



**WATER QUALITY ASSESSMENT USING OPTIMAL MULTI-
OBJECTIVE WASTE-LOAD ALLOCATION APPROACH: The case of
Little Akaki River**

A Dissertation by:

MULUGETA YILMA TSEGAYE

Submitted to the School of Chemical and Bio Engineering in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY (ENVIRONMENTAL ENGINEERING STREAM)

Advisors: Dr. Ing. Zebene Kiflie (Associate Professor)
 Doz. Dr. Andreas Windsperger
 Dr. Nebiyeleul Gessese

Addis Ababa Institute of Technology, Addis Ababa university
Addis Ababa, Ethiopia

April, 2019

Addis Ababa University
School of Graduate Studies
School of Chemical and Bio Engineering

**WATER QUALITY ASSESSMENT USING OPTIMAL MULTI-
OBJECTIVE WASTE-LOAD ALLOCATION APPROACH: The case of
Little Akaki River**

A Dissertation by:

MULUGETA YILMA TSEGAYE

Approved by:

Dr. Ing. Zebene Kiflie _____ (Main Advisor)

DECLARATION

Addis Ababa University
School of Graduate Studies
School of Chemical and Bio Engineering
Environmental Engineering Post Graduate Program

As members of Examining Board of the final PhD Dissertation public defense, we certify that we have read and evaluated the dissertation prepared by Mulugeta Yilma, entitled "Water Quality Assessment Using Optimal Multi-Objective Waste-load Allocation Approach: the case of Little Akaki River" and recommended that it can be accepted as fulfilling the dissertation requirements for the Degree of Doctor of Philosophy in Chemical Engineering (Environmental Engineering Stream).

By: MULUGETA YILMA TSEGAYE

Approved by the Examining Board:

Signature

Date

School Dean:

Main Advisor:

External Examiner:

Dr. Shashidhar



27/2/2019

Internal Examiner:

S e i f u A d m a s s u



4 - 4 - 2 0 1 9

ABSTRACT

Water Quality Assessment using Optimal Multi-Objective Waste-Load Allocation Approach:
The case of Little Akaki River

Mulugeta Yilma Tsegaye

Addis Ababa University, December 2018

In Addis Ababa, indiscriminant waste disposal from domestic, industrial and commercial sources affect the water quality of the Little Akaki River. Moreover, this situation is limiting the usability of the River. Consequently, in many studies, the River is regarded as one of the most polluted Rivers in Ethiopia. On the other hand, there are inadequate comprehensive studies on the river principally due to insufficient research fund. In addition, existing studies on the River mainly focus on concentration measurement of certain constituents and their comparison against local and international standards. This approach has limited the public and policy makers from knowing the exact pollution status of the River. Furthermore, a number of studies conducted on the pollution problem and ongoing efforts to enforce existing environmental regulations have not effectively restored the river water quality. As a result, its pollution problem has increasingly worsened.

Therefore, in this study, the application of *statistical multivariate analysis* for regular and economical water quality assessment, *water quality index analysis* for summarizing the water quality situation and *optimal waste-load allocation modeling* as a tool for decision-making are sought.

For *statistical multivariate analysis*, twenty-seven locations from the River and tributaries were sampled and analyzed in October/November 2015. Afterwards, multivariate statistical tools were used to investigate data from measurements and laboratory analysis. Accordingly, the cluster analysis divided the sampled sites into three according to level of their pollution. This result indicates that water quality variation was caused because of the difference in land-use conