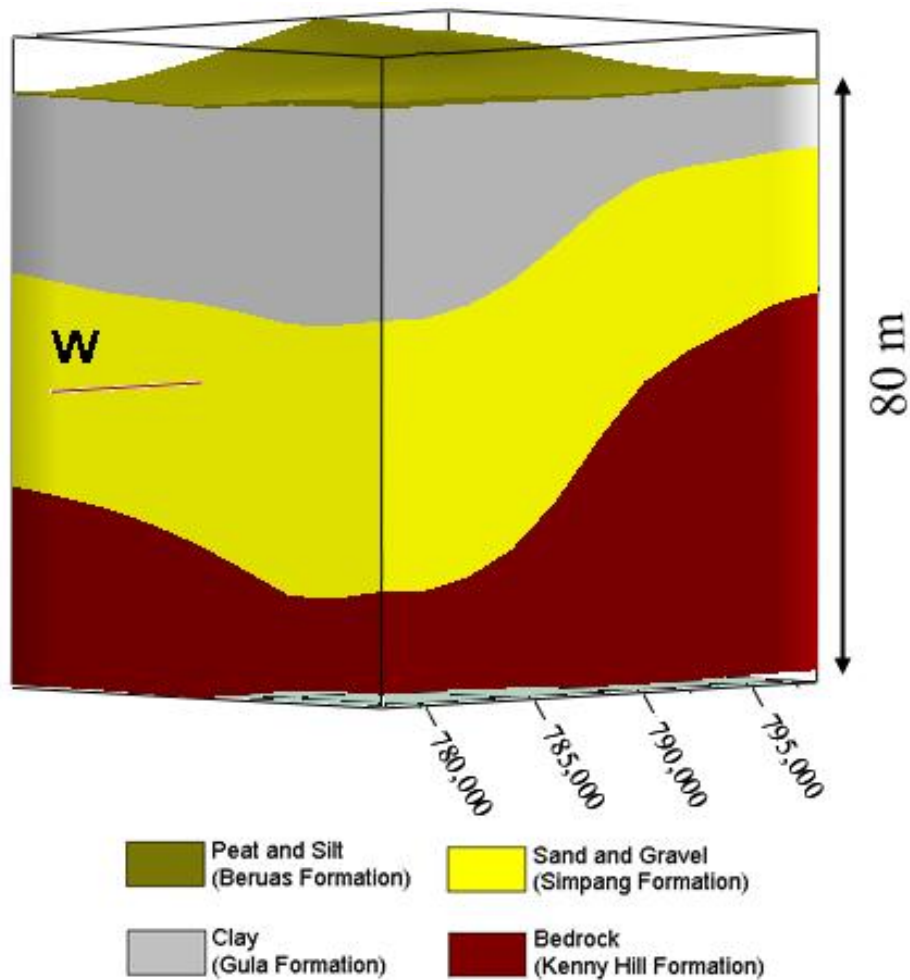


Figure 4.21: Panel diagrams represent the 3D view of the resistivity cross-sections along the study area.

Fence diagram of the resistivity sections in the study area is shown in Figure 4.22. The upper layer corresponding to the thin layer consists of silt and peat which is along the study area particularly in the northern part. The second layer corresponds to the clay layer which has the maximum depth of 25 m. However, an increase in the depth and thickness of the layer is found toward the west part of the study area. The third major layer is characterized by high resistivity values and extends to depth of about 60 m which have been shown in the Figures 4.19-4.22. This layer is composed mainly of sandy gravel sediments and forms the confined Quaternary aquifer in the study area (Figure 3.9). At some localities, the third layer exhibits a relatively low resistivity,

particularly at the west parts of the study area. This feature was due to the intrusion of sea water from the west of the study area. The fourth layer is characterized by higher resistivity values and extends to the base of the sections. In the west part of the study area the lower resistivity is found in this layer (bedrock) which is due to the penetration of sea water to layer.

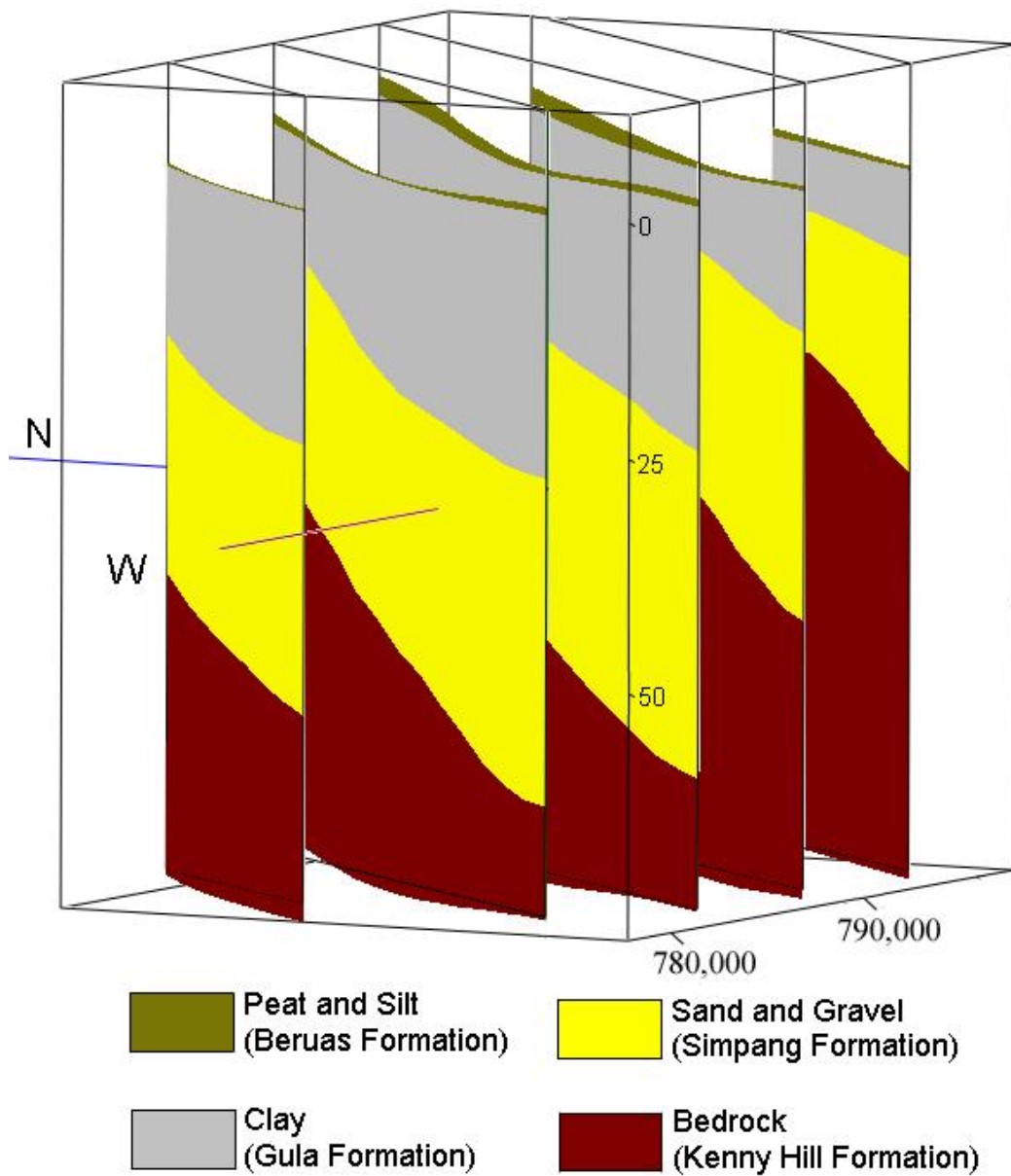


Figure 4.22: Fence diagram of study area

4.2.5 Time Lapse Electrical Resistivity Monitoring of Groundwater Level

The final portion of this section was a test to determine whether time-lapse resistivity could detect changes in the thin top layer and detection groundwater level in the study area. In the present experiment the line was constituted by 64 electrodes organized to have a 40 and 80 m long ERI line (Fig. 4.23) with a 1 m electrode spacing (41 and 61 electrodes). ERI measurements were performed every month from first of February to August continued at the same line position. Four boreholes (B1-B4) with the maximum depth of 3 m along the four ERI profiles are created by using hand auger to obtain groundwater level. The ERI survey traverses were oriented NE-SW. The image was obtained from the inversion model subsurface resistivity, which was obtained from data series measured and was compared with another model (Figure 4.24). Figure 4.24 shows the resistivity images of monitoring profile 3. Based on the monitoring of well no 3 (Table 4.1) water levels dominate at the depth of 39, 84, 114.5, 114.5, 113, 105 and 98 cm which are corresponding to the resistivity pseudosection with resistivity value 33 ohm.m in figure 4.24.

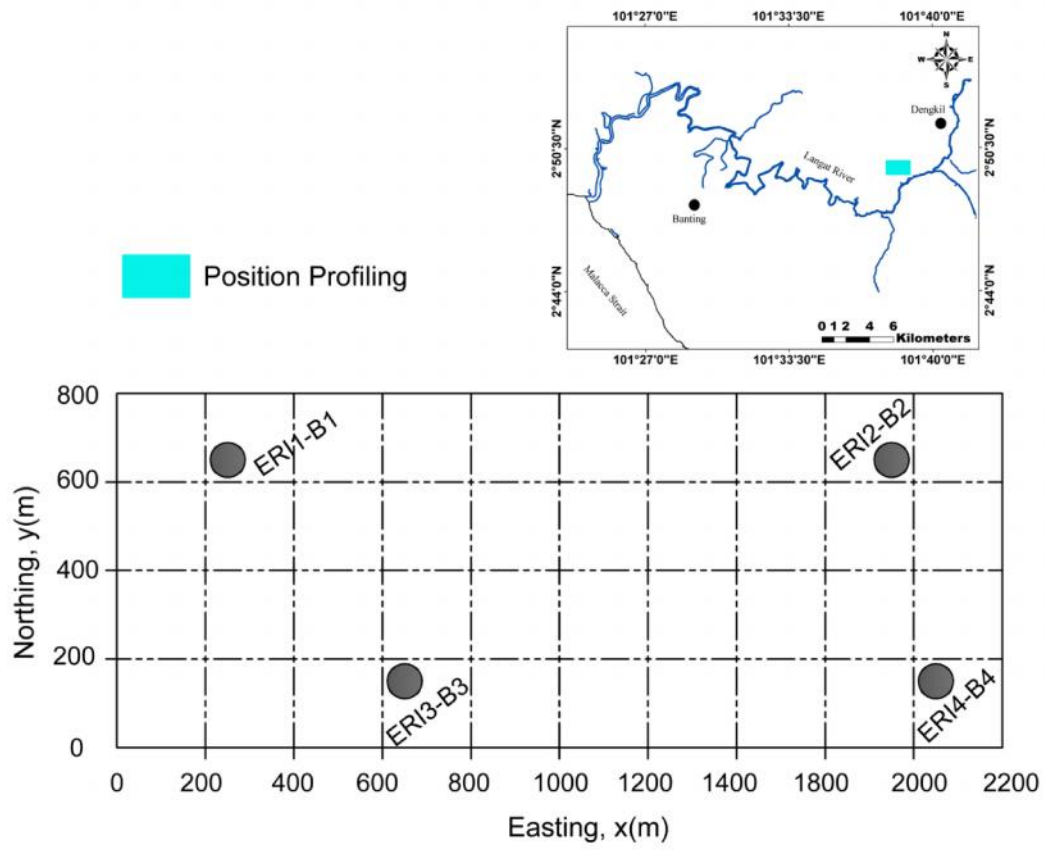


Figure 4.23: Field set up groundwater level monitoring

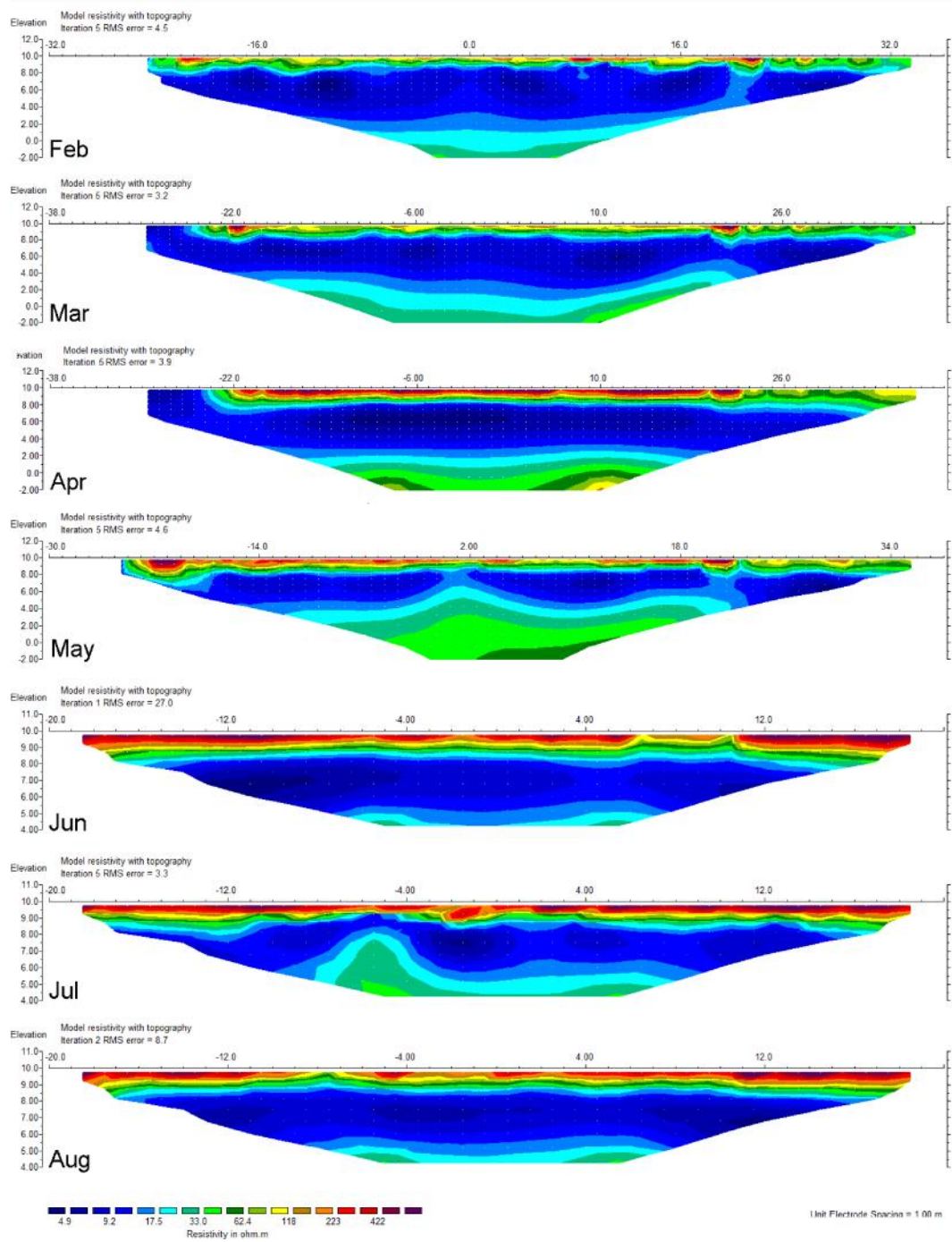


Figure 4.24: Resistivity model for monitoring profile 3 from February to August 2013

Table 4.1. Statistical Data for borehole no 3

| Time | Rainfall (mm) | Water Level (cm) |
|----------|---------------|------------------|
| February | 193.6 | 39 |
| March | 184.2 | 84 |
| April | 95.5 | 114.5 |
| May | 67 | 114.5 |
| June | 67 | 113 |
| July | 170.5 | 105 |
| August | 146.6 | 98 |

4.3 Hydrogeochemical Study

A detailed observation of water chemistry was made from 2008 to 2013 once a year in October and for 17 boreholes to cover all the different lithological units. Figure 4.25 and Table 4.2 displays the locations and physical properties of the wells in the study area, respectively. Groundwater sampling was taken as a required step to resolve spatial and temporal variations in water quality and as a mean to explain resistivity profiling data. Water wells were excavated with several casing volumes of water and sampled only after they had been continuously pumped for long periods of time to accommodate sampling and to remove stagnant water. Additionally, groundwater level measurements were carried out in March and October. The physical and hydrogeochemical results are given in Table 4.3. All water samples were collected within the aquifer and analysed for 31 different parameters.

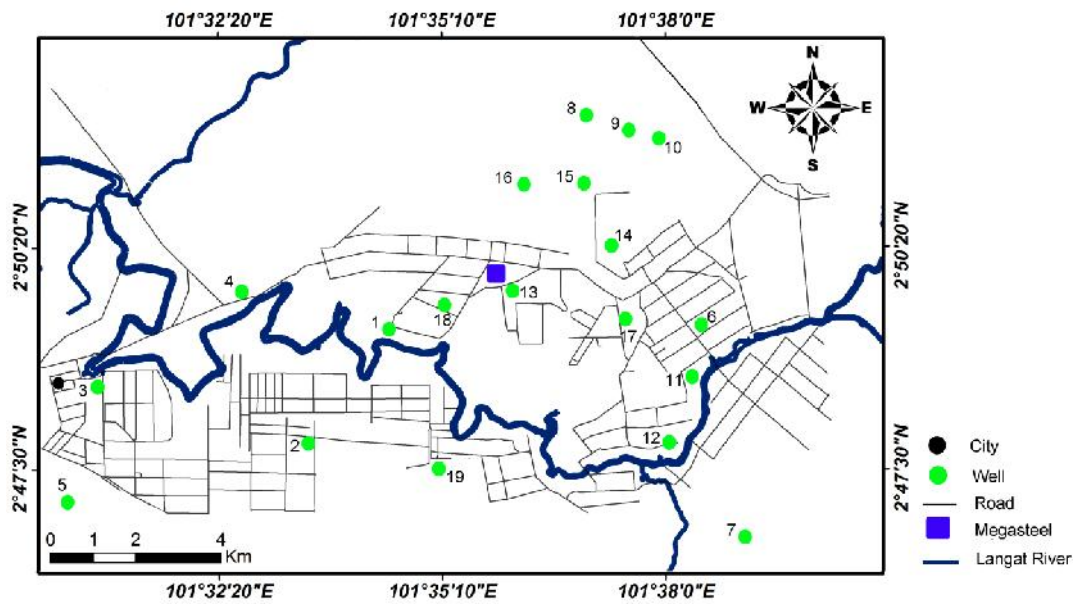


Figure 4.25: Well locations for groundwater sampling in the study area

Table 4.2: List of existing well stations which is used for groundwater sampling (Mineral and Geosciences Department of Malaysia)

| No | Elevation (masl) | Depth to water (m) 2013 | Depth of Bedrock (m) | Depth of Borehole (m) |
|----|------------------|-------------------------|----------------------|-----------------------|
| 1 | 9 | 5.04 | 58 | 60.5 |
| 2 | 12 | 3.85 | 67 | 70 |
| 3 | 7 | 1.55 | 58.5 | 61 |
| 4 | 8 | 2.11 | 38 | 39.5 |
| 5 | 8 | 3.84 | 65 | 67 |
| 6 | 9 | 0.93 | 36.25 | 38.5 |
| 7 | 10 | 1.4 | 25 | 27 |
| 8 | 11 | 0.68 | 27 | 33 |
| 9 | 11 | 0.49 | 31 | 36 |
| 10 | 12 | 1.14 | 27 | 34 |
| 11 | 11 | 1.2 | 36 | 38 |
| 12 | 10 | 1.16 | 40 | 42 |
| 13 | 8 | 2.58 | 32 | 35 |
| 14 | 8 | 2.31 | 30 | 32 |
| 15 | 11 | 1.05 | 27 | 33 |
| 16 | 12 | 0.13 | 31.8 | 34 |
| 17 | 10 | 0.85 | 36.5 | 37 |

Table 4.3. Chemical composition of groundwater samples from 2008 to 2013 (A3)

4.3.1 Water Chemistry

The chemical analyses of the samples collected from the confined aquifer were interpreted. The evaluation of the water quality of the aquifer has been performed based on the measurement of TDS, and the significant ion contents determine whether the quality of the groundwater has varied from saline to fresh groundwater, where saline groundwater includes TDS content in more than 1000 mg/L. The TDS concentrations in groundwater vary over two to three orders of magnitude. The classification scheme for categorizing groundwater, which is based on the total dissolved solids, is shown in Table 4.4 (Freeze and Cherry, 1977). The results of the TDS analyses indicate that most of the west of the area exhibits brackish water with high TDS values (>1000 mg/L), where it is closest to the coastal area. The observations of borehole 3, located roughly 11 km away from the coastline, indicate that the TDS value is more than 1,000 mg/L for saline water. This borehole also shows that the value of chloride content is above 1,000 mg/L. High measured electrical conductivity (23500 $\mu\text{S}/\text{cm}$ corresponding to 2013) supports the high salinity of groundwater from this borehole. This conductivity value is higher than the limit for the saline water baseline (Rhoades, 1982). According to the national water quality standard for Malaysia (NWQS) (DOE, 2006) were used to evaluate Langat Basin quality. According to the Table 4.3 the overall classification of the study area, water quality classified as V and IV which is suitable for agricultural irrigation only and requires extensive treatment for drinking for the west of the study area and classified I and IIA represents water bodies of good quality.

Figures 4.26 shows the contour map of distribution TDS value during the study from 2008 to 2013. The results of the TDS analyses indicate that most of the west of the area exhibits brackish water with high TDS values (more than 1000 mg/L), where it is closest to the coastal area. Regarding the obtained results, the TDS has increased roughly from 2008 to 2013 in the southwest and towards the west of the study area.

Demonstrated in Table 4.3 and contours (Figures 4.26), TDS has roughly increased in the southwest and west of the study area from 2008 to 2013.

Table 4.4. Groundwater classification based on Total dissolved solids (Freeze and Cherry, 1977).

| Class of water | TDS (mg/L) |
|----------------|----------------|
| Fresh water | 0-1,000 |
| Brackish water | 1,000-10,000 |
| Saline water | 10,000-100,000 |
| Brine water | >100,000 |

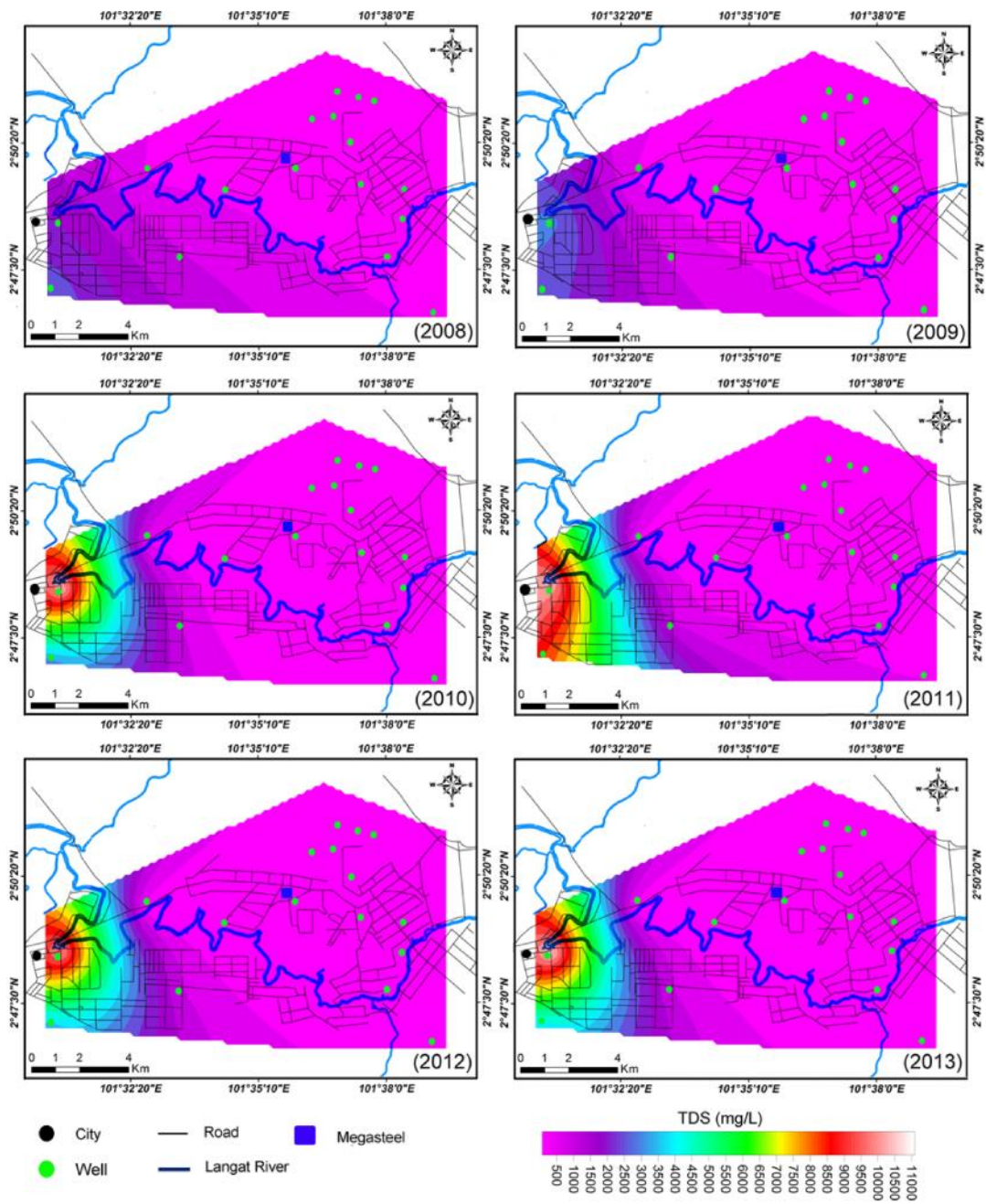


Figure 4.26: Contour map of Total dissolved solid level in Langat basin from 2008 to 2013

The major ion composition of groundwater is used to classify groundwater into various types based on the dominant cations and anions. The composition of the dominant ions can be displayed graphically by several methods. Piper and Stiff diagrams as two multivariate graphical methods are applied to study of water-quality. The Piper diagram

is an effective chemical graphical representation employed in water samples in hydrogeological studies. A Piper diagram (Figure 4.27) was created for the Langat Basin area by using the analytical data obtained from hydrochemical analyses. The general trends in aquifer chemistry are indicated by the arrows. Figure 4.27 shows that the average water chemistry from 2008 to 2013 in the study area ranged from calcium bicarbonate water in the wells to water with a more neutral ionic distribution, found in the streams in siliciclastic rocks. The average annual water chemistry analyses of the Langat Basin catchment is plotted on the Piper diagram. Based on the results for the concentrations of cations, the water type was shown to be Na+k in terms of calcium type. An abundance of sodium in the area can be interpreted as marking the intrusion of seawater into the ground- and/or surface waters. Therefore, salinity plays a significant role in controlling the exchangeable cations in the study area. However, the results for anions cannot be interpreted in the same way. Anion concentrations are not limited to a specific type of water. The chemistry of water samples is more controlled, with a higher concentration of HCO_3 and Cl (analyses of samples yielded a straight line in anion triangles close to the HCO_3 -Cl side) due to the type of bedrock and salinity in the area. Total ions measurement accuracy was evaluated through computing the ion balance error which was within $\pm 10\%$ (Table 4.3). Generally, the classification of water is marine and deep ancient groundwater (on the right side) over the 6 years analyses of water chemistry obtained from the catchment indicated by groundwater facies (central quadrilateral of the piper diagram).

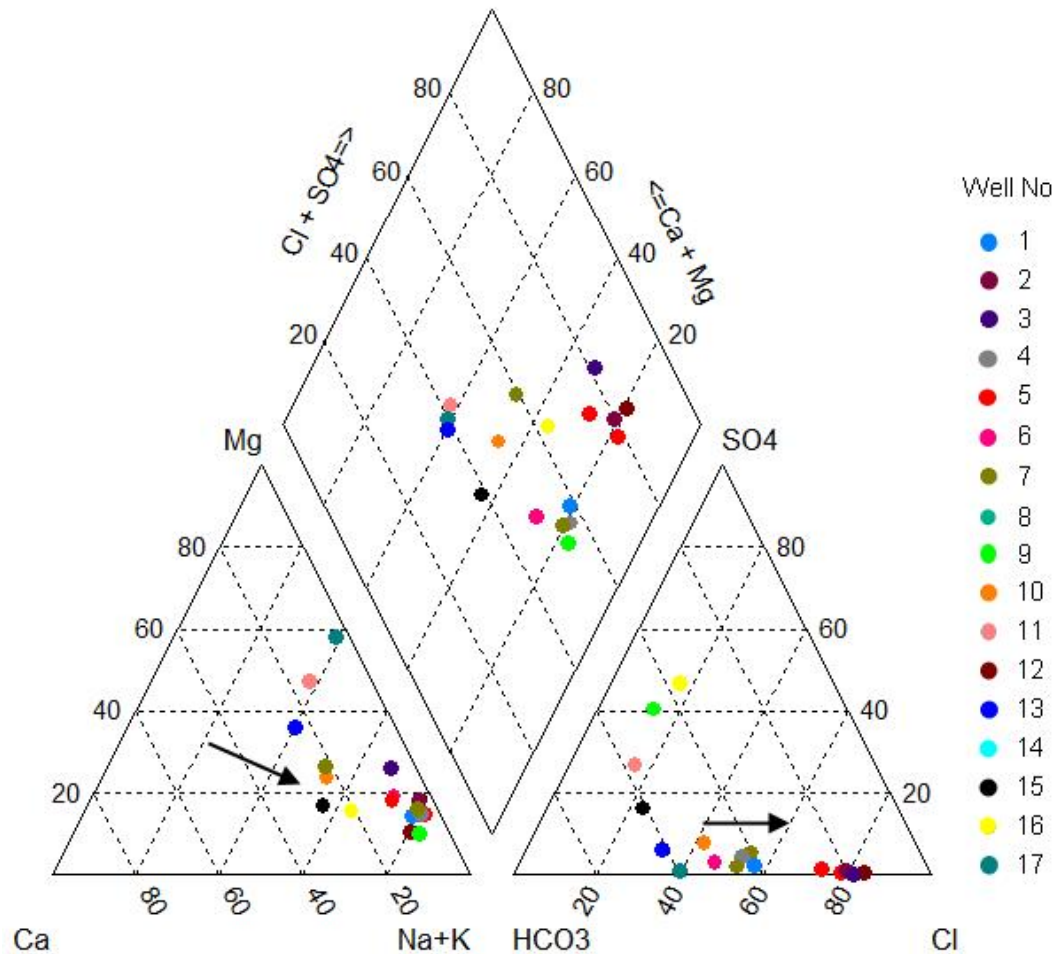


Figure 4.27: Piper diagram showing the average chemical composition of the groundwater samples along different wells in the study area from 2008 to 2013. The detailed chemical data is listed in Table 4.3

Stiff diagrams are known as the most popular application of profile plots on water resources. The shape of the Stiff diagram was employed to show if there were differences in the chemistry of the source water and that of the acquiring ground water. Concentrations are expressed as meq/L in Stiff pattern diagrams. They are used to speed up the evaluation involving different water samples from various sources successfully. The resulting points are connected to provide an irregular polygonal pattern. The size of the pattern is roughly equal to the content of total ionic. The patterns acquired by using the Stiff approach to interpret images and plot average of water parameters from 2008 to 2013 analyses of 17 samples of water from wells in research area are shown in Figures

4.28. The intuitive and quick identification of water types related to geology along various flow paths is enabled by this data visualization. The patterns present the predominant ions that define the chemical type of the water. Cations and anions profile are plotted in Figure 4.28 which are shown on the left and the right of the vertical zero axis respectively. The shape of each diagram in the same classification is just like the others as presented in the Stiff diagrams. Therefore, the evaluation of the variation of water quality at a single position over a period of time through the eyes is simpler. Figure 4.29 presents the Stiff diagrams annotated on map of the study area that illustrate various chemistry (polygon shapes) for each piezometer situated in the alluvium. Water quality in this area is divided to two main groups, I and II, that group I includes water with overall low TDS and EC concentrations and group II consists of higher TDS and EC values. Within the first group, wells 8, 9 and 16 have the lowest EC, due to low concentrations of dissolved ions such as Cl^- , HCO_3^- and Ca. In the second group, the highest EC and TDS values are observed for the well numbers 3 and 5 which is caused by seawater intrusion. More than 50% of groundwater is a NaCl water type at all sample locations that was indicated by the Stiff plots. Usually, NaCl type of groundwater can be found close to the coastal area since this type of water contains seawater in many cases. The mean EC and TDS values of well 3 and well 5 are high, mainly due to higher concentrations of Cl, Na, HCO_3^- and Mg in both water wells. This is obviously observed in the Stiff diagrams and shown in Figures 4.28-4.29. Moreover, wells 3 and 5 situated at around 9 km from the shoreline, and monitoring wells such as 1, 2, 4 and 6, can be more or less affected by sea water.

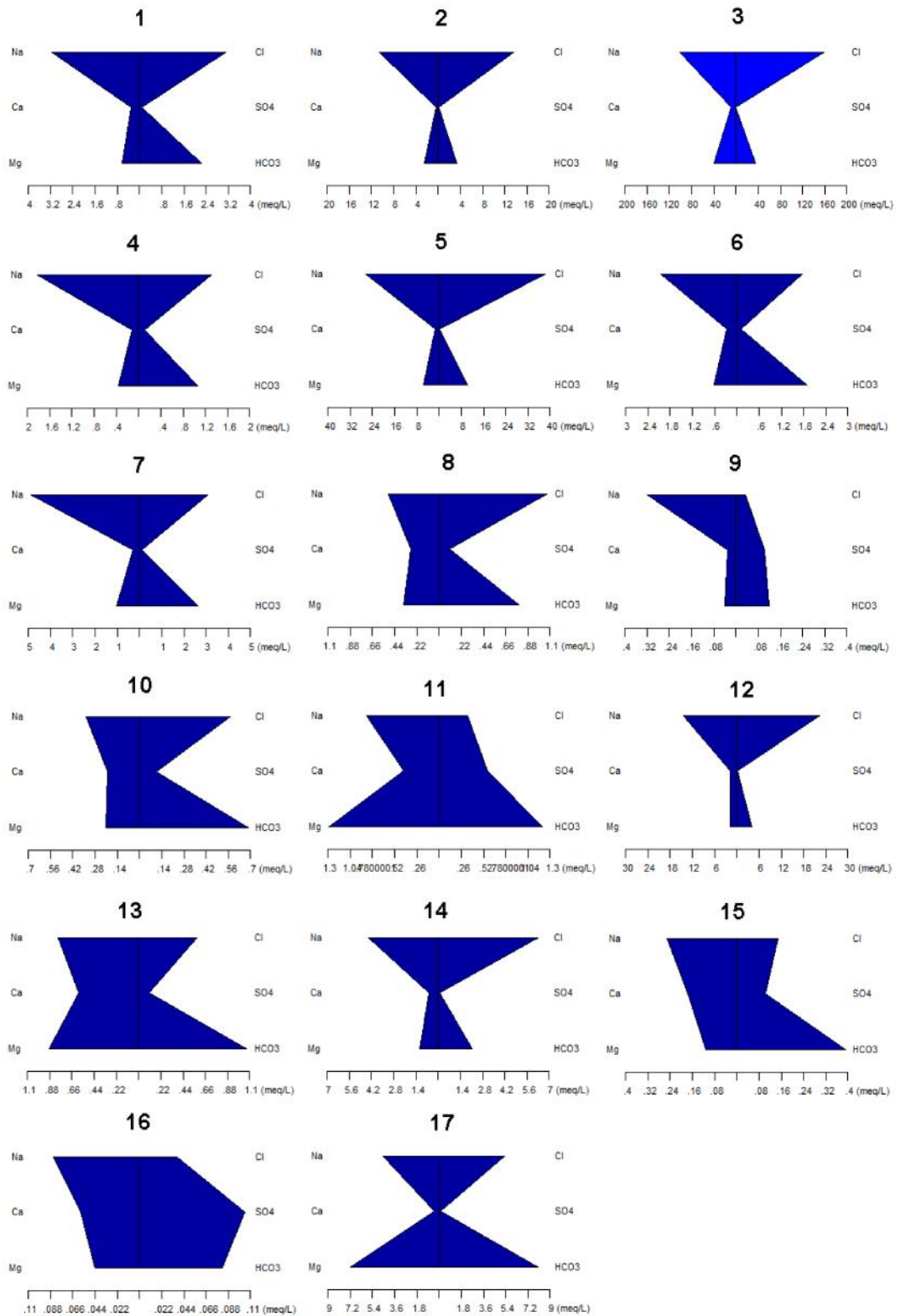


Figure 4.28: Stiff diagram showing the average chemical composition of the groundwater samples along different wells in the study area from 2008 to 2013. The detailed chemical data is listed in Table 4.3

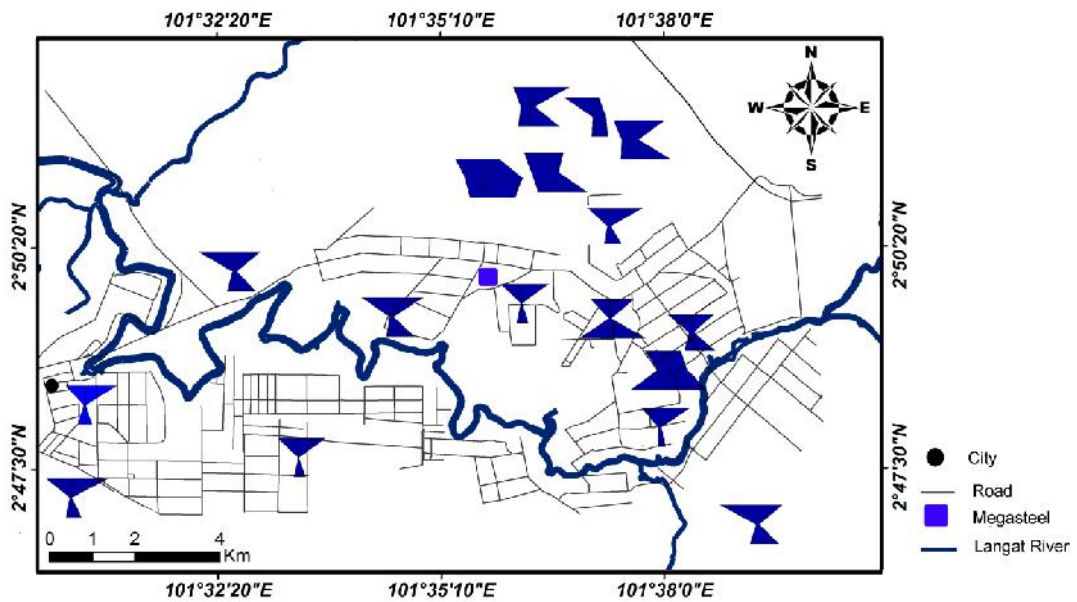


Figure 4.29: Areal display of Stiff diagrams for wells from the study area. Diagrams are centered on their well locations.

4.4 Groundwater Level Monitoring

The measurements of groundwater level in the study area were carried out in 17 boreholes twice a year from 2008 to 2013. In general, the field surveys were planned to correspond to the rainy and dry seasons, namely October and March.

Based on the observations from existing wells, groundwater level contours were prepared for March and October from 2008 to 2013 (Figures 4.03 and 4.31). The water level during the study period varied from a minimum of 3.55 MASL in the well 4 (March, 2012) to a maximum of 11.14 MASL in the well 10 (October, 2010). The main flow directions in the study area were from north-east towards south-west, and from south to north.

In the centre of the study area, the water levels in some monitoring wells may not be closely related to natural variation. This is observed for wells placed close to the Mega Steel Company, which overexploits the aquifer. The actual volume extracted by

the Company is questionable. However, the groundwater pumping records show that a mean of 23,300 m³/day was extracted from some wells in the past (Minerals and Geosciences Department, 2002).

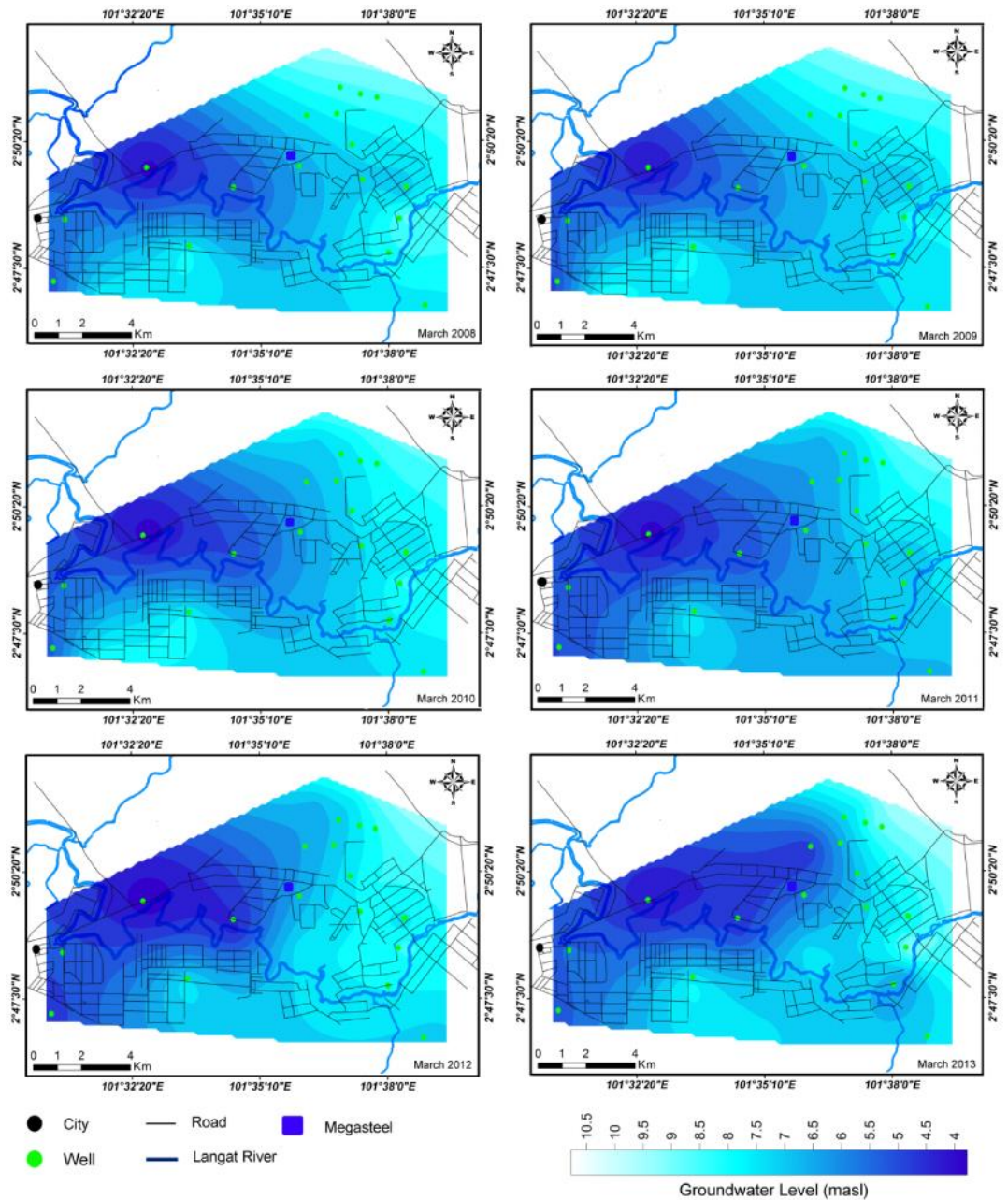


Figure 4.30: Groundwater level (masl) in the study area in March, 2008 to 2013

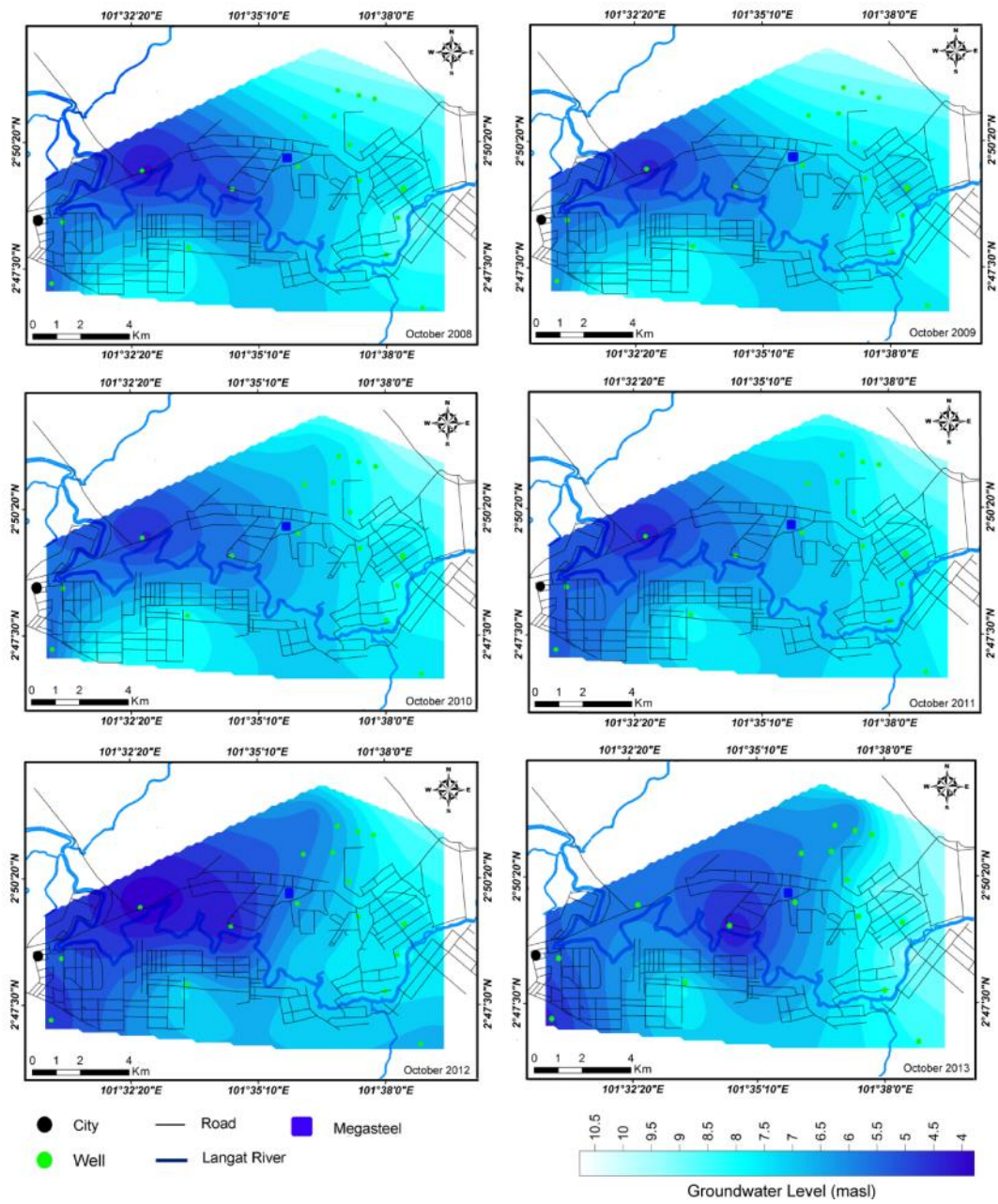


Figure 4.31: Groundwater level (masl) in the study area in October, 2008 to 2012

4.5 Summary

In this chapter the electrical resistivity and hydrogeochemical analyses results have been employed to investigate the groundwater. The results consist of the inverted ERI image, cross section of study area based on resistivity results, chemical analyses of groundwater samples and monitoring of groundwater. Level combination of the

resistivity investigation with borehole data and the geological mapping provide comprehensive understanding of the thickness and depth of aquifer and bedrock surface. Therefore, a combined interpretation of the obtained results from the geoelectrical and hydrogeochemical approaches have been performed and discussed. These results will lead to groundwater resource managements.

For the processing of the geophysical data, the RES2DINV software was used and the obtained results were demonstrated in this section. The 2D multi-electrode resistivity survey has developed important 2D sections with high quality structural resolution that leads to significantly identified layering of various lithological units, identification of structures, depth to bedrock and weathering profiles. These structures are found to be quite essential for investigation on groundwater. Resistivity approach is cost effective to characterize the subsurface and it helps to define the aquifers thickness and to map the bedrock precisely with a relatively shallow depth. The results of the ERI inversion in the area demonstrate a representative and clear image of the aquifer. The mapping the bedrock and of the aquiferous sand and gravel zones that are located under the clay layer were determined in the survey. The interpolation of an interpreted geoelectrical model generates the 2D and 3D map of the area. The cross section of geoelectrical model and lithological data obtained from existing boreholes illustrate the good correlation between thickness of aquifer varying from 10 to 30 m (and up to 40 m) and the depth of the bedrock changes from 20 to 60 m. The lower resistivity values indicate the aquifer in the west of the area contained brackish water. In addition, the time lapse monitoring survey of geoelectrical resistivity has been successful in mapping the groundwater level in the part of the study area.

Chemical analyses shows, the concentration of TDS and EC has gradually increased during the monitoring in the collected groundwater samples at the west of the study area. High concentrations of TDS and EC found in the aquifer at the west part of

the study area and water are not safe for human consumption. While the lower concentration of TDS and EC was observed in the east parts that make the water safe for human consumption. Moreover, water level monitoring was carried out to determine the groundwater flow showing that the general flow is in to south-western part of research study.

Regarding geoelectrical resistivity, lower resistivity value shows that the west of the area was affected by sea water intrusion. High TDS is caused to decrease in the resistivity values. On the basis of acquired electrical resistivity profiles, the hydrogeochemical results indicate that the groundwater in the aquifer within the study area is both fresh and brackish. The fresh water was indicated with low TDS and chloride concentrations the groundwater. Chloride and TDS concentration tend to be higher in the water sample in the zone marine soil deposit. However, the concentration is within the accepted limit for human consumption. The geoelectrical resistivity model illustrates a low resistivity value less than 28 ohm.m in the zones the aquifer filled by salt/brackish water. This value depends on the percentage of salt water in the aquifer and soil characteristics of the aquifer. The resistivity value shows an increase landward as a decrease of seawater mixture.

This research shows that 2D resistivity surveys are capable as an effective method to delineate the subsurface lithology of the study area comparing with borehole data results. The hydrogeological units and inverted resistivity tomograms with borehole data have been well matched. The knowledge concerning water quantity and quality can be helpful for managing land and groundwater resources.

Chapter 5: Prediction and Assessment of Groundwater Depth and Quality

5.1 Introduction

Artificial neural network (ANN) and adaptive neuro fuzzy inference system (ANFIS) models were applied as tools to predict groundwater parameters in this chapter. They use several various architectures of neural networks and membership functions for forecasting fluctuations of water level and quality. The correlation coefficient (R) and the mean square error (MSE) determine the accuracy for these approaches. The obtained data from various time periods at the same location can validate the ANN and ANFIS model for modelling. The predicted water level fluctuation and quality are compared with the actual data to examine its accuracy. A site-dependent model to predict water table and quality variation at the selected sites was provided in this portion of the study.

In previous decades, artificial intelligence (AI) methods, such as ANNs and the ANFIS were applied as strong tools and accurate solutions to many of the extremely difficult challenges faced by water sciences and hydrology, and this usage has increased. The growing AI approaches have the ability to fill the gaps of the measurements and to forecast future values without long observation data (Karimi et al., 2013).

The detailed discussions on neural network modelling have been done earlier in chapter 3. In the improved neural network, there is an input layer, where input data is introduced to the network and an output layer, and also two hidden layers as intermediate layers. In the improved ANFIS, the same employed input parameters with ANN model were chosen for evaluation of the target parameter. In this study, the performed ANN models are of three networks (section 3.5.2.5) and four algorithms

(section 3.5.2.7) that were usually applied to solve regression problems which were the models typically including three layers, namely an input, hidden and output layer. ANFIS models were performed with two common membership functions (section 3.5.3.1).

5.2 A Neural Network and Adaptive Neuro Fuzzy Inference System for Modelling of Groundwater Level (Single Well)

Application of the ANN and ANFIS models to simulate the groundwater level has been targeted in this section. The training process was undertaken with 84 monthly data sets covering the period between 2007 and 2013. The processing was applied to the FNN, CFN and RNN for various epoch numbers. Similar sets of input and output data were applied in the ANFIS modelling, including Gaussian and Generalized bell MFs for each input, which were found to be satisfactory for the processing of the model. The number, step size and shape of the MFs as predefined internal ANFIS factors were efficient, making it possible for the ANFIS model to achieve the performance objective. The selected well for modelling of groundwater fluctuations in the study area is displayed in the Figure 5.1. It has been selected due to following reasons:

- 1) It is located far from the coastal area
- 2) It is not influenced by Mega Steel Factory which pumps a wide range of water
- 3) Data collations are more convenient compared to other wells
- 4) This well is located in the area with lower slope

This area is flat lowland where oil palm plantations and urban lands are the dominant land uses. The elevation of selected well is 10 masl.

5.2.1 Data Description

According to the water balance equation, the changes in the amount of water within any hydrological system can be given in terms of the difference between the inflow and outflow:

$$X - Y = S$$

Equation (5.1)

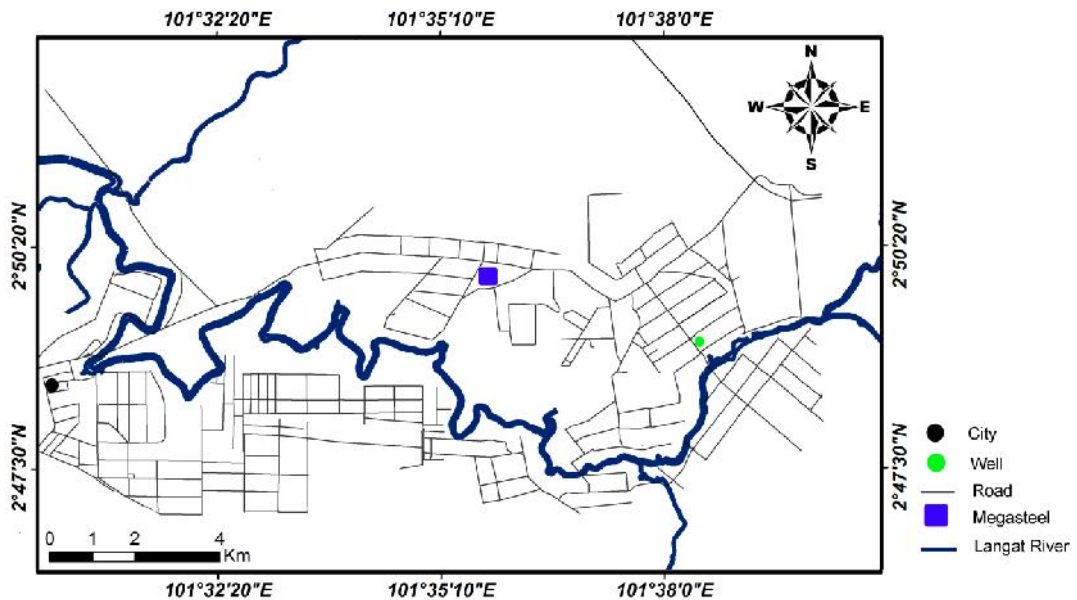


Figure 5.1: Location of the selected well for modelling of groundwater level in the study area.

where X shows inflow like precipitation, Y is outflow like surface run-off, evaporation, infiltration and groundwater flow, and S shows water level variations. In this study, monthly groundwater levels, precipitation, evaporation, humidity, maximum temperature and minimum temperature data for the selected single well were identified and trained with ANNs and ANFIS. These data are known as efficient parameters in the fluctuation of groundwater levels (Lallahem et al., 2005). The tests and obtained results were derived through the Matlab R2012a computer program. Figure 5.2 displays the minimum, maximum and mean monthly rainfall in the study area from 2007 to 2013.

The maximum rainfall occurs in the September to December with a mean of 433 mm. The highest amount of rainfall (480 mm) occurs in December 2012. The minimum rainfall occurs in the February with a mean of 94 mm. In addition, the number of rainy days in a year ranges from 140 to 210 days. Figure 5.3 shows the plot the monthly precipitation and groundwater level measured over 6 year period of the study which was between 2007 and 2013. The figure also shows monthly depth to groundwater measurements at the Langat Basin observation well during the study. The depth of the groundwater is well correlated to the precipitation depth throughout the certain period, as is clearly shown in Figure 5.3. The time series of water table depth and of rainfall at the selected well from January 2007 to December 2013 indicates that there is clear correlation between them. Temperature has a principal role in the water budget as it impacts evaporation (Te et al., 1988). Monthly minimum, maximum and mean temperature measurements at the meteorological station of the study area are represented in Figure 5.4. A slight increasing trend is found in values, which change steadily over the years as identified by the bold line in Figure 5.3. This minor trend is neutralized when compared with longer data sets, and significant temperature variation over a long period of time was not observed. Furthermore, the monthly evaporation and humidity in the study area are represented in Figures 5.5 and 5.6, respectively. The average monthly relative humidity in the range of 77% to 85% varies from place to place of research area and from month to month. The minimum range of average relative humidity varies from 67% in February to 79% in November. The maximum range of mean relative humidity varies from 82% in June to 89% in November. In Peninsular Malaysia, the lowest relative humidity occurs in January and February while the highest relative humidity normally happens in November (MMD, 2013). Consequently, the monthly values for the variations of evaporation, humidity, rainfall, minimum and maximum temperature and variation of water level, for 2007 to 2013,

were identified and trained with ANNs and ANFIS for modelling of groundwater level in the study area. Therefore, the numbers of input and output data were arranged at five and one, respectively. The data was randomly included in the training (70% of all data) and testing (30% of all data) of the data sets. The statistical properties of the data used are represented in Table 5.1.

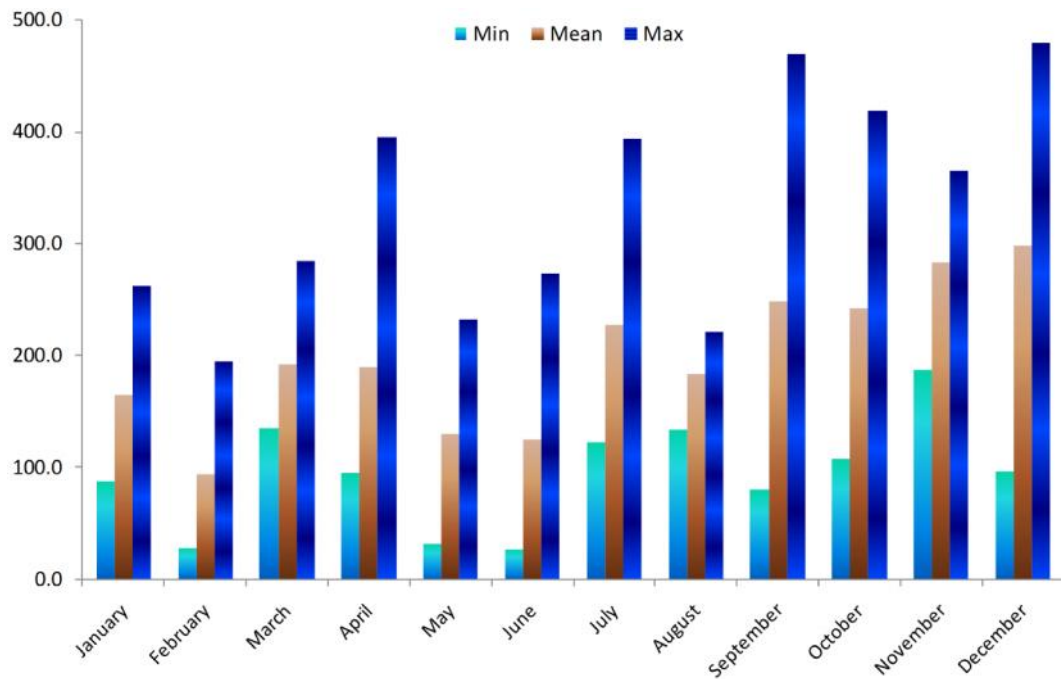


Figure 5.2: Monthly rainfall in the study area from 2007 to 2013.

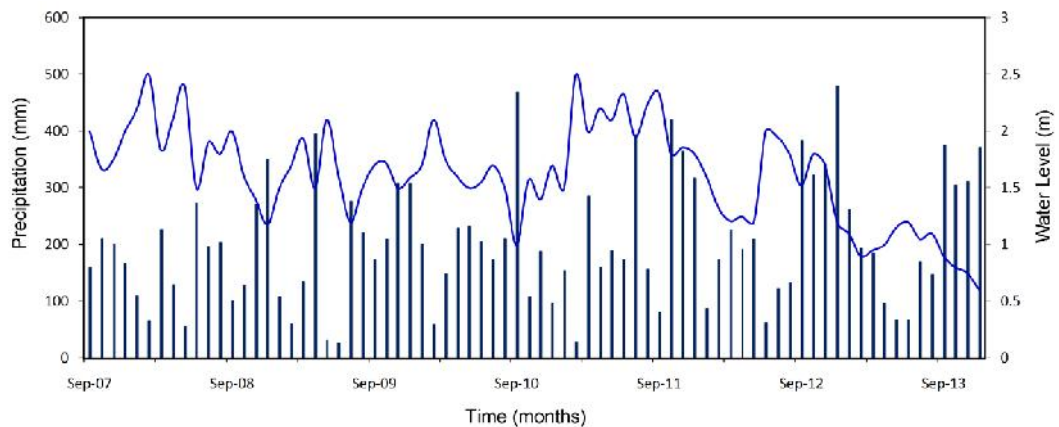


Figure 5.3: Monthly precipitation (mm) and depth to groundwater level (m) for observation well from 2007 to 2013 in the study area.

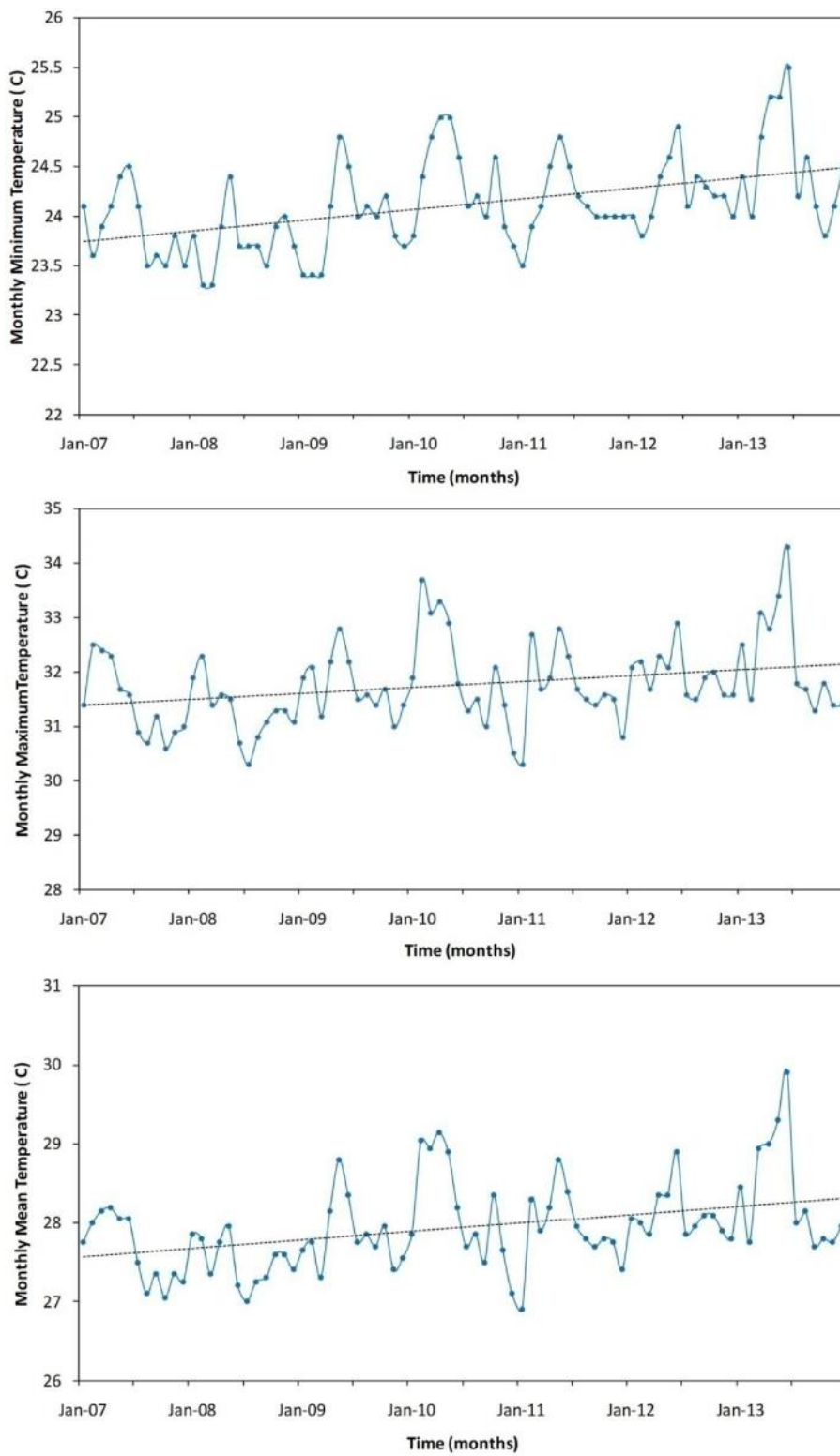


Figure 5.4: Minimum, maximum and average monthly temperature (°C)

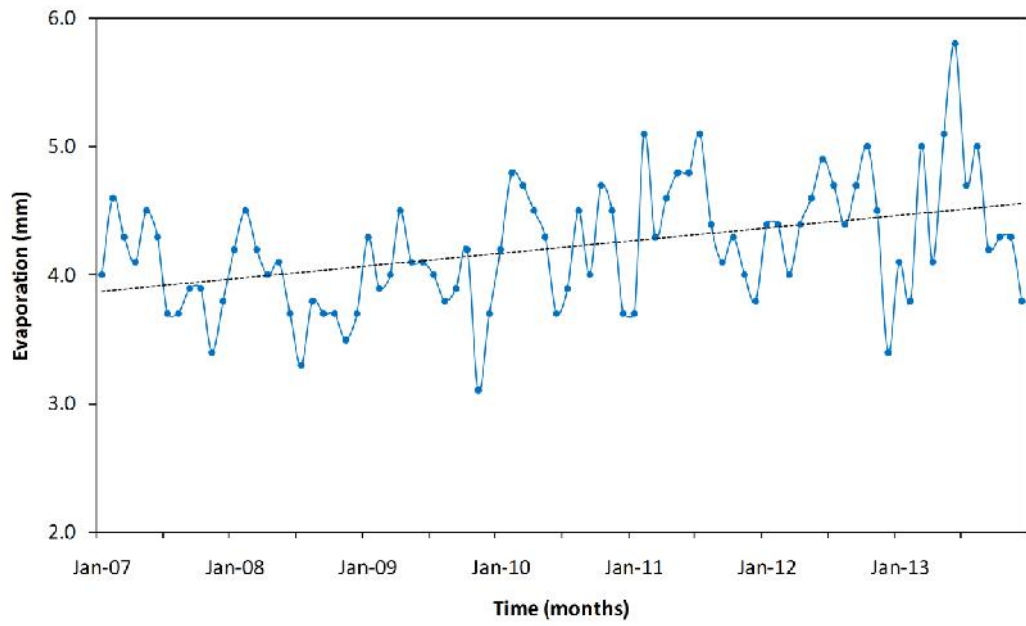


Figure 5.5: Average monthly evaporation (mm) in the study area.

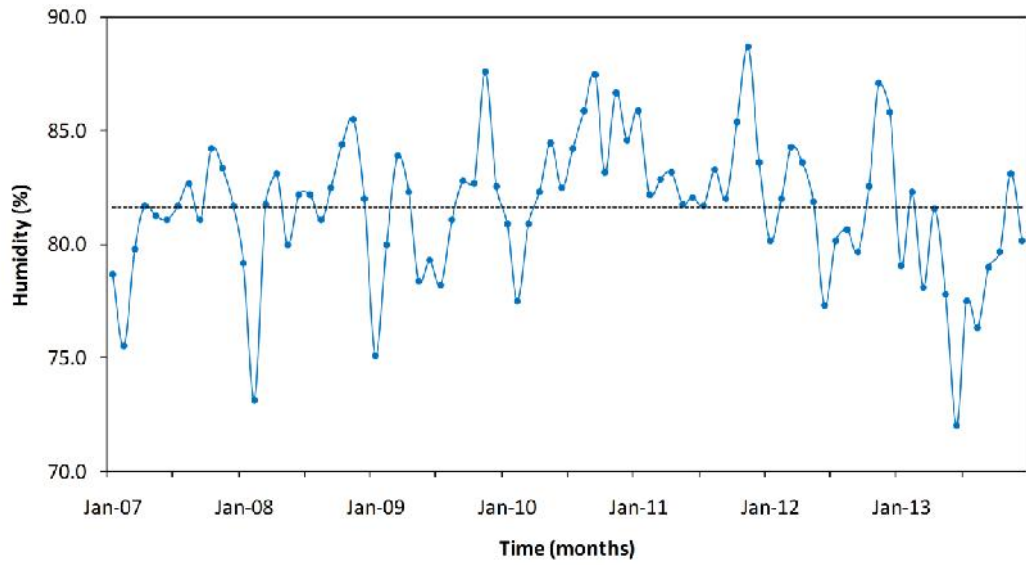


Figure 5.6: Average monthly humidity (%) in study area

Table 5.1. Characteristic of monthly data sets

| Data sets | Unit | mean | minimum | maximum | standard deviation | coefficient of variation | skew | Kurtosis |
|-----------------|------|--------|---------|---------|--------------------|--------------------------|-------|----------|
| Rainfall | mm | 190.81 | 26.9 | 480 | 105.5 | 55.29 | 0.7 | 0.16 |
| Humidity | % | 81.78 | 72 | 88.7 | 3.06 | 3.74 | -0.58 | 1.25 |
| Evaporation | mm | 4.21 | 3.1 | 5.8 | 0.48 | 11.51 | 0.4 | 0.44 |
| Min Temperature | C | 24.1 | 23.3 | 25.5 | 0.47 | 1.95 | 0.61 | 0.22 |
| Max Temperature | C | 31.78 | 30.3 | 34.3 | 0.78 | 2.45 | 0.66 | 0.67 |
| Water Level | m | 1.72 | 0.9 | 2.7 | 0.41 | 24 | 0.08 | -0.48 |

5.2.2 Building and Training the Network

The number of neurons was chosen as twice the number of inputs for the hidden layer. After trial and error, this choice appears to be the most useful. The regularization algorithm can help selecting the suitable number of neurons and therefore improve generalization, however too much memory is necessary to handle as big a set of inputs as the one employed in this project. When the regularization algorithm was attempted on the data, the computer ran out of memory. Therefore, trial and error after the greatest selection of inputs sets modified the final number of neurons in the hidden layer.

For ANN model, the transfer function in the hidden layer was set to a sigmoid, because the initial evaluation which has shown the sigmoid function provides better results comparing to the other transfer functions, although the pure linear transfer function has been carried out in the output layer. Multiple layers of neurons with nonlinear transfer functions enable the network to understand nonlinear and linear relationships involving input and output vectors. The final input required for generating the network and algorithms item is the training function to be employed. Three networks from section 3.5.2.5 and four algorithms obtained from section (3.5.2.7) have been chosen for the training function in the present study. The number of epochs was

set to 500 and the goal to $1e-8$, and the network was trained with the training set of data (data ranging from 2007 to 2013 were available).

5.2.3 Membership Functions (MFs)

The number of types of MFs and the number of MFs assigned to every input parameter for ANFIS analyses was selected by trial and error. Figure 3.22 illustrates the common membership functions for entire input parameters. The Gaussian and Generalized bell membership function presented in Figure 3.22 and equation 3.34 and 3.35 were employed as the MFs for this research. The membership grades are between 0 and 1 for all data in the database.

5.2.4 Training and Testing Results for ANN Model

Neural network model has developed to estimate water level with applying the database of 84 months of test data. For developing the models a three-layer back propagation ANN with five input nodes in the input layer and two hidden layers have been chosen. Regarding the results from several test runs, the number of hidden nodes was determined. The sigmoid function as the activation function was employed. The structure of this three layer ANN model, which uses five previous months data set as inputs layer has been shown in Figure 5.7. These parameters exist in the input file. Another file consisting of the measured water level depth (output) is applied as the target file. Using trial and error, all three networks derive an optimum network and parameter configuration. The network was trained continually by updating weights until error goal is $1e-8$ or 500 epochs. Selected values of learning parameters to obtain the best performance of network when correlation coefficient (R) is highest and MSE error

is the lowest are illustrated in Table 5.2. The total data set, all training and testing related to the R and the MSE, between the target unit and the output of the all networks and algorithms for the observation well, are demonstrated in Table 5.3. The best overall performance was resulted from the achievement of feed-forward network trained with the Levenberg-Marquardt algorithm which were presented in Table 5.3 and by the cascade forward network trained with the same algorithm known as the second best was demonstrated by their small MSE error and higher correlation.

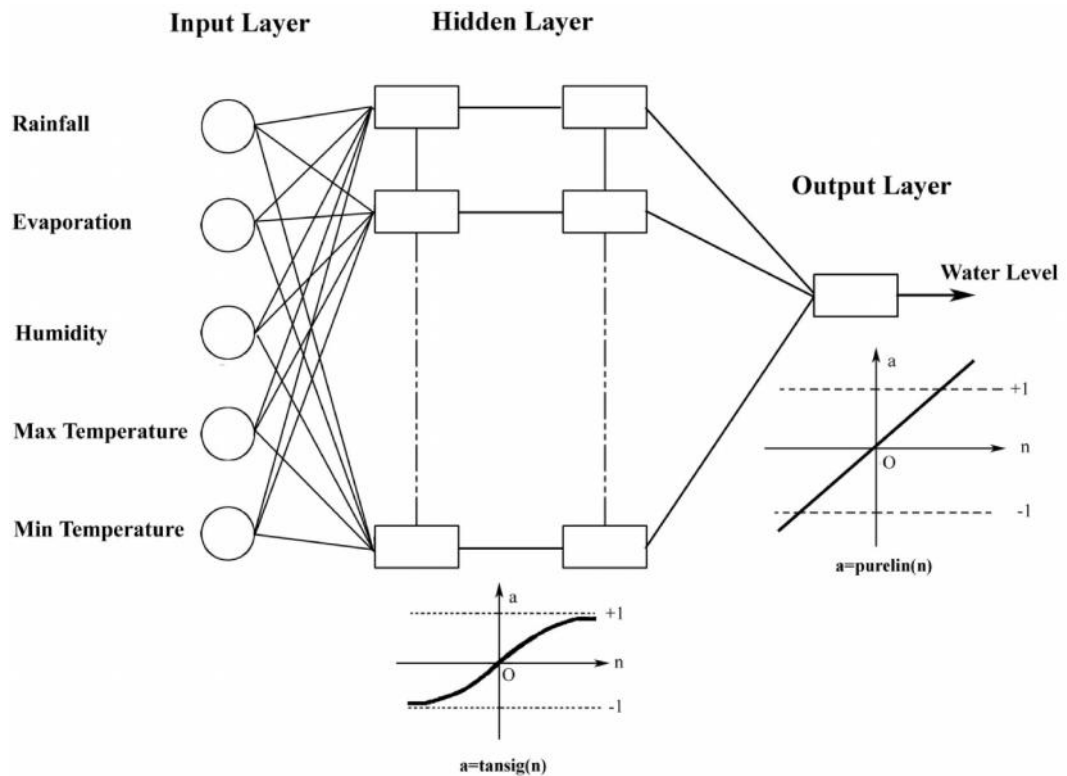


Figure 5.7: General conceptual neural network for the water level computation in the study area.

Table 5.2. Learning factors at the maximum R and minimum MSE of data set

| FNN with LM algorithm | Epoch | All | | Training | | Testing | | |
|-----------------------|-------|-----------|-----|----------|------|---------|------|-------|
| | | Structure | R | MSE | R | MSE | R | MSE |
| | 500 | (8-10-1) | 0.9 | 0.044 | 0.99 | 0.009 | 0.76 | 0.229 |

Table 5.3. Comparison of performance of ANN models developed for all, training and testing periods

| Method | Algorithm | Structure | Epoch | All | | Training | | Testing | |
|--------|-----------|-----------|-------|-------|-------|----------|-------|---------|-------|
| | | | | R | MSE | R | MSE | R | MSE |
| FNN | LM | (8,10,1) | 500 | 0.839 | 0.048 | 0.998 | 0.001 | 0.443 | 0.265 |
| | GDX | (8,10,1) | 500 | 0.672 | 0.087 | 0.648 | 0.097 | 0.645 | 0.087 |
| | SCG | (7,9,1) | 500 | 0.708 | 0.072 | 0.915 | 0.021 | 0.463 | 0.367 |
| | RP | (7,9,1) | 500 | 0.78 | 0.064 | 0.865 | 0.042 | 0.654 | 0.119 |
| RNN | LM | (8,9,1) | 500 | 0.761 | 0.086 | 0.956 | 0.05 | 0.477 | 0.314 |
| | GDX | (8,9,1) | 500 | 0.692 | 0.107 | 0.694 | 0.118 | 0.494 | 0.112 |
| | SCG | (8,10,1) | 500 | 0.694 | 0.104 | 0.696 | 0.109 | 0.488 | 0.125 |
| | RP | (8,10,1) | 500 | 0.737 | 0.073 | 0.828 | 0.059 | 0.669 | 0.084 |
| CFN | LM | (9,11,1) | 500 | 0.811 | 0.058 | 0.97 | 0.008 | 0.5 | 0.165 |
| | GDX | (9,11,1) | 500 | 0.629 | 0.088 | 0.754 | 0.066 | 0.526 | 0.19 |
| | SCG | (8,11,1) | 500 | 0.751 | 0.081 | 0.845 | 0.049 | 0.546 | 0.14 |
| | RP | (9,11,1) | 500 | 0.787 | 0.061 | 0.848 | 0.045 | 0.592 | 0.112 |

A recurrent neural network trained with the gradient descent algorithm was investigated as the most unsuitable network. This might be becomes more complex training algorithms are necessary than RNN (Coulibaly et al., 2001). The FNN with LM algorithm compared with the other techniques has generated a much higher correlation R, lower mean square error, MSE. The evaluation of the ANN model ability compared with the different networks and algorithms and referred to the training phase was shown in the form of a scatter plot in figures 5.8-5.10. This presents the simulation of groundwater level and also MSE error of each model in the training step to obtain the greatest input admixture in the form of hydrographs. Furthermore, Figures 5.11-5.16 displays the observed depths to groundwater and simulated ones per month and from 2007 to 2013 for each network and algorithm. Figures 5.11-5.16 also shows the depths to groundwater for all three networks by different training algorithms in compared with the observed groundwater levels. The obtained results indicate almost great correlation between these data sets, hence indicate the capability of ANN model for water level depth predicting in the specific conditions of this research. This ANN model as the water level prediction model was employed for this well site.

The absolute difference between the predicted and the observed model demonstrates variation of water level depths with FNN, CFN and RNN for seven years:

- Water level depths for FNN model vary from 0.000203 m for November 2011 with the LM algorithm to 0.750358 m for April 2007 with the SCG algorithm.
- Water level depths for CFN model vary from 0.000613 m for February 2008 with the LM algorithm to 0.840088 m for March 2013 with the SCG algorithm.
- Water level depths for RNN model vary from 0.028682 m for January 2007 with the LM algorithm to 0.813794 m for February 2013 with the GDX algorithm.

The actual and simulated groundwater levels for all of the networks are extremely well matched, as is displayed in these results. The FNN and CFN simulation with LM algorithm are closer to the related actual values in comparing with other networks, which is confirmed by the hydrographs and scatter plots. Better performance of these networks might be due to the structure of the networks.

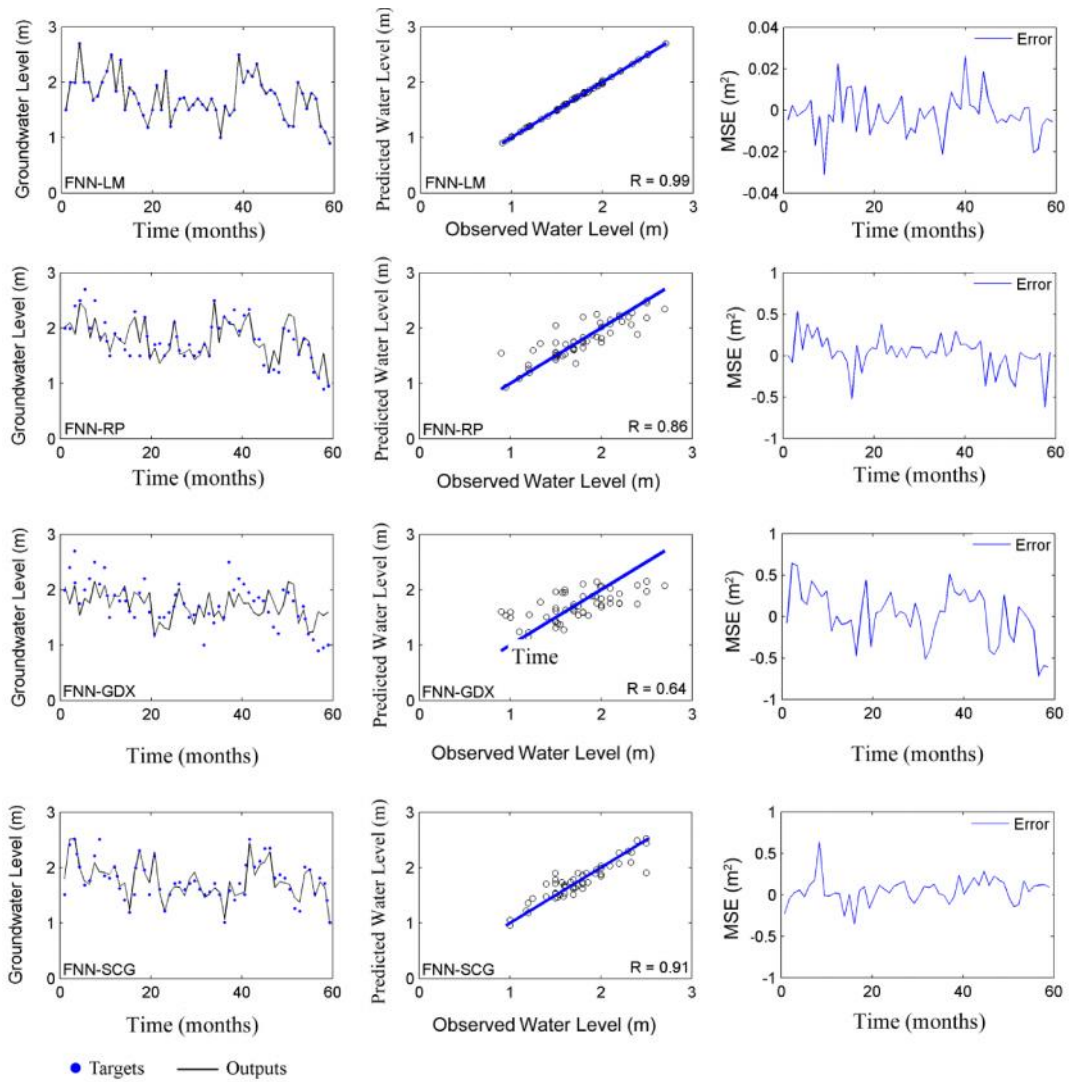


Figure 5.8: Scatter plots of the observed and simulated water levels at training period for FNN with different algorithms

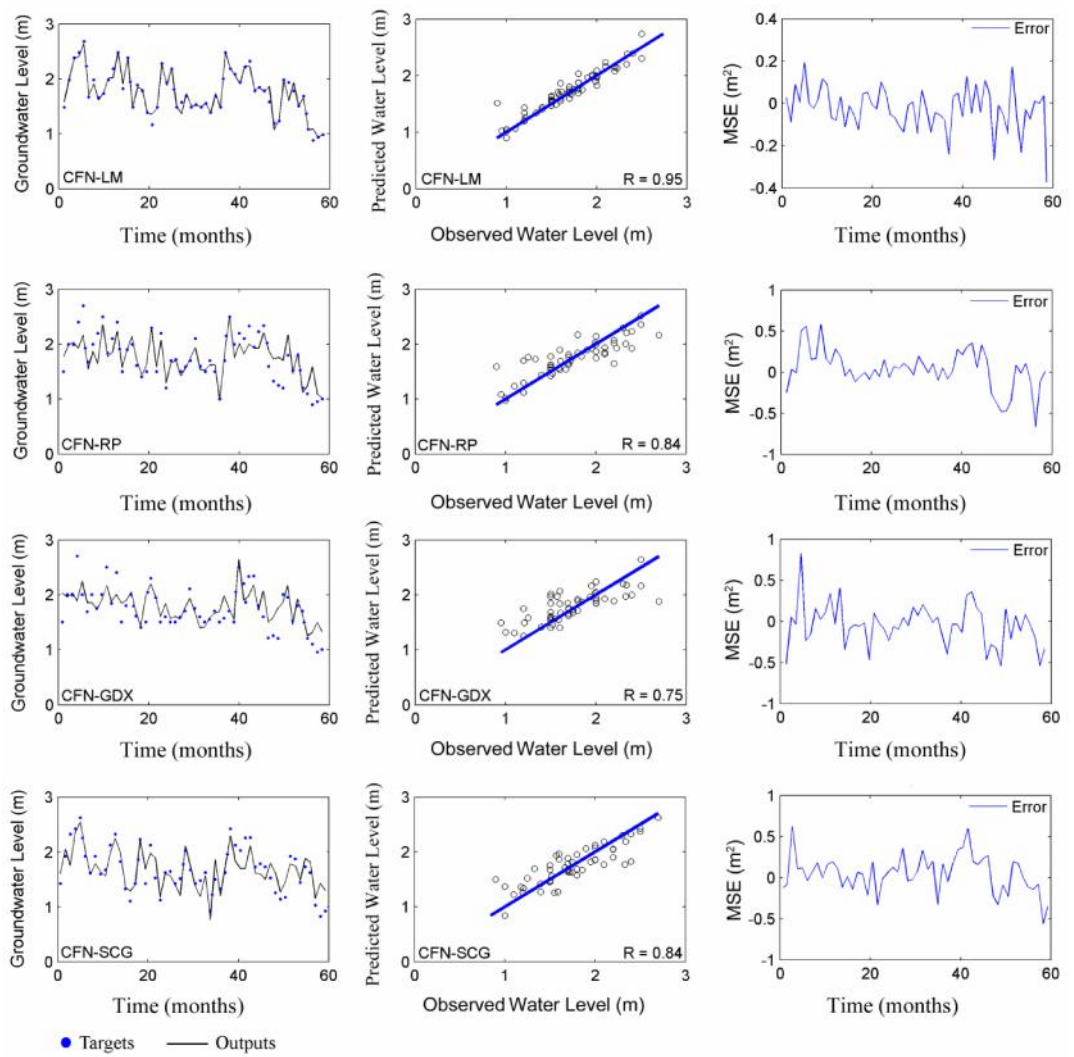


Figure 5.9: Scatter plots of the observed and simulated water levels at training period for CFN with different algorithms

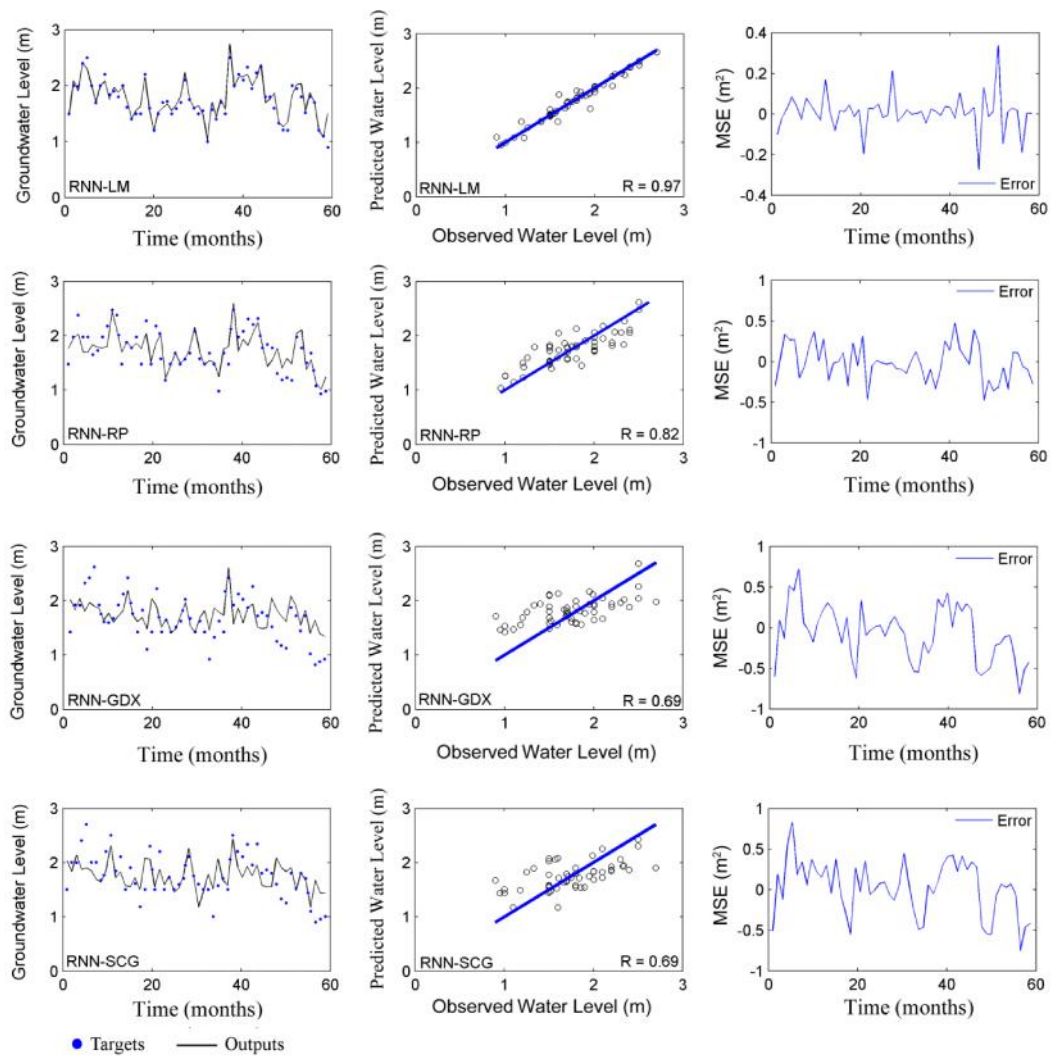


Figure 5.10: Scatter plots of the observed and simulated water levels at training period for RNN with different algorithms

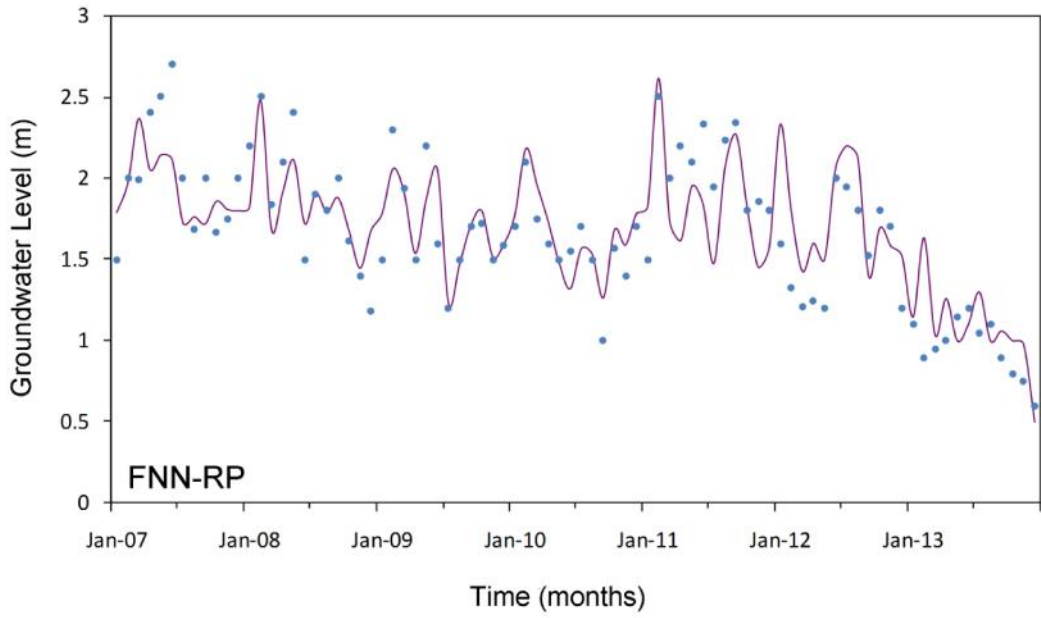
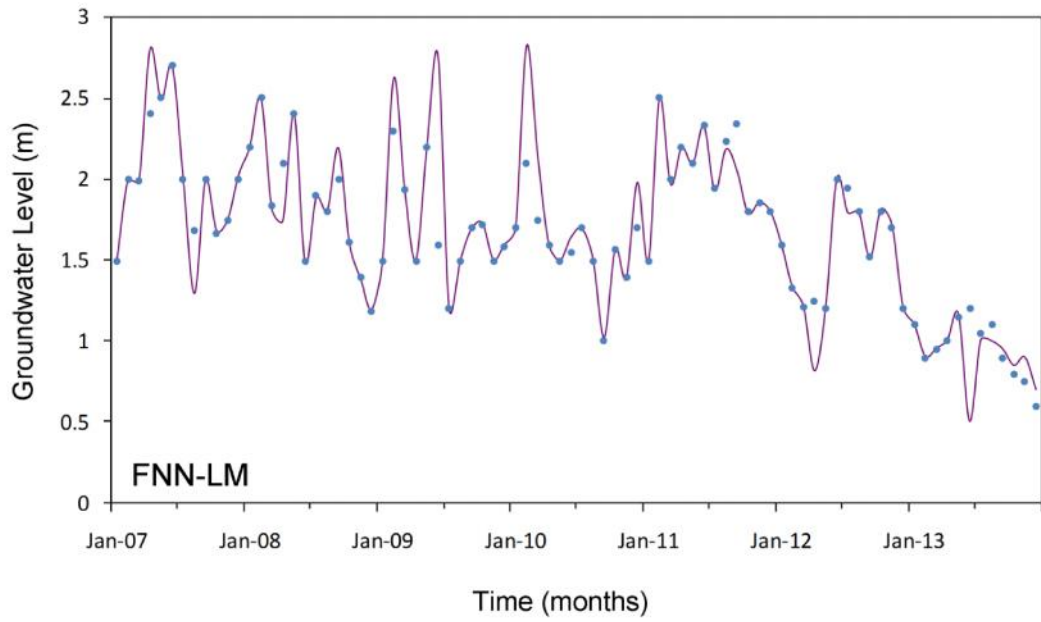


Figure 5.11: The result of the models and the actual value for the FNN approach with LM and RP algorithms.

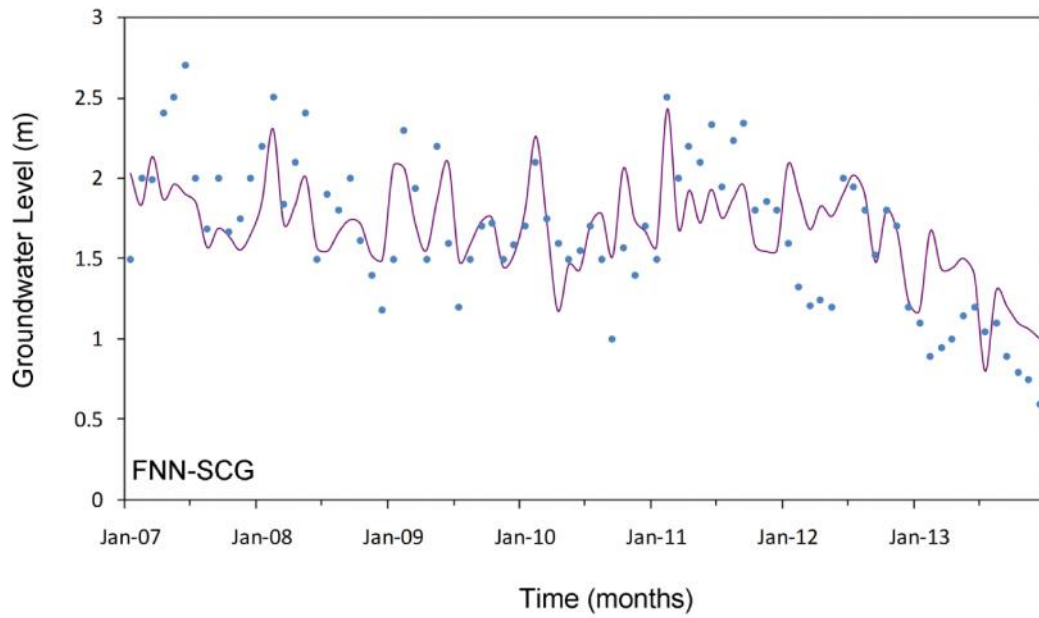
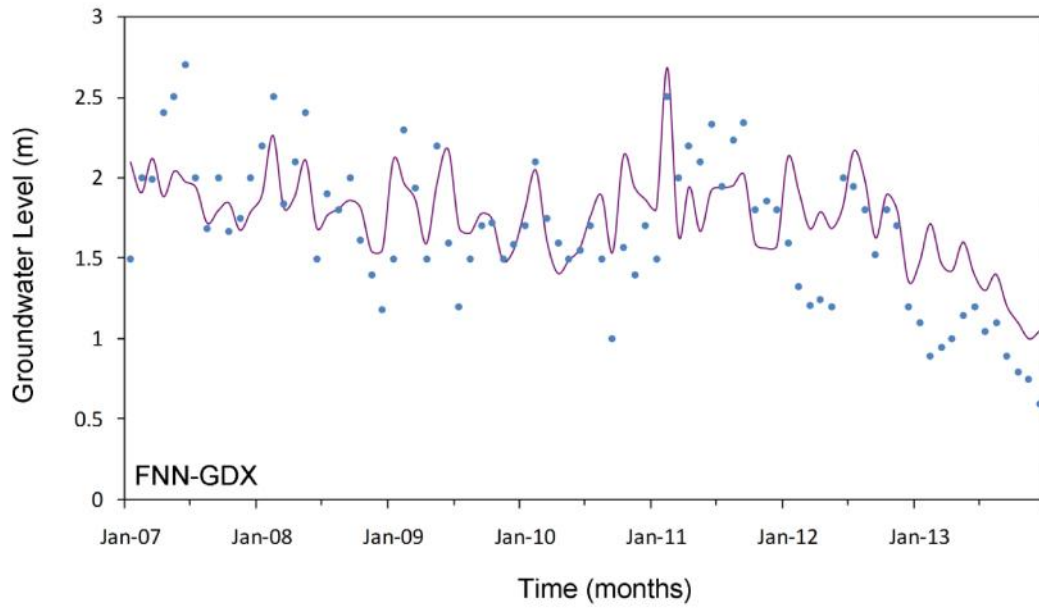


Figure 5.12: The result of the models and the actual value for the FNN approach with GDX and SCG algorithms.

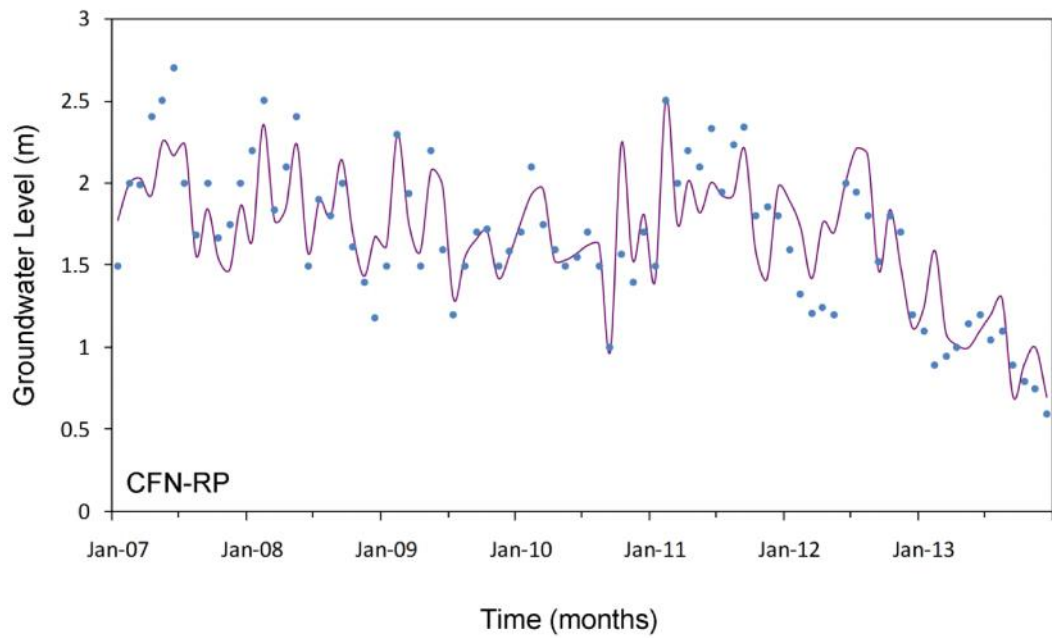
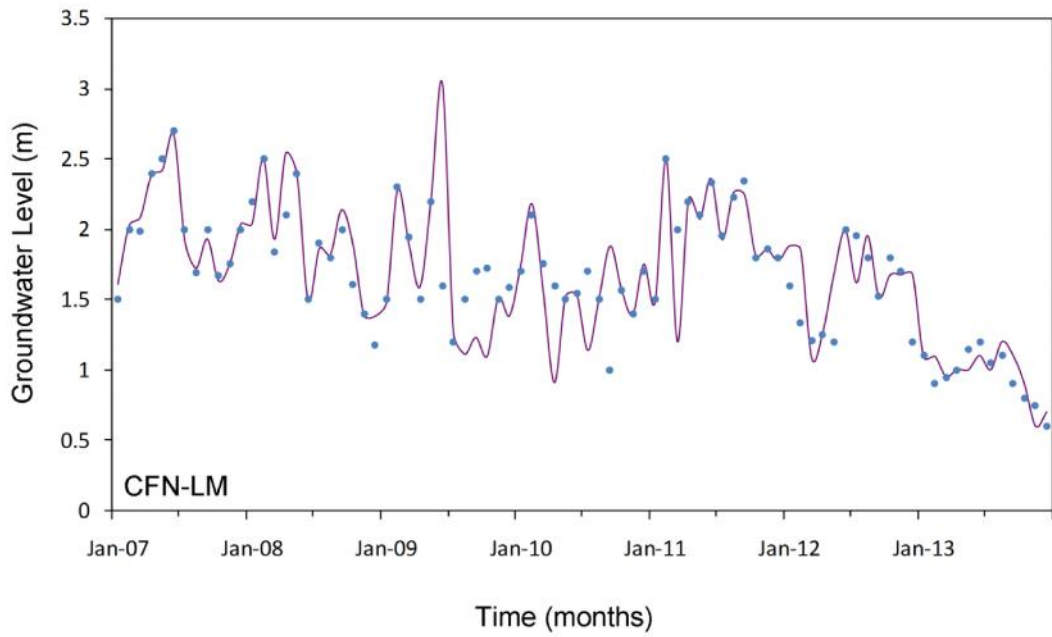


Figure 5.13: The result of the models and the actual value for the CFN approach with LM and RP algorithms.

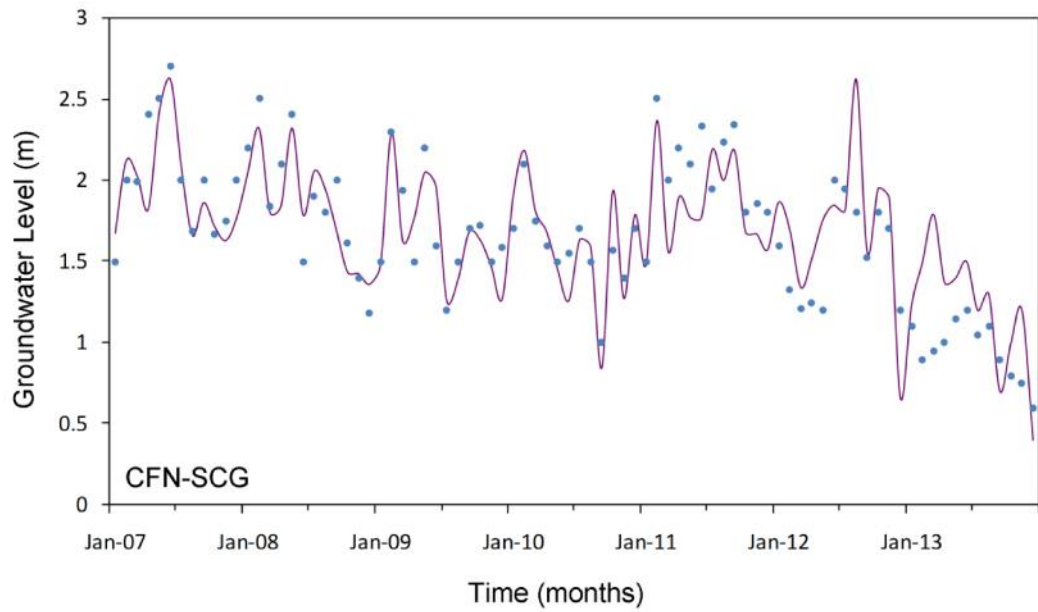
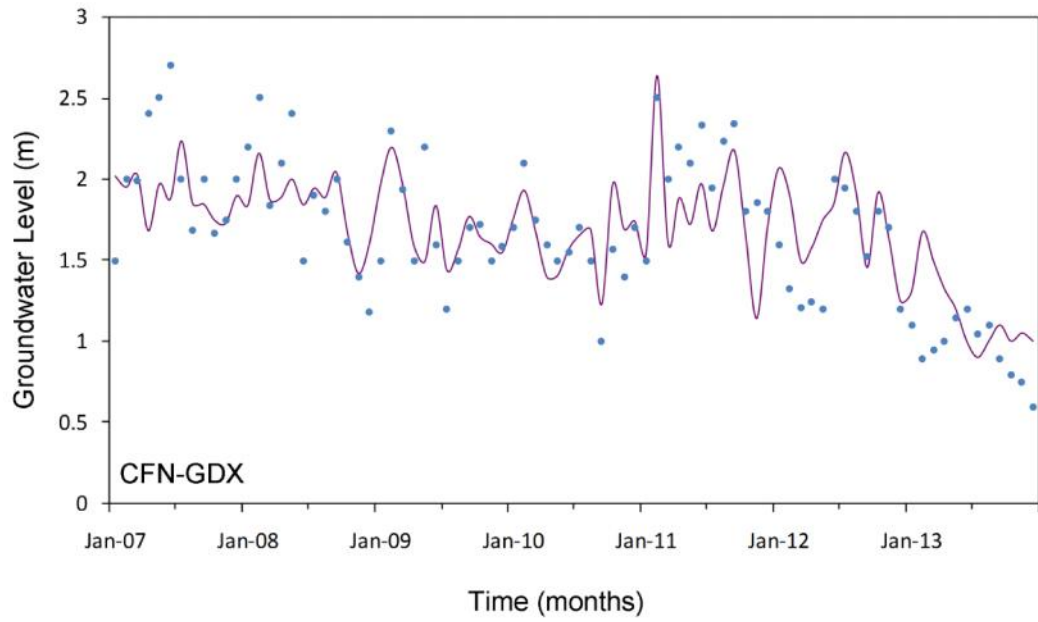


Figure 5.14: The result of the models and the actual value for the CFN approach with GDX and SCG algorithms.

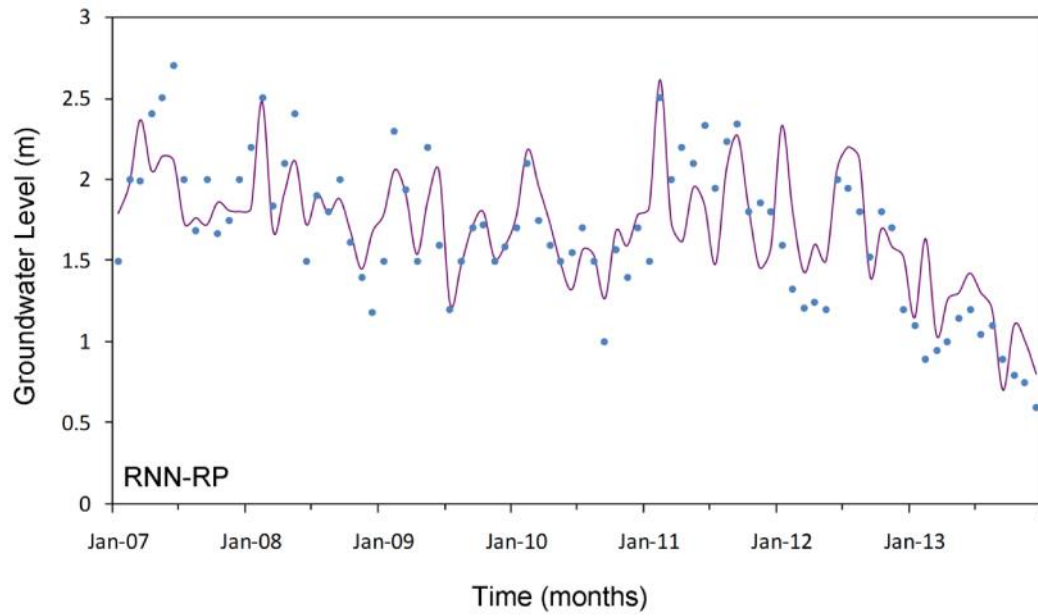
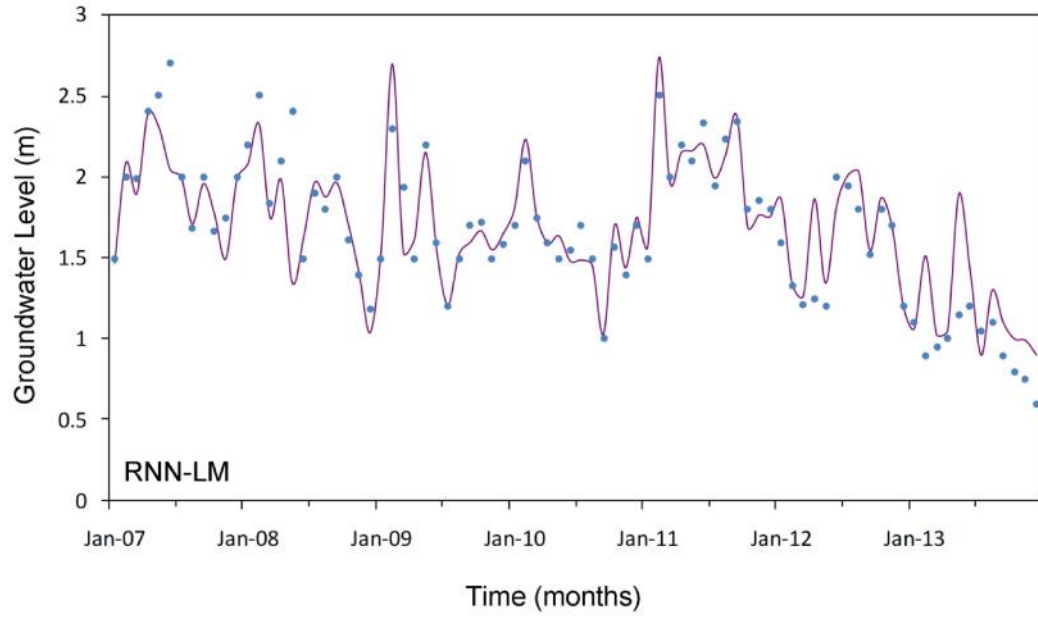


Figure 5.15: The result of the models and the actual value for the RNN approach with LM and RP algorithms.

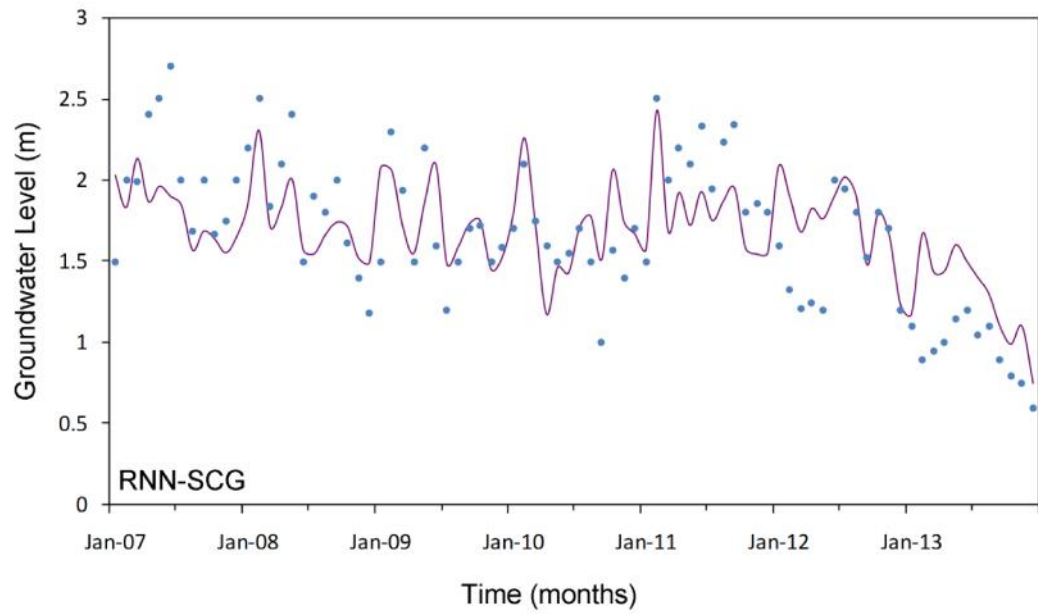
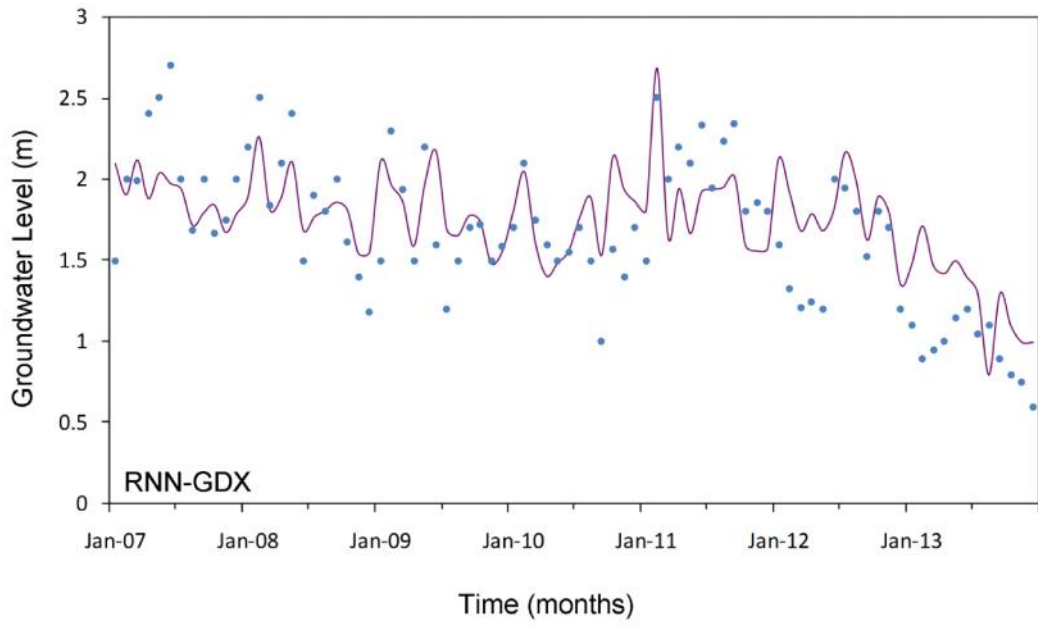


Figure 5.16: The result of the models and the actual value for the RNN approach with GDX and SCG algorithms.

5.2.5 ANFIS for the Prediction of Water Level

For prediction of water level fluctuations in the research area, an adaptive neuro-fuzzy inference system (ANFIS) model was developed. The training and testing of the network were carried out based on the same data employed in ANN model. Five input parameters were chosen to develop ANFIS which are rainfall, humidity, evaporation, minimum temperature and maximum temperature. The types of MFs were chosen Gaussian and generalized bell functions and number of MFs assigned to each input parameters are determined by trial and error, for ANFIS evaluation. The number, step size and shape of the MFs as predefined internal ANFIS elements were effective, making it possible for the ANFIS model to achieve the efficiency goal. Figures 3.22 illustrate the MFs for all input parameters. For all the models of the considered approaches, the network was trained for 500 epochs. The final topology of the network was obtained after considering the preprocessing, activation function, number of neurons, number of hidden layers, training algorithm, and so on.

5.2.6 Training and Testing Results for ANFIS Model

For obtaining a data set (input-output data pairs) and divide it into training and testing data sets are resulted in ANFIS modeling procedure starts. The data is normalized in order to make it appropriate for the training process which was carried out by mapping each term to a value between 0 and 1. For training the ANFIS, the normalized data was used as input and output data. The training data set is employed to obtain the primary assumption variables for the MFs by similarly spacing each of the MFs. An existence threshold value for the error between the real and desired output was identified. Using the least-squares approach, the consequent variables were found and consequently the error for each data pair was found. If this performance is larger than

the threshold value, the assumption variables employing the gradient decent method will be updated. Then to compare the model with real system, the testing data set was applied. If the model does not represent the system as a result, then a lower threshold value will be applied.

Table 5.4 presents a comparison between the predictions accuracy of the ANFIS and measured ones from field tests for all, training and testing data set. The obtained results from Table 5.4 in comparison with Table 5.3 interpret that ANFIS with the Gaussian membership function and generalized bell was efficiently comparable with the FNN, CFN and RNN in the entire data set. It clearly shows that ANFIS model performance provides improvement over that ANN models performed. The scatter plot shown in Figure 5.17 presents the analyses of the capability of the ANFIS models, compare to different MFs which refers to the training steps. This illustrates the simulation of groundwater level and also MSE error of each model in the training step in gaining the greatest input admixture in the form of hydrographs. The predicted values of the models and the real value were matched. The predicted output of ANFIS with two common membership functions which are Gaussian and generalized bell MFs, along with the corresponding actual values for all of the data were presented in Figure 5.18.

Table 5.4. Comparison of performance of ANFIS models developed for all, training and testing periods

| Method | Epoch | All | | | Training | | Testing | |
|------------------------------|-------|-----------|------|--------|----------|--------|---------|-------|
| | | Structure | R | MSE | R | MSE | R | MSE |
| ANFIS (Gaussian MFs) | 500 | 2 | 0.99 | 0.0076 | 0.99 | 0.0001 | 0.87 | 0.024 |
| ANFIS (Generalized bell MFs) | 500 | 2 | 0.99 | 0.0043 | 0.99 | 0.0001 | 0.9 | 0.018 |

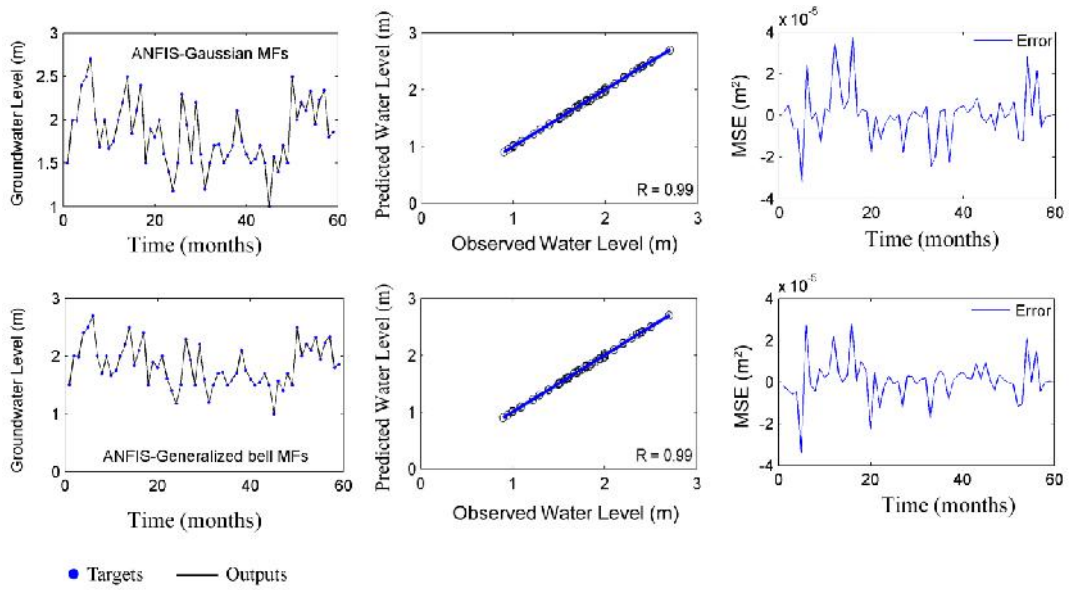


Figure 5.17: The result of the models and the actual value for the ANFIS approach with Gaussian and Generalized bell MFs.

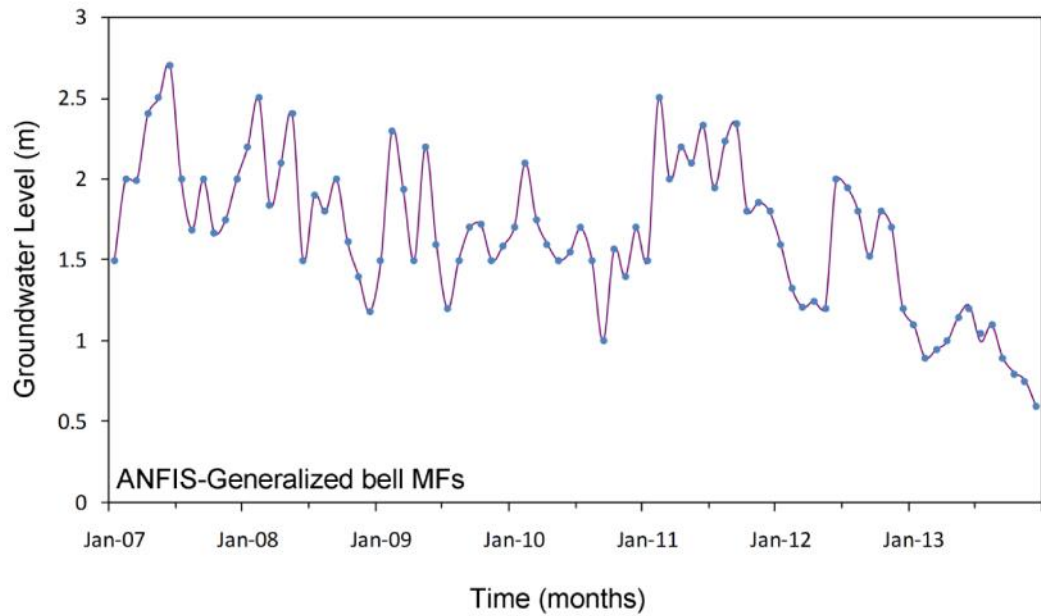
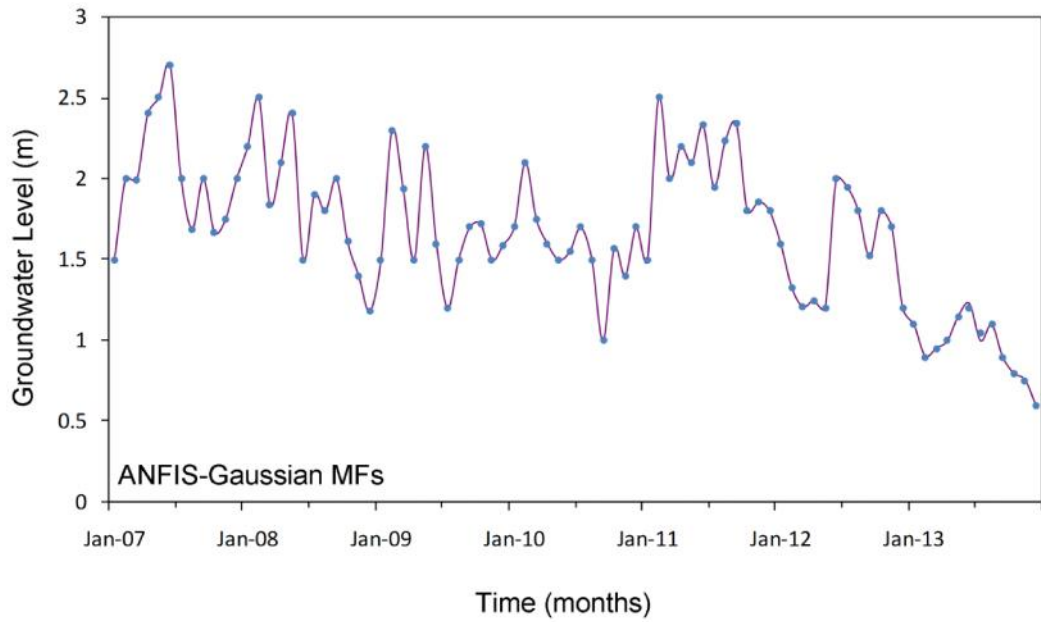


Figure 5.18: The result of the models and the actual value for the ANFIS with Gaussian and Generalized bell MFs.

Figures 5.11-5.16 and 5.18 show that the predicted values from ANFIS had the similar results with ANN (FNN, CFN and RNN), and matched the measured values much better than those obtained from the three methods. In general, for this research ANFIS is more beneficial compared to ANN since ANFIS model trains considerably

faster with supplying better predictions.

Table 5.3 and 5.4 illustrate the correlation R and mean squared error (MSE) of predicted versus measured results. A much higher correlation R and a lower mean squared error were generated by ANFIS in compare with the other methods, as shown in Table 5.3 and 5.4. It is indicated that the ANFIS predictions are much accurate in compare with other methods. Consequently, the ANN and ANFIS models can be efficiently applied to simulate the groundwater level fluctuations.

5.3 A Neural Network and Adaptive Neuro Fuzzy Inference System for Modelling of Groundwater Quality

In this section, ANN and ANFIS models are developed as forecasting tools to calculate of optimal total dissolved solid (TDS) and electrical conductivity (EC) concentration by using the values of other existing water quality parameters.

Groundwater quality modelling is essential to the development and operation of management projects. Protection of the water quality is a significant and complex engineering challenge. The main pollutants are raised from agriculture, intrusion of sea water and others activities in the research area as major point and non-point pollution sources. The essential knowledge of groundwater systems should be improved, when the ground water demand was increased by different activities that they threat the quality of groundwater. As a result, the simulation analyses of groundwater quality for common development situations are essential, based on importance of groundwater problem in water resource modelling. Furthermore, simulation of the water quality is a major element to manage the quality of water and a manager is capable to select an option that satisfies a large number of defined conditions.

Knowledge based systems and Expert known as most popular artificial intelligence (AI) tools are being used in a number of computer-based applications supporting real-time continuous water quality monitoring equipment. Recently, AI has supplied tools to select the equipment and optimize the operations, as well as solve the problems which involve the large amounts of data that a human is not easily able to cope with in the decision making process. A very favourable environment for the introduction of other powerful AI tools is created by AI systems together with an ever-increasing number of sophisticated purpose-built computer software packages. The ANNs and ANFIS techniques are well suited to the model complex problems where the model variables (input-output) have an unknown relationship.

5.3.1 The Water Quality Parameters

In this section of research the data set was collected through continuous monitoring of five wells in the study from 2008 to 2013 as it is shown in Figure 5.19. All water samples were collected within the aquifer; they were analyzed for 31 various parameters. The analyses of water quality were carried out in-situ applying a handy analyser, basic chemical assessment on site, and laboratory analyses. The parameters such as pH, conductivity, water level, temperature and total dissolved solids were identified in the field. 1 L sampling bottles were used to collect all samples and they were maintained in a refrigerator at 4°C and brought to the laboratory to measure the chemical analyses. Prior to the analyses, the standard approach of preservation was done to add a few drops of concentrated nitric or acetic acid which was done after the sample has been obtained and filtered to reduce the pH to 2 (Hiscock, 2009). One of the basic parameters in water quality measurements is known as TDS, which is the total amount of solids in mg/L. Total dissolved solids (TDS) and electrical conductivity (EC) were considered as the water quality parameters of interest; they were chosen because of their

importance in the Langat Basin. A high concentration of TDS was found at >1000 mg/L in the study area. The obtained data shows that TDS was highly influenced by the intrusion of sea water from the west of the study area. As a general water quality indicator, EC is associated with major water quality parameters because of the dilution effect of stream flow. In addition, TDS is considered as one of the principal factors and it determines the application of groundwater for any objective. The EC and TDS value ranges from 114 to 23500 $\mu\text{S}/\text{cm}$ and 70 to 11244 mg/L respectively. The presence of a high content salt in the water because of the intrusion of saline from the sea is a basic problem with EC. Lower EC and TDS concentrations have been found in the eastern and northern parts of the study area where the wells are not affected by seawater intrusion. A high correlation was found between EC and TDS, particularly in highly saline areas. The EC was found to be significantly correlated with TDS at five stations in the present study. Hence, these findings indicate the serious problems of the water treatment station placed in the area. Therefore, to model the two water quality parameters of TDS and EC at five boreholes in the study area, ANN and ANFIS techniques were developed. For these analyses, the data of thirteen water quality parameters consisting of Ca^{2+} , Na^+ , Mg^{2+} , K^+ , pH, Cl^- , HCO_3^- , TDS, EC, Fe, NO_3^- , SiO_2 and Turbidity (NTU) at the five boreholes for the time period 2008 to 2013 were selected.

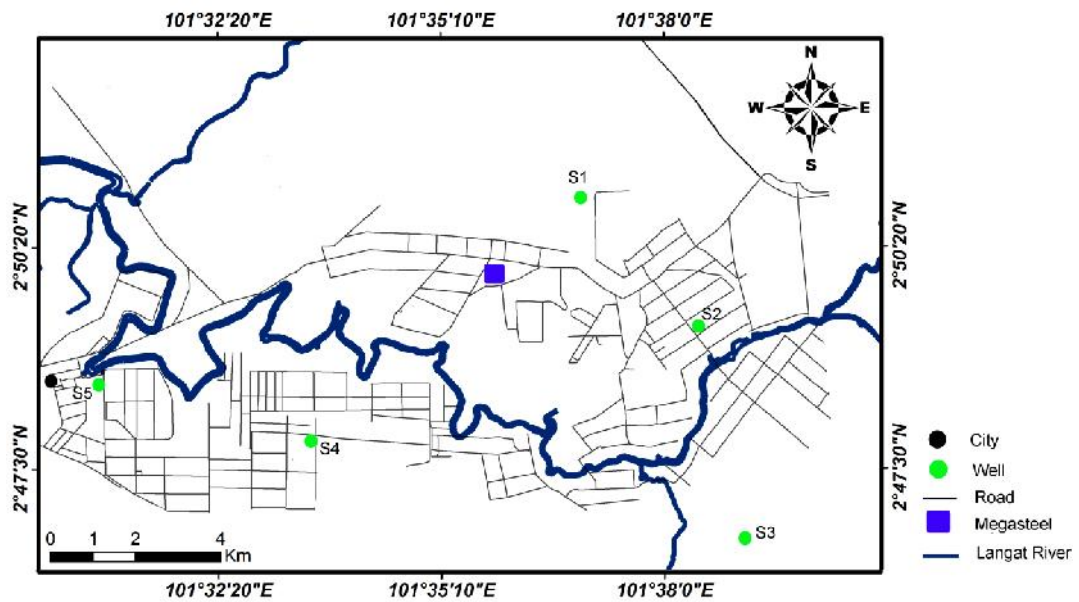


Figure 5.19: Location of the selected wells for modelling of groundwater quality (TDS and EC) in the study area.

5.3.2 Statistical Analyses

The Analyses of Variances (ANOVA) test is one of the most powerful mathematical tests for groundwater monitoring. One-way ANOVA analyses were carried out applying statistical analyses (SPSS) software. This technique is properly used to evaluate background data to compliance well data and is also useful for the conditions where constituents are detected in many wells at varying concentrations. Moreover, this technique is employed for comparing the means of various groups of observations (i.e., from various wells on the same sampling date) and for determining significance differences among the groups; the data can be further analysed to identify contrasting aspects of the data set, if there are significance differences. Either a parametric or a nonparametric method can be used for analyses performances. The summary of the ANOVA results for input parameters compared at 5% significance level and Least Significant Difference (LSD) test are indicated in Table 5.5.

5.3.3 Input Parameters

The network ability to generate sensible estimations is always influenced by selected input parameters, as a data-driven design. Adequate analyses of the variables managing the system analysed is consequently required to developing a reliable network. Regarding to the experiments and ANOVA analyses, the parameters of water quality were defined as input for the learning techniques. 70% of the data corresponds to train and the 30% to test samples for the experiments. Boundary of the eleven groundwater quality parameters as input parameters values were demonstrated in Table 5.6. They were measured at each sampling site and their statistical characteristics obtained from 2008 to 2013.

Table 5.5. Results of the statistical analyses to show the effect of the factors on samples properties

| | Sum of Squares | Degrees of freedom | Mean Square | F | Significance |
|----------------|----------------|--------------------|-------------|------|--------------|
| Between Groups | 19324251.03 | 10 | 819611.62 | 5.25 | 0 |
| Within Groups | 117302798.4 | 319 | 140819.15 | | |
| Total | 136627049.5 | 329 | | | |

Table 5.6. The statistics of geochemical parameter variables for five boreholes in the study area

| Geochemical parameters | Min | Max | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Skew | Kurtosis |
|------------------------|------|------|--------|----------------|------------|----------------------------------|-------------|------|----------|
| | | | | | | Lower Bound | Upper Bound | | |
| Ca | 3.50 | 293 | 21.80 | 54.41 | 9.93 | 1.48 | 42.11 | 4.63 | 22.94 |
| Mg | 2.30 | 549 | 67.48 | 149.93 | 27.37 | 11.49 | 123.46 | 2.78 | 6.51 |
| Na | 11 | 2802 | 358.4 | 743.67 | 135.77 | 80.70 | 636.09 | 2.63 | 5.81 |
| K | 5.20 | 234 | 33.27 | 52.95 | 9.66 | 13.50 | 53.05 | 2.59 | 6.77 |
| HCO3 | 1 | 2679 | 354.26 | 684 | 124.88 | 98.85 | 609.67 | 2.81 | 6.66 |
| Cl | 9 | 6118 | 807.46 | 1730.07 | 315.86 | 161.44 | 1453.48 | 2.62 | 5.61 |
| NO3 | 0.50 | 46 | 8.45 | 9.94 | 1.81 | 4.74 | 12.16 | 2.54 | 7.19 |
| pH | 6.60 | 9 | 7.62 | 0.79 | 0.14 | 7.32 | 7.92 | 0.33 | -1.17 |
| SiO2 | 7.10 | 93 | 22.14 | 19.37 | 3.53 | 14.91 | 29.38 | 2.40 | 5.94 |
| Turbidity | 1 | 209 | 30.23 | 43.93 | 8.02 | 13.83 | 46.64 | 2.80 | 9.35 |
| Fe | 0.70 | 58 | 10.96 | 13.85 | 2.52 | 5.78 | 16.13 | 2.68 | 7.23 |

5.3.4 Training and Testing Results

In this section the developed ANN and ANFIS were trained and tested employing the accessible techniques and data sets as discussed in the previous sections. In this section, two networks (FNN and CFN) with LM algorithm were resulted in the performed ANN models which have shown the more accurate result comparing with the other algorithms. The structure of three layers ANN model, which uses the eleven water quality parameters data set as inputs layer and two parameters (TDS and EC) data set as output layer have been shown in Figure 5.20. Figure 5.20 shows that each layer comprises a number of simple processing elements named neurons. ANFIS models were performed two membership functions which are Gaussian and Generalized bell MFs.

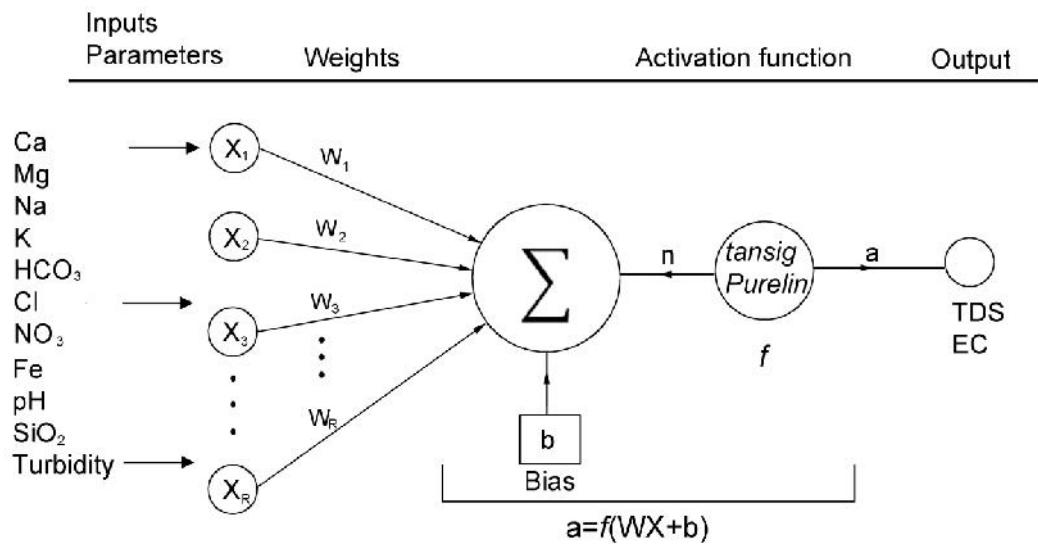


Figure 5.20: Non-linear model of a neuron

The data division into training and testing data set were done randomly from the chronologically arranged observations. 70% and 30% of the all of observation data were used by the training data and testing data set respectively. The obtained and the tests results were derived through a computer programming which is the Matlab R2012. This chapter contains the study of the of ANN and ANFIS capability to estimate TDS and EC which is based on another water quality parameters. To conceptualize the catchment behaviour, the developed models for water quality evaluation might be used. This clearly assists in assessing the relative importance of the parameters.

Table 5.7 shows that the total data set consists of training and testing data set which are corresponding to the R and the MSE, between the output of the hidden unit and the output of the ANN and ANFIS, for simulation EC and TDS. Accurate results were obtained from these variables comparing with other variables in the initial analyses. The MSE values for TDS and EC validated in the range of 0.054 to 0.096 and 0.031 to 0.096, for all of the data respectively with average of 0.073 and 0.064. In the training process the better prediction ability and generalization can be illustrated for

MSE and R, ANFIS using the Gaussian membership functions and the generalized bell membership functions comparing with the FNN and CFN methods as shown in Table 5.7. The evaluation of the ability of the ANN and ANFIS models, comparing with different networks and membership functions which referring to the training steps, are displayed in the form of a scatter plot (Figure 5.21 and 5.22). This shows the groundwater quality simulation and MSE error of every model in the training phase to achieve the best input combination in the form of hydrographs. The outputs of the ANN and ANFIS water quality at the five sampling stations were compared over the period 2008 to 2013 (Figure 5.23. and 5.24); all indices demonstrated similar results and were sensitive to the level of water quality parameter variations. However, the values of water quality proposed in this research generate a more stringent output. Table 5.7 and Figures 5.21-5.24 show the ability of the ANN and ANFIS models for simulating of the concentration values of EC and TDS at the Langat Basin was perfect. Consequently, the ANN and ANFIS models can be successfully employed for the simulation of water quality parameters. The ANFIS simulation was found more accurate than FNN and CNN simulation.

Table 5.7. Comparison of performance of models developed for all, training and testing periods

| Total Dissolved Solid | | | | | | | | |
|------------------------------|-----------|-------|------|-------|----------|-------|---------|-------|
| Method | Structure | Epoch | All | | Training | | Testing | |
| | | | R | MSE | R | MSE | R | MSE |
| FNN | (8-10-1) | 500 | 0.99 | 0.08 | 1 | 0.024 | 0.98 | 0.54 |
| CFN | (8-11-1) | 500 | 0.99 | 0.096 | 1 | 0.055 | 0.97 | 0.67 |
| ANFIS (Gaussian MFs) | 2 | 500 | 0.99 | 0.062 | 1 | 0.007 | 0.98 | 0.024 |
| ANFIS (Generalized bell MFs) | 2 | 500 | 0.99 | 0.054 | 1 | 0.009 | 0.98 | 0.011 |
| Electrical Conductivity | | | | | | | | |
| Method | Structure | Epoch | All | | Training | | Testing | |
| | | | R | MSE | R | MSE | R | MSE |
| FNN | (8-10-1) | 500 | 0.99 | 0.096 | 1 | 0.038 | 0.87 | 0.21 |
| CFN | (9-11-1) | 500 | 0.99 | 0.094 | 1 | 0.03 | 0.97 | 0.11 |
| ANFIS (Gaussian MFs) | 2 | 500 | 0.99 | 0.037 | 1 | 0.004 | 0.98 | 0.017 |
| ANFIS (Generalized bell MFs) | 2 | 500 | 0.99 | 0.031 | 1 | 0.015 | 0.99 | 0.01 |

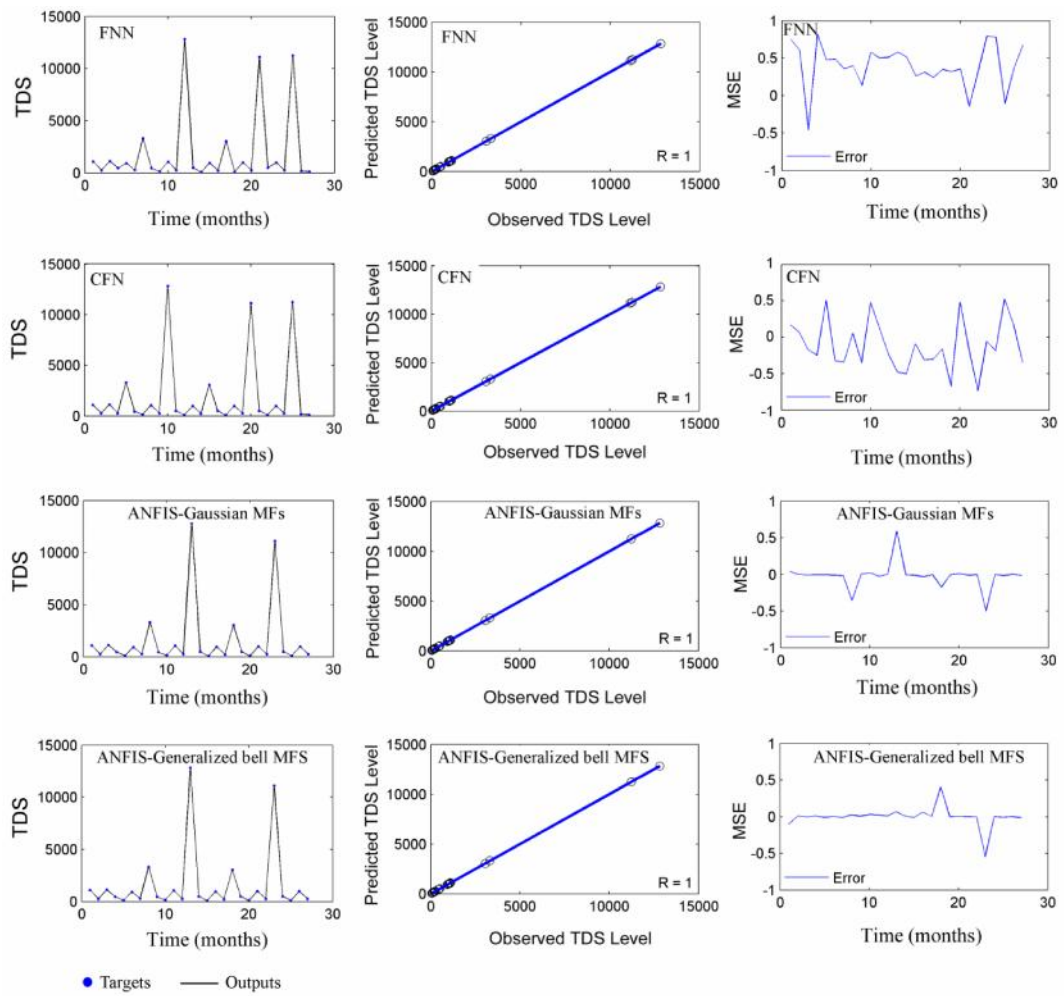


Figure 5.21: Scatter plots of the observed and forecasted water quality at training period for TDS.

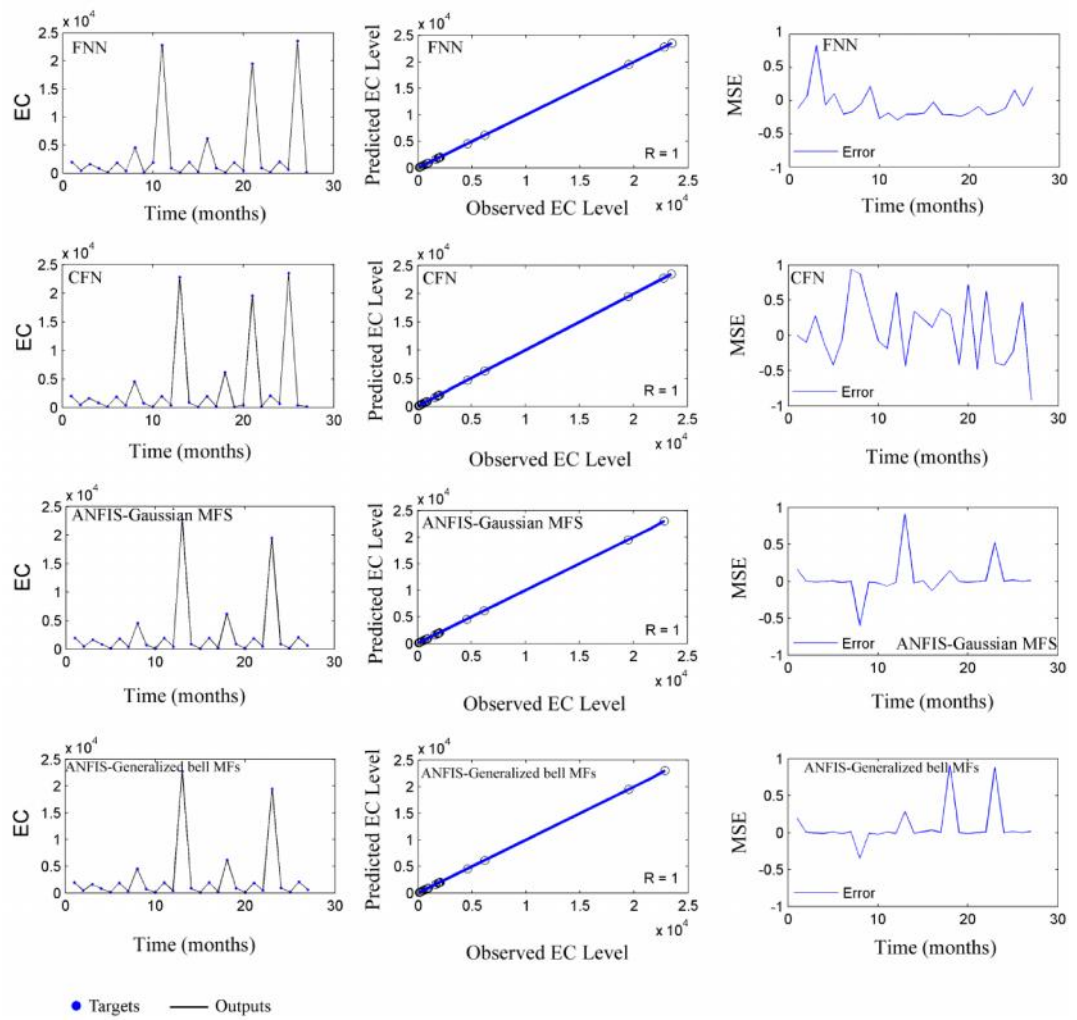


Figure 5.22: Scatter plots of the observed and forecasted water quality at training period for EC.

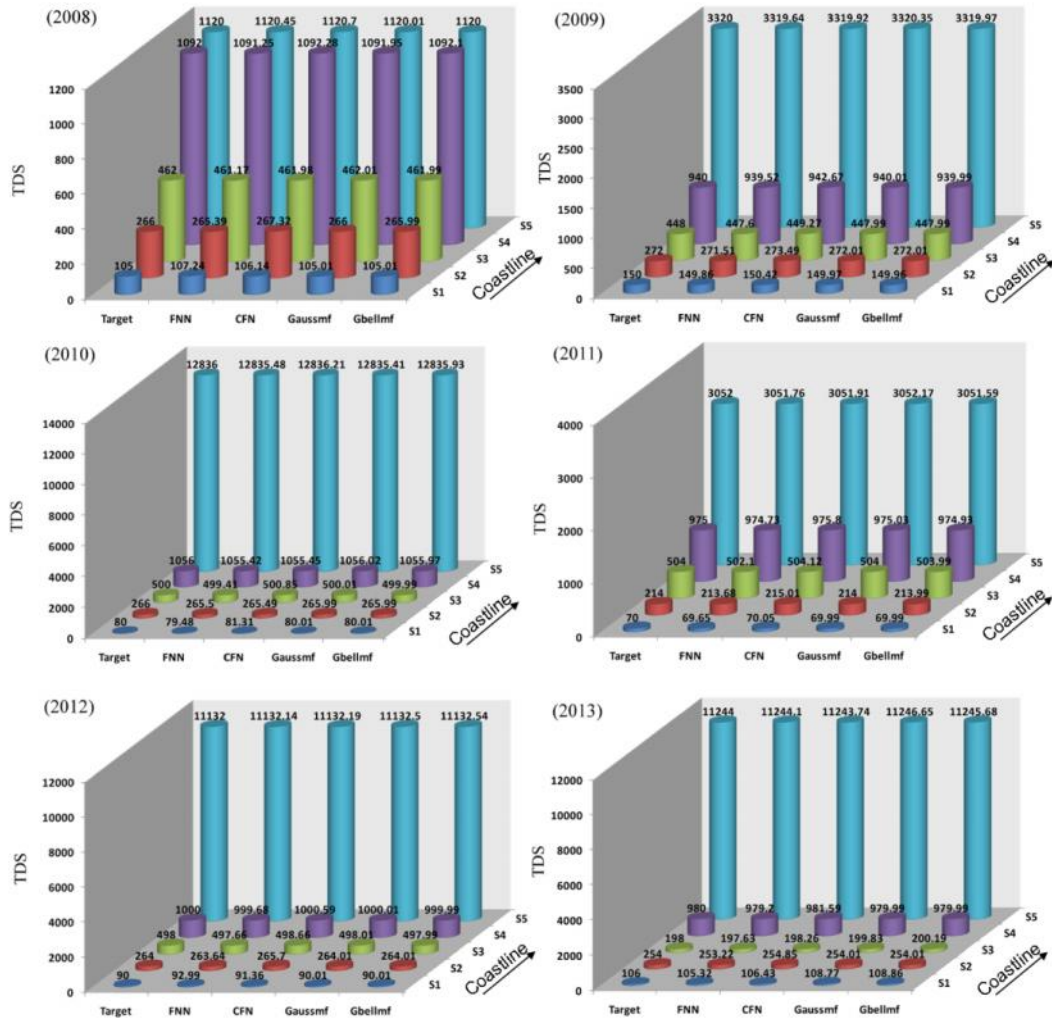


Figure 5.23: Comparison of results to observed TDS quality with different networks

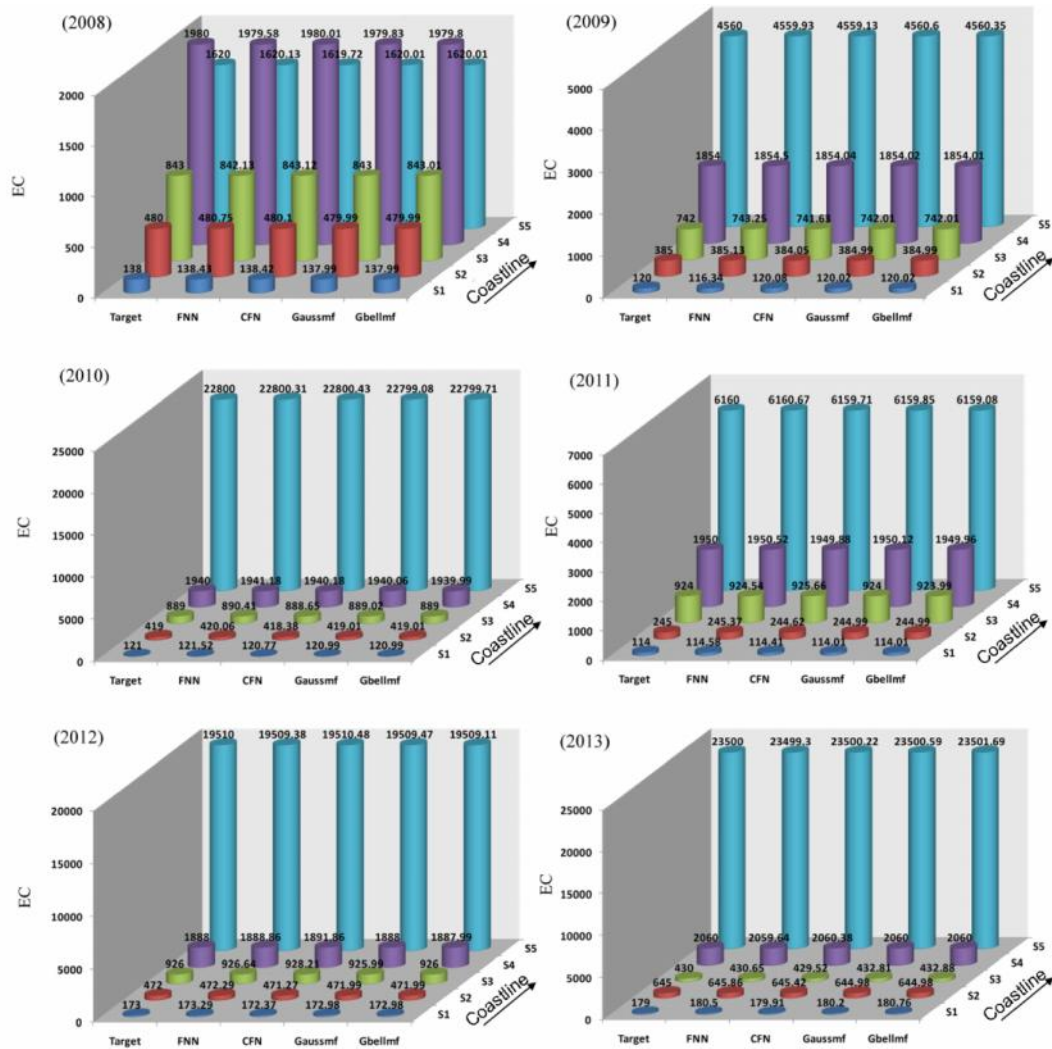


Figure 5.24: Comparison of results to observed EC quality with different networks

5.4 Prediction Groundwater Level and Quality with Experimental Data

In the present section of this chapter, the obtained results from previous section were used for predicting the water level, TDS and EC for a period of 2014 to 2025 in the study area. Therefore, “Forecast” function and “Linear Trend” in Microsoft Excel software are used to create the artificial data with the uncertainty of 15%. The rainfall, min and max temperature, humidity and evaporation factors which were discussed in section 5.2 and Table 5.6 are used for modelling the ground water level and water

quality parameters, respectively. To create the artificial data the ground water pumping by Mega steel company is assumed to be constant ($23,300 \text{ m}^3/\text{day}$). In addition, the slope of the area is found to be less than 5%. Two architectures (Feed-Forward Neural Network and Cascade Forward Network) with Levenberg-Marquardt for neural network and two Gaussian and generalized bell MFs for ANFIS were applied for modelling due to more accurate obtained result in the previous section.

Figure 5.25 present the obtained result to predict ground water level over 12 years, from 2014 to 2025. Consequently the comparison between the results from various models identifying the general trend of ground water level is gradually decreasing in the study area.

Sodium and chloride are the parameters most affecting TDS and EC. The prediction is carried out upon increasing of water qualities parameters specially these two factors from the west of the study area for 2008 to 2013 and continues trend. The prediction of the value of EC and TDS with ANN and ANFIS methods is illustrated in Figures 5.26 and 5.27 for 2019 and 2025. Figures 5.26 and 5.27 which displays approximate similar results for ANN and ANFIS and defines increasing of TDS and EC values is enhanced over the time gradually.

Figure 5.28 shows the boundary of TDS values for 2008 and 2013 and predicted TDS values of 2019 and 2025. TDS values are less than 1000 mg/L in the right side of the defined boundary line whereas the values are more than 1000 mg/L on the left side. It confirms that the intrusion of sea water to aquifer is increasing over the time from 2008 to 2025. The distance of specified TDS boundaries line was found to be 10 km, 13.6 km, 14.7 and 16.4 from the coastal line for 2008, 2013, 2019 and 2025 respectively. It is predicted that the boundary of drinking and brackish water reach to 16.8 km far from the sea.

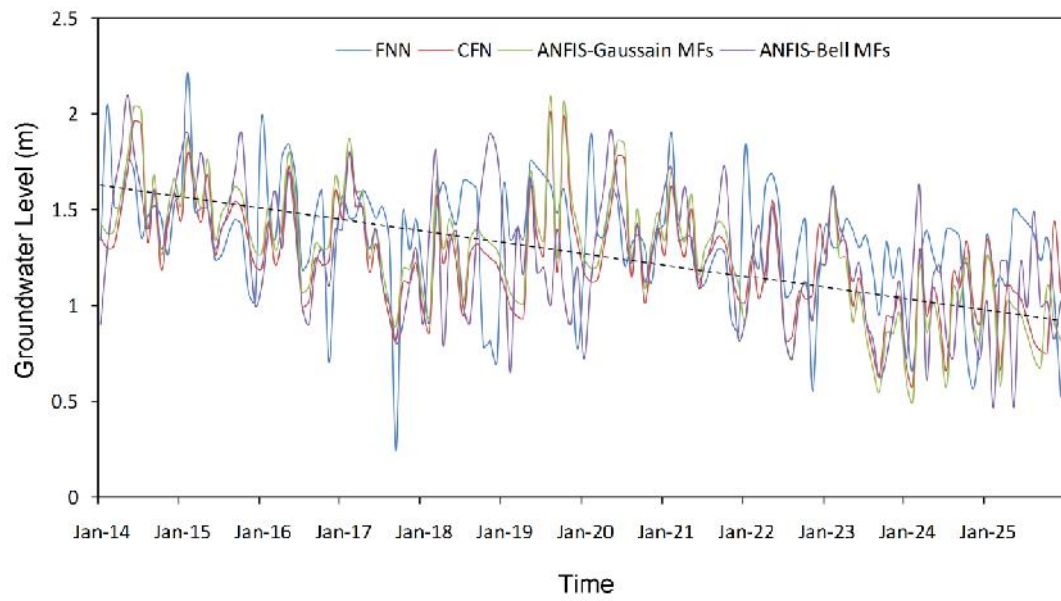


Figure 5.25: The prediction of water level in the selected well for FNN and CFN with LM algorithm and ANFIS ANFIS with gaussian and generalized bell MFs from 2014 to 2025.

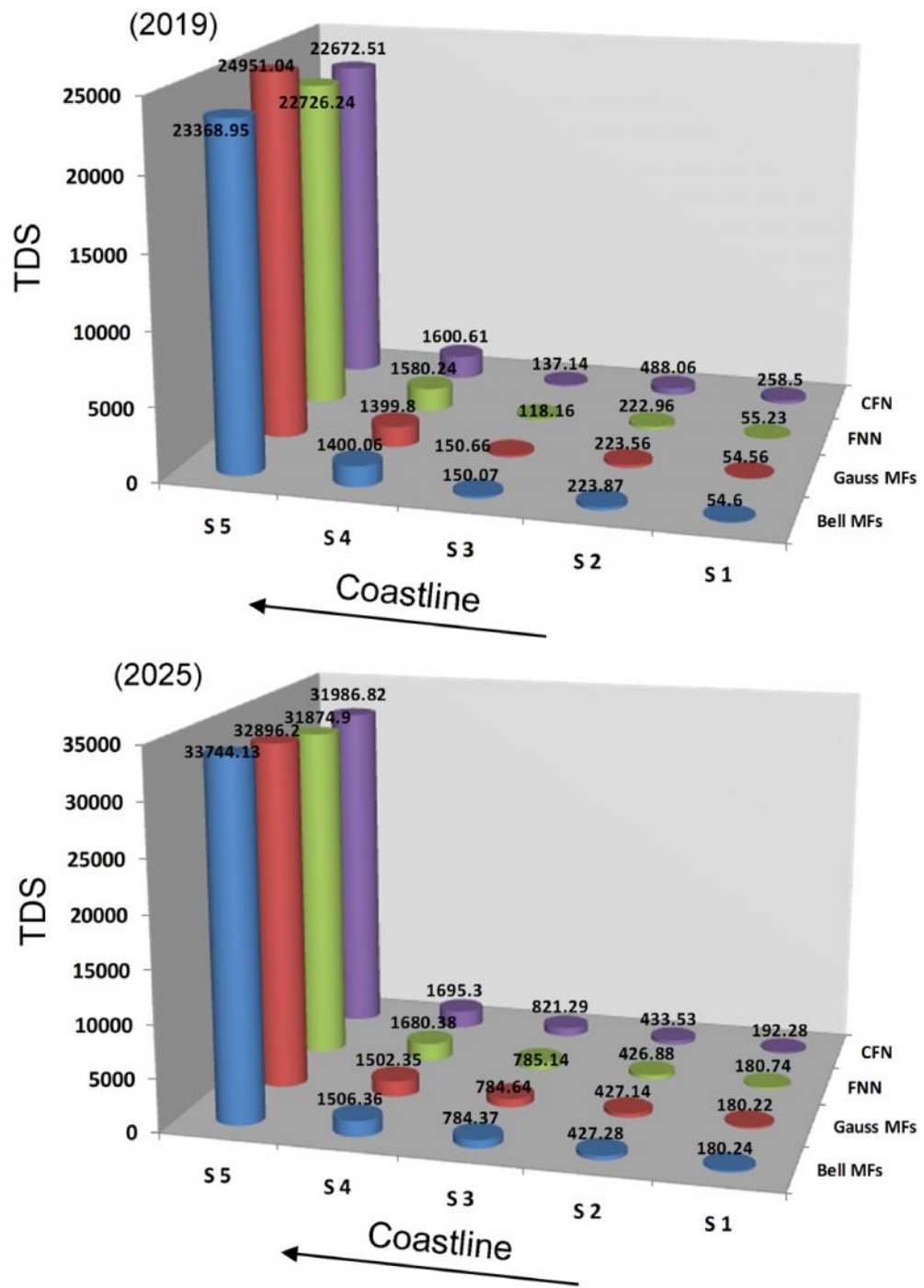


Figure 5.26: Predicted TDS results to various models for 2019 and 2025

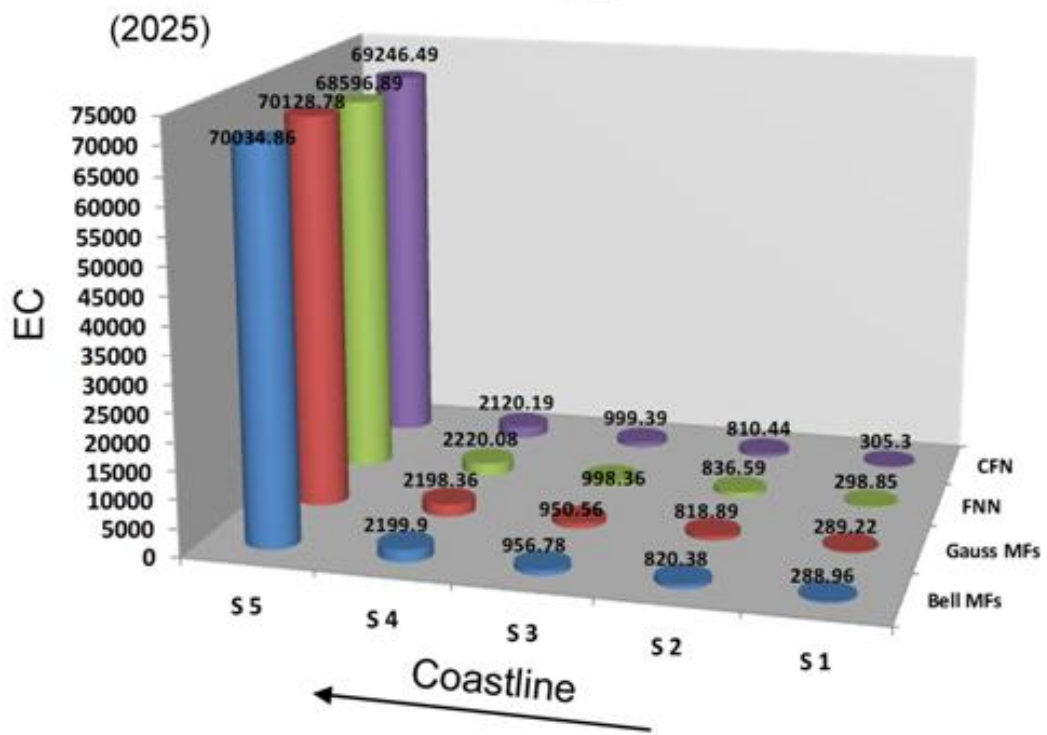
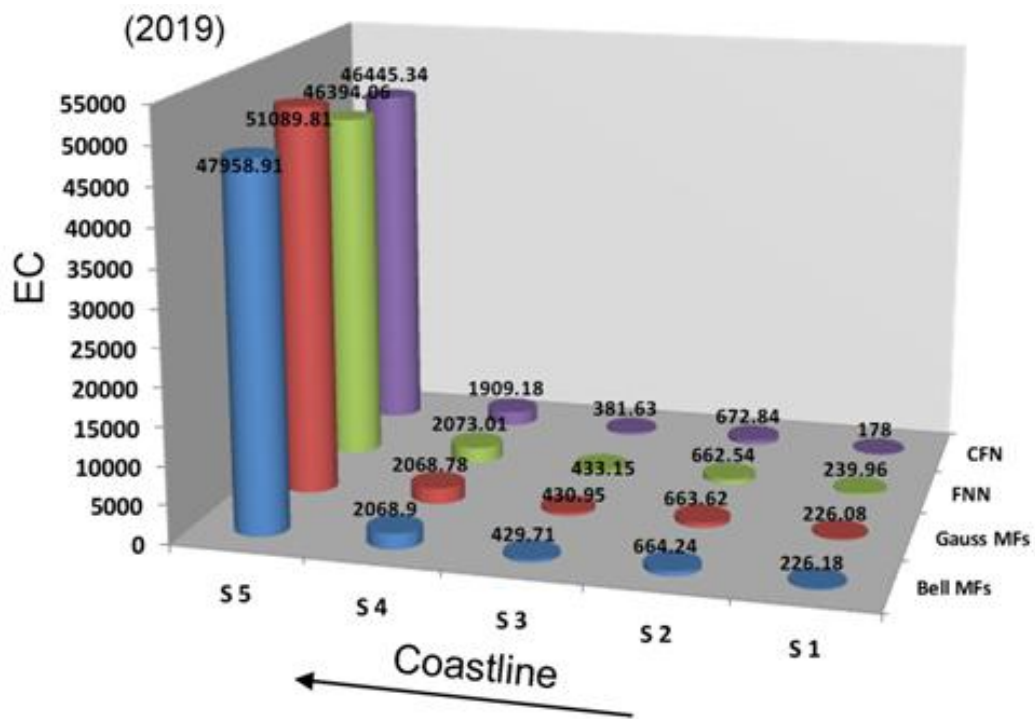


Figure 5.27: Predicted EC results to various models for 2019 and 2025

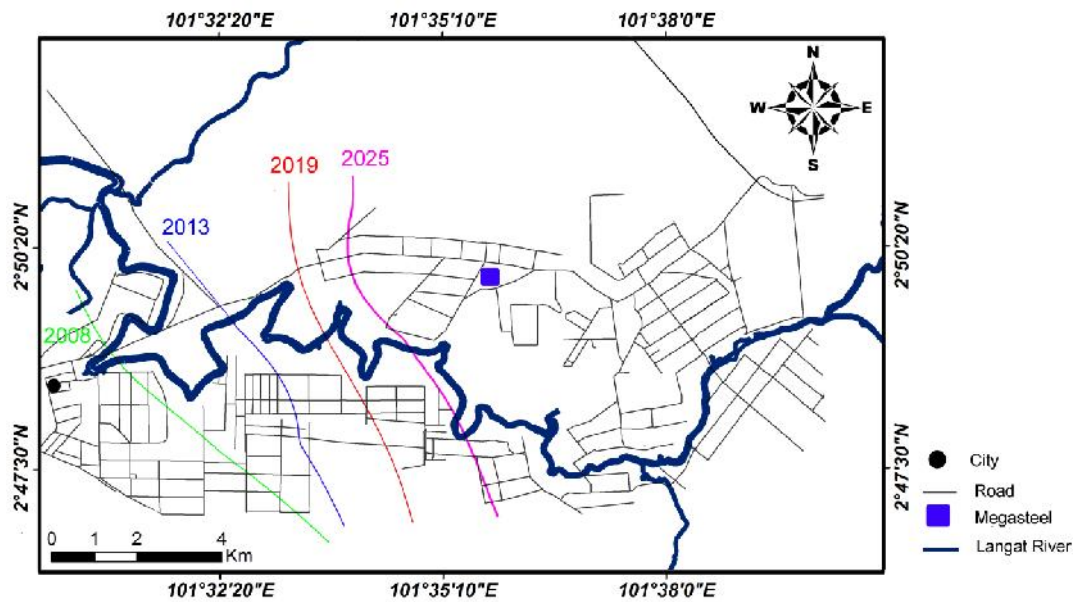


Figure 2.28: Defined boundray line of TDS for 2008, 2013 and predicted TDS line for 2019 and 2025. Right and left side of boundray 1000 mg/L refer to less and more than 1000 mg/L.

5.5 Summary

Developing a computational intelligence model to improve the water level and quality prediction in the high risk area was discussed in this chapter. In the present research, the models developed based on the ANN and ANFIS calculation methods are extremely satisfying in simulating the variations in monthly groundwater in Langat Basin, Malaysia. The selected methodology used to evaluate the performance of different architectures, algorithms and membership functions for determination of the most appropriate predictive model. In general, both models were done very well in prediction of water level and quality data. The performance of these methods has been examined using a database of groundwater variances covering a seven-year period. In this study, the Matlab as the most appropriate software was selected to process ANN and ANFIS application technique. Indeed this is very reliable and widely tested in ANNs and ANFIS. The obtained result for the training and testing of the data sets, shows that all the models result in reasonable agreement with the observed data. The

most suitable structure for ANN proved to be a 9-10-1 and 9-11-1 feed forward network trained with the LM algorithm as it demonstrated the most precise predictions of the groundwater level. However, the ANFIS models show the best performance in comparing the correlation coefficient R and MSE, followed by the FNN, CFN and RNN models, respectively.

The simulated depth for groundwater in all the observation data, with MSE of 0.0043–0.107 m², was successfully represented by the obtained results. The minimum reported error belonged to the ANFIS model. In comparing the fewer computational complexities in training the network and quicker convergence to a solution, ANFIS shows the better result than FNN, CFN and RNN. Besides, the accuracy of the groundwater level simulation with minimum calculations compared to the FNN, CFN and RNN is improved by the ANFIS model.

The performance evaluation criteria, MSE and the correlation coefficient were analyzed on a six years database of water quality parameters. Moreover, results successfully represent the network's simulated value to groundwater quality in all observation data with a mean square error (MSE) of 0.031–0.096 m². An ANOVA analyses and descriptive statistics allowed the production of a sensitivity analyses of the selected input for the learning techniques. The outcomes for the training and testing of data sets, obviously illustrate the capability of the ANN and ANFIS model to simulate the values of TDS and EC very well, utilizing the values of the other existing water quality parameters that are presented as inputs. The ANFIS model with generalized bell function with an MSE of 0.031 – 0.054 m² showed the best performance and it has been determined as the best fitting model compared to the performance of the other models. The ANFIS model resulted in the lowest value of MSE. Consequently, all modelling approaches were satisfactory. The results demonstrated that ANFIS requires fewer computational complexities to train the network and it converges to a solution faster

than the FNN and CFN. Moreover, it improved the precision of water quality parameter simulation with minimal computation compared to the FNN and CFN.

In General, the outcomes of the research are acceptable and show that neural networks and neuro fuzzy system can be a helpful tool for simulation in the case of groundwater hydrology studies. Consequently, the investigation demonstrated in this research supports us with a reliable tool to understand the various management strategies and how various climatic tools cause groundwater levels and quality to respond. These tools can be applied to determine the optimal policies that can promote sustainable management of accessible water resources when complete data on the hydrological system is not available.

Regarding to the obtained results from 2008 to 2013, ground water level and quality such as TDS and EC characterizations, for 2014 to 2025 have been predicted using FNN and CFN architectures with LM algorithm and ANFIS model with Gaussian and Generalized bell MFs. Concluding from the previous section, the predicted model with ANFIS model is more suitable than ANN method.

Chapter 6: Conclusions and Recommendations

6.1 Conclusions

Integrated geoelectrical resistivity, hydrogeochemical and artificial intelligence methods were used to investigate groundwater and associated problems in the Langat Basin, south-western Selangor, Malaysia. The research involved the acquisition, processing, modelling and interpretation of the electrical resistivity, hydrogeochemical, artificial neural networks and adaptive neuro fuzzy inference system datasets. The problems that the this study project is involved with, the goals and objectives, the methodologies, test of hypotheses and the achieved results have been discussed in detail in the previous chapters. The results and discussions are briefly discussed together with conclusions with principle research objectives, in this chapter.

The findings of this dissertation provided an approach which assists in better defining the groundwater resources available for distribution on Langat Basin.

The ERI technique is chosen as the most suitable geophysical approach for imaging the subsurface structure. A cost-effective and rapid approach was used in this investigation to achieve wide area subsurface information. Site accessibility, portability, non-invasiveness, fast investigating speed, operator safety, and good resolution are the other significant advantages of ERI.

Data combination of resistivity surveys and hydrogeochemical analyses demonstrates more accurate depictions and interpretations of subsurface structure in the Langat Basin.

Electrical resistivity Imaging datasets were obtained at the research area. The measured apparent resistivities were inverted to generate a map of the possible true resistivities of the subsurface structure. The information concerning the subsurface

structure was estimated to obtain by the ERI technique and its associated processes based on the resistivity field distribution which was resulted from the inverted image.

The depth of bedrock, thickness and depth of aquifer were imaged by resistivity profiling with Wenner electrode arrays which has shown acceptable accuracy. In general, the maximum depth of investigation with Wenner array was approximately 65 m for this research. In the research area, the bedrock was clearly imaged due to the fact that the resistivity surveys can map the shape of aquifer geometry. The basement depth has changed in specific region. In the interpretation of 2D resistivity imaging has been estimated that the variation aquifer thickness is from 10 to 30 m (and up to 40 m) with the depth ranging from 10 to 65, and the depth of the bedrock is between 30 m and 65 m. East to west of the study area depth and thickness of layers increased gradually but in the direction of north to south there is not significant variation and thickness and depth roughly remains constant. Resistivity model supply comparatively superb correlation involving the resistivity pseudosection and lithological data which is obtained from boreholes. As a result of 2D resistivity modelling, the ERI technique is an acceptable method for practical application to shallow subsurface structure determination.

In this study, an improving geoelectrical interpretation for content of salt/brackish in the aquifer has been carried out. It was concluded that resistivity differences may unequivocally detect in the research area, based on the statistical analyses of the inverted resistivity datasets. In general, higher resistivity values were obtained at the upper part of the Langat River comparing with the lower part of Langat river which was affected by intrusion of sea water in the aquifer from the west of the study area. The value of resistivity reduced drastically when the ratio of sea water in pore soil increased. Geoelectrical model through the resistivity value is capable predict a degree of percentage of sea water.

Monitoring the groundwater level of existing wells demonstrates the overall water flow is into the southwest in the study area. In the central part of the study area, water level flow has a different trend due to wells which are located close to a Mega Steel Company with extensive water withdrawals from the aquifer.

The depth slice visual images of resistivity distribution successfully mapped the zone with high percentage of sea water mixture in the aquifer. The resistivity value increased to the landward and decreased to the seaward. Seaward is located in the west of the area while the percentage of sea water has decreased into the east. The obtained results from water chemical analyses support this interpretation

The ANN and ANFIS models at a single well site are capable to forecast water level. Therefore, the performance of these models, benefits from greater correlations between the input and target variables. Using appropriate variables in the model is a key to achieve at successful ANN and ANFIS modelling. The correct prediction of water level fluctuations at a well site is the advantage of employing ANN and ANFIS water level prediction models. These ANN and ANFIS models are capable to be used for running the past data to fill up the missing period of the water level depth measurements at a well site. As a result, much more complete data sets for water resource research can be provided by these models. This can be helpful for farmers, environmental engineers, and water resource planners to get the better way to foresee the groundwater movement and feasible water level fluctuation in order to take essential action in a more effective way.

Regarding to groundwater level prediction in training the ANN, three types of architecture with four algorithms and two types of membership functions (MFs) for ANFIS models were employed to produce the synthetic data. These architectures are called Feed-Forward Neural Network (FNN), Elman Neural Network (RNN) and

Cascade Forward Network (CFN) four common training algorithms, i.e., Levenberg-Marquardt (LM), Gradient descent with momentum and adaptive learning rate back-propagation (GDX), Resilient back propagation (RBP) and Scaled conjugate gradient (SCG) for ANN model and two common MFs, i.e., Gaussian and generalized bell MFs have been compared for water level fluctuations. The developed models were trained and tested on data sets which were collected from years 2007 to 2013 with different desired target and number inputs.

Regarding to predicted model, the ground water level is gradually decreasing.

In order to predict the parameters of the groundwater quality such as total dissolved solids and electrical conductivity in training the ANN two types of architecture which are FNN and CFN with LM algorithm and two types of MFs which are Gaussian and generalized bell MFs for ANFIS models, were used. They generate the synthetic data from aquifer in the 5 boreholes to cover the study entirely. The improved models were trained and tested on data sets that were collected from 2008 to 2013 with different number inputs and desired targets. Real field data sets were used to check the ability of trained ANN and ANFIS for data simulation.

Based on the modelling of 5 selected wells from 2014 to 2025 it is concluded that the groundwater quality hazard can be increased in the west of the study area. Hence, average values of EC and TDS reach at 47972 $\mu\text{S}/\text{cm}$, 23429 mg/L for station no 5 respectively in 2019. Whereas, average values of EC and TDS are 69501 $\mu\text{S}/\text{cm}$ and 32625 mg/L in 2025.

The type of MFs and the number of MFs are essential factors to build the ANFIS architecture. These factors are selected empirically by trial and error.

ANFIS models are greatly affected by the number of MFs for each input variable comparing to the number of training data.

One of the ways in order to avoid over-fitting for ANFIS models, is using small number of MFs first because a greater number of MFs leads to a small error on the training set although does not necessarily produce a small error on the test set.

For both ANNs and ANFIS models, visual results are well consistent with mean square error (MSE), and correlation coefficient (R) can be generally used as a good indicator.

ANFIS and ANN provide similar predictions to simulate TDS and EC data and also they can serve as a simple, reliable and predictive tool for prediction of water quality. Furthermore, ANFIS may be more efficient than FNN and CFN as ANFIS trains much faster than FNN and CFN while they provide similar results. Consequently, ANFIS models found to be more accurate and desirable for simulating purposes.

In general, this research has managed to integrate several computational and modelling methods that would prove to be appropriate for different institutions which are directly involved in the water resource management in Malaysia.

6.2 Recommendations

New approaches in the application of geophysical, hydrogeochemical and artificial intelligence techniques in investigation of groundwater have been opened up by the research in the south-western of Selangor State, Malaysia. The following approaches are recommended for the upcoming studies.

The proposed method (ERI) has been employed and tested and the result has proven the efficiency of this method. Consequently, using this technique in other sites with various geological structures is useful to confirm the applicability of this technique.

The present research included of part of Langat Basin around the Langat River where possibility was affected by sea water intrusion from the western part. Consequently, applying the proposed probability of other parts of Langat Basin suggested that it creates a probability image for the sea water intrusion.

In this study, statistical analyses have been employed to predict the groundwater quality and level. Moreover, Genetic Expression Programming (GEP) can be applied to many geosciences problem including those covered in the current research

The applied methods have been successful for mapping the aquifer, chemical monitoring, detecting the bedrock, predicting water quality and predicting the groundwater level. In future studies, a similar approach could be used to the other location that they will lead to novel results in the Peninsula Malaysia.

Regarding to lacking data in the other area in Langat Basin, similar study is strongly recommended creating the conceptual model in the out of the study area.

Reference

- ABEM. (2010). Instruction Manual Terrameter SAS 4000 / SAS 1000 LUND Imaging System.
- Abu-Hassanein, Z. S., Benson, C. H., & Blotz, L. R. (1996). Electrical resistivity of compacted clays. *Journal of Geotechnical Engineering*, 122(5), 397-406.
- Anderson, D., & McNeill, G. (1992). Artificial neural networks technology. *Kaman Sciences Corporation*, 258, 13502-13462.
- Antanasijevi , D., Pocajt, V., Povrenovi , D., Peri -Gruji , A., & Risti , M. (2013). Modelling of dissolved oxygen content using artificial neural networks: Danube River, North Serbia, case study. *Environmental Science and Pollution Research*, 20(12), 9006-9013.
- Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME*, 146(1), 54-62.
- Ay, M., & Kisi, O. (2011). Modeling of dissolved oxygen concentration using different neural network techniques in Foundation Creek, El Paso County, Colorado. *Journal of Environmental Engineering*, 138(6), 654-662.
- Barker, R. (1989). Depth of investigation of collinear symmetrical four-electrode arrays. *Geophysics*, 54(8), 1031-1037.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*: Oxford university press.
- Boghici, R. (2003). A field manual for ground-water sampling. *Texas Water Development Board, Austin, TX*.
- Boyd, C. (2002). Standardize terminology for low-salinity shrimp culture. *Global Aquaculture Advocate*, 5(5), 58-59.
- Boyd, C. E. (2000). *Water quality: An introduction*: Springer.
- Campanella, R., & Weemees, I. (1990). Development and use of an electrical resistivity cone for groundwater contamination studies. *Canadian Geotechnical Journal*, 27(5), 557-567.
- Chang, F.J., & Chang, Y.T. (2006). Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 29(1), 1-10.
- Cherkassky, V., Krasnopolsky, V., Solomatine, D. P., & Valdes, J. (2006). Computational intelligence in earth sciences and environmental applications: Issues and challenges. *Neural networks*, 19(2), 113-121.
- Cosenza, P., Marmet, E., Rejiba, F., Jun Cui, Y., Tabbagh, A., & Charlery, Y. (2006). Correlations between geotechnical and electrical data: A case study at Garchy in France. *Journal of applied geophysics*, 60(3), 165-178.

- Coulibaly, P., Anctil, F., Aravena, R., & Bobée, B. (2001). Artificial neural network modeling of water table depth fluctuations. *Water Resources Research*, 37(4), 885-896.
- Dayhoff, J. E. (1990). *Neural network architectures: an introduction*. New York: Van Nostrand Reinhold Co.
- De Vita, P., Agrello, D., & Ambrosino, F. (2006). Landslide susceptibility assessment in ash-fall pyroclastic deposits surrounding Mount Somma-Vesuvius: Application of geophysical surveys for soil thickness mapping. *Journal of applied geophysics*, 59(2), 126-139.
- Dedecker, A. P., Goethals, P. L., Gabriels, W., & De Pauw, N. (2004). Optimization of Artificial Neural Network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium). *Ecological Modelling*, 174(1), 161-173.
- Delleur, J. W. (2010). *The handbook of groundwater engineering*: CRC press LLC.
- Dey, A., & Morrison, H. (1979). Resistivity modeling for arbitrarily shaped three-dimensional structures. *Geophysics*, 44(4), 753-780.
- Drever, J. I. (1997a). *The geochemistry of natural waters: surface and groundwater environments* (3rd ed.): Prentice Hall.
- Drever, J. I. (1997b). *The geochemistry of natural waters: surface and groundwater environments*.
- Dreyfus, G. (2005). *Neural networks: methodology and applications*: Springer.
- Edwards, L. (1977). A modified pseudosection for resistivity and IP. *Geophysics*, 42(5), 1020-1036.
- El-Qady, G., & Ushijima, K. (2001). Inversion of DC resistivity data using neural networks. *Geophysical Prospecting*, 49(4), 417-430.
- Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2), 179-211.
- Farmaki, E. G., Thomaidis, N. S., Simeonov, V., & Efstathiou, C. E. (2013). Comparative use of artificial neural networks for the quality assessment of the water reservoirs of Athens. *Journal of Water Supply: Research & Technology-AQUA*, 62(5).
- Fausett, L. (1994). *Fundamentals of neural networks: architectures, algorithms, and applications*: Prentice-Hall, Inc.
- Filik, U. B., & Kurban, M. (2007). A New Approach for the Short-Term Load Forecasting with Autoregressive and Artificial Neural Network Models. *International Journal of Computational Intelligence Research*, 3(1).
- Freeze, R. A., & Cherry, J. A. (1977b). *Groundwater*: Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Gallant, S. I. (1993). *Neural network learning and expert systems*: MIT press.

- Grasby, S., Hutcheon, I., & Krouse, H. (1997). Application of the stable isotope composition of SO₄ to tracing anomalous TDS in Nose Creek, southern Alberta, Canada. *Applied geochemistry*, 12(5), 567-575.
- Günther, T. (2004). *Inversion methods and resolution analysis for the 2D/3D reconstruction of resistivity structures from DC measurements*. University of Mining and Technology, Freiberg
- Hagan, M. T., Demuth, H. B., & Beale, M. (1996). *Neural network design*. PWS, Boston, MA.
- Hagan, M. T., & Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *Neural Networks, IEEE Transactions on*, 5(6), 989-993.
- Han, H.G., Chen, Q.I., & Qiao, J.F. (2011). An efficient self-organizing RBF neural network for water quality prediction. *Neural networks*, 24(7), 717-725.
- Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*: MIT press.
- Haykin, S. (1999). *Neural networks*. Prentice Hall, ISBN 0, 13(273350), 1.
- Heincke, B., Günther, T., Dalsegg, E., Rønning, J. S., Ganerød, G. V., & Elvebakk, H. (2010). Combined three-dimensional electric and seismic tomography study on the Åknes rockslide in western Norway. *Journal of applied geophysics*, 70(4), 292-306.
- Hem, J. D. (1985). *Study and interpretation of the chemical characteristics of natural water* (Vol. 2254): Department of the Interior, US Geological Survey.
- Hiscock, K. (2009). *Hydrogeology: principles and practice*: John Wiley & Sons.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359-366.
- Hutchison, C. S., Tan, D. N. K., & Malaya, U. (2009). *Geology of Peninsular Malaysia*: University of Malaya.
- Islami, N. (2011). Integrated geoelectrical resistivity, hydrogeochemical, and soil property analysis for groundwater investigation in north Kelantan, Malaysia, PhD. Thesis, University of Malaya.
- Jang, J.S. (1992). Self-learning fuzzy controllers based on temporal backpropagation. *Neural Networks, IEEE Transactions on*, 3(5), 714-723.
- Jang, J.S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, FMan and Cybernetics, IEEE Transactions on*, 23(3), 665-685.
- Jang, J.S., & Sun, C.T. (1995). Neuro-fuzzy modeling and control. *Proceedings of the IEEE*, 83(3), 378-406.
- JICA & MDGM. (2002). The study on the sustainable groundwater resources and environmental management for the Langat Basin in Malaysia. Japan International Cooperation Agency (JICA) and Mineral and Geoscience Department Malaysia (MDGM) Report, Vol 3

- Karim, M. H. (2006). Groundwater resources in Malaysia: Issues and challenges. Technical papers, volume 3. Minerals and Geoscience Department Malaysia
- Karimi, S., Kisi, O., Shiri, J., & Makarynskyy, O. (2013). Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. *Computers & Geosciences*, 52, 50-59.
- Karmokar, B. C., Mahmud, M. P., Siddiquee, M. K., Nafi, K. W., & Kar, T. S. (2012). Touchless written English characters recognition using neural network. *Int J Comput Organ Trends*, 2(3), 80-84.
- Kearey, P., Brooks, M., & Hill, I. (2009). *An introduction to geophysical exploration*: John Wiley & Sons.
- Keller, G. V., & Frischknecht, F. C. (1966). Electrical methods in geophysical prospecting.
- Khalil, B., Ouarda, T., & St-Hilaire, A. (2011). Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis. *Journal of Hydrology*, 405(3), 277-287.
- Kirsch, R. (2007). Groundwater geophysics: Springer.
- Koefoed, O., & Patra, H. (1979). *Resistivity sounding measurements*: Elsevier Scientific Publ.
- Kolmogorov, A. N. (1957). The representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Doklady Akademii Nauk SSSR*, 114(5), 953-956.
- Kuras, O. (2002). *The capacitive resistivity technique for electrical imaging of the shallow subsurface*. University of Nottingham.
- Lallahem, S., Mania, J., Hani, A., & Najjar, Y. (2005). On the use of neural networks to evaluate groundwater levels in fractured media. *Journal of Hydrology*, 307(1), 92-111.
- Lashkarbolooki, M., Shafipour, Z. S., & Hezave, A. Z. (2013). Trainable cascade-forward back-propagation network modeling of spearmint oil extraction in a packed bed using SC-CO₂. *The Journal of Supercritical Fluids*, 73, 108-115.
- Li, Y., Migliaccio, K., & Li, Y. (2010). *Water quality concepts, sampling, and analyses*: CRC Press Inc.
- Lippmann, R. P. (1987). An introduction to computing with neural nets. *ASSP Magazine, IEEE*, 4(2), 4-22.
- Loke, M. (2013). Tutorial: 2D and 3D electrical imaging surveys. *Geotomo Software, Malaysia*.
- Maier, H. R., & Dandy, G. C. (2000). Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental modelling & software*, 15(1), 101-124.

- Malaysian Meteorological Department (MMD). (2013). <http://www.met.gov.my>
- Mamdani, E. H., & Assilian, S. (1975). An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, 7(1), 1-13.
- Manap, M. A., Sulaiman, W. N. A., Ramli, M. F., Pradhan, B., & Surip, N. (2013). A knowledge-driven GIS modeling technique for groundwater potential mapping at the Upper Langat Basin, Malaysia. *Arabian Journal of Geosciences*, 1-17.
- Masters, T. (1993). *Practical neural network recipes in C++*: Morgan Kaufmann.
- Mathworks. (2013). Foundations of Fuzzy Logic. <http://www.mathworks.com/help/fuzzy/foundations-of-fuzzy-logic.html>.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.
- McGillivray, P., & Oldenburg, D. (1990). Methods For Calculating Fréchet Derivatives And Sensitivities For The Non-Linear Inverse Problem: A Comparative Study. *Geophysical Prospecting*, 38(5), 499-524.
- Merrick, N. (1997). A new resolution index for resistivity electrode arrays. *Exploration Geophysics*, 28(1/2), 106-109.
- Mohanty, S., Jha, M. K., Kumar, A., & Sudheer, K. (2010). Artificial neural network modeling for groundwater level forecasting in a River Island of Eastern India. *Water Resources Management*, 24(9), 1845-1865.
- Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, 6(4), 525-533.
- DOE. (2006). *National water quality standard for Malaysia (NWQS)*. Kuala Lumpur: The Ministry of Helth.
- Nasiman, S., & Nazan, A. (1997). Alluvial aquifer in Selangor: a threatened water resource. *J ENSEARCH*, 10(3/4), 195-201
- Negnevitsky, M. (2005). *Artificial intelligence: a guide to intelligent systems*: Pearson Education.
- Niesner, E., & Weidinger, J. (2008). Investigation of a historic and recent landslide area in Ultrahelvetic sediments at the northern boundary of the Alps (Austria) by ERT measurements. *The Leading Edge*, 27(11), 1498-1509.
- Omar, S. N. S., & Mansor, S. (2004). GIS/RS for landslides zonation in Pos Slim-Cameron Highlands district, Peninsula Malaysia. *Disaster Prevention and Management*, 13(1), 24-32.
- Onkal-Engin, G., Demir, I., & Engin, S. N. (2005). Determination of the relationship between sewage odour and BOD by neural networks. *Environmental modelling & software*, 20(7), 843-850.
- Openshaw, S., & Openshaw, C. (1997). *Artificial intelligence in geography*: John Wiley & Sons, Inc.

- World Health Organization. (2004). *Guidelines for drinking water quality: recommendations* (4th ed. Vol. 1): Geneva: WHO.
- Patnode, H., & Wyllie, M. (1950). The presence of conductive solids in reservoir rocks as a factor in electric log interpretation. *Journal of petroleum technology*, 2(02), 47-52.
- Perrone, A., Iannuzzi, A., Lapenna, V., Lorenzo, P., Piscitelli, S., Rizzo, E., & Sdao, F. (2004). High-resolution electrical imaging of the Varco d'Izzo earthflow (southern Italy). *Journal of applied geophysics*, 56(1), 17-29.
- Ramli, M. F., Sulaiman, W. N. A., Yusoff, M. K., Low, Y. Y., & Manap, M. A. (2005). Open source geographical resources analysis support system (GRASS) for landslide hazard assessment. *Disaster Prevention and Management*, 14(4), 522-532.
- Reynolds, J. M. (2011). *An introduction to applied and environmental geophysics*: John Wiley & Sons.
- Rhoades, J. (1982). *Reclamation and management of salt-affected soil after drainage*. Paper presented at the Proc. First Annual Western Provincial Conf. Rationalization of Water and Soil Resources and Management, Alberta: Lethbridge.
- Riedmiller, M., & Braun, H. (1993). *A direct adaptive method for faster backpropagation learning: The RPROP algorithm*. Paper presented at the Neural Networks, 1993., IEEE International Conference on.
- Ripley, B. (1996). *Pattern recognition and neural networks*. Cambridge Uni. Press, Cambridge.
- Robins, N., Jones, H., & Ellis, J. (1999). An aquifer management case study—the chalk of the English South Downs. *Water Resources Management*, 13(3), 205-218.
- Robinson, E. S. (1988). *Basic exploration geophysics*. Wiley, Chichester.
- Roy, A., & Apparao, A. (1971). Depth of investigation in direct current methods. *Geophysics*, 36(5), 943-959.
- Rumelhart, D. E., McClelland, J. L., & Group, P. R. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, Volumes 1 and 2: Cambridge, MA: MIT Press.
- Samarasinghe, S. (2006). *Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition*: CRC Press.
- Scanlon, B. R., Paine, J. G., & Goldsmith, R. S. (1999). Evaluation of electromagnetic induction as a reconnaissance technique to characterize unsaturated flow in an arid setting. *Groundwater*, 37(2), 296-304.
- Shahbudin, S., Hussain, A., El-Shafie, A., Tahir, N. M., & Samad, S. (2009). Adaptive-neuro fuzzy inference system for human posture classification using a simplified shock graph *Visual Informatics: Bridging Research and Practice* (pp. 585-595): Springer.

- Silvester, P. P., & Ferrari, R. L. (1996). *Finite elements for electrical engineers*: Cambridge university press.
- Singhal, B., & Gupta, R. P. (2010). *Applied hydrogeology of fractured rocks*: Springer.
- Sugeno, M. (1985). *Industrial applications of fuzzy control*: Elsevier Science Inc.
- Sun, T., Pan, S.b., & Li, Y.j. (2004). Application of artificial neural network model to groundwater quality assessment and classification. *Hydrogeology and Engineering Geology*, 31(3), 58-61.
- Te, C. V., Maidment, D. R., & Mays, L. W. (1988). *Applied hydrology*. McGraw Hill.
- Telford, W. W. M., & Sheriff, R. E. (1990). *Applied geophysics* (Vol. 1): Cambridge University Press.
- Todd, D. K., & Mays, L. W. (2005). *Groundwater hydrology edition*: Wiley, New Jersey.
- Verma, A., & Singh, T. (2013). Prediction of water quality from simple field parameters. *Environmental Earth Sciences*, 69(3), 821-829.
- Wasserman, P. D. (1989). *Neural computing: theory and practice*: Van Nostrand Reinhold Co.
- Waxman, M., & Smits, L. (2003). Electrical conductivities in oil-bearing shaly sands. *SPE Reprint Series*, 107-122.
- Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10), 1550-1560.
- Werbos, P. J. (1994). *The roots of backpropagation: from ordered derivatives to neural networks and political forecasting* (Vol. 1): John Wiley & Sons.
- Wetzel, R. (2001). Limnology, 3 E. Lake and River Ecosystems. *Academic Press*, 525 B Street, Ste. 1900, San Diego, CA 92101, USA. 850, 850.
- Yang, S.M., Li, F.-M., Suo, D.-R., Guo, T.-W., Wang, J.-G., Song, B.L., & Jin, S.L. (2006). Effect of long-term fertilization on soil productivity and nitrate accumulation in Gansu Oasis. *Agricultural Sciences in China*, 5(1), 57-67.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
- Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I. *Information sciences*, 8(3), 199-249.

List of Publications and Papers Presented

Journal Papers

1. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2014) Groundwater quality assessment of a fresh water wetland in the Selangor (Malaysia) using electrical resistivity and chemical analysis. *Water Science and Technology: Water Supply*, 14(2), 255-264.
2. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Application of the artificial neural network and neuro fuzzy system for assessment of groundwater quality. *CLEAN - Soil, Air, Water Journal* (In Press).
3. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Electrical resistivity imaging and hydrochemical analysis for groundwater investigation in Kuala Langat, Malaysia. *Hydrogeological Science Journal* (In Press).
4. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Integrated Geoelectrical and hydrogeochemical investigation for mapping the aquifer at Langat Basin, Malaysia. *Acta Geophysica* (Major Revision).
5. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Application of an artificial intelligence system for the simulation of groundwater level. *Environmental Earth Sciences* (Under Review).
6. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Artificial Neural Network technique for modeling of groundwater level in the Langat Basin, Malaysia. *Water Science and Technology: Water Supply* (Under Review).
7. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Imaging shallow subsurface lithology in Dengkil, Malaysia, based on electrical resistivity surveys. *Sains Malaysiana Journal* (Under Review).

8. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami Application of ANN, ANFIS and SVR in simulation of water quality in Langat Basin, Malaysia. Hydrogeological Science Journal (Under review).

National and International Conferences

1. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2014). Time lapse electrical resistivity monitoring survey for imaging the groundwater level, Langat basin, Malaysia, (ICBEEES), Malaysia.
2. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2014). Integrated resistivity survey and chemical analysis to study groundwater in the Selangor, Malaysia. Ninth Mathematics and Physical Sciences Graduate Congress 2014, Malaysia (9th MPSGC, **Malaysia**)
3. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2014). Integrated geoelectrical resistivity and geochemical survey to study groundwater: a case study from Selangor, Malaysia (76th EAGE Conference & Exhibition, **Netherlands**)
4. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2013). Electrical Resistivity Imaging to detect the bedrock profile in Langat Basin, Malaysia, National Geosciences Conference, **Malaysia**
5. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2013). Investigation of the groundwater aquifer by resistivity method at Dengkil, Malaysia, The 12th Asia Pacific Physics Conference, **Japan**.
6. **Mahmoud Khaki**, Ismail Yusoff and Nur Islami (2012). Experimental of two dimensional resistivity tomography of dyke model, 8MPSGC, **Thailand**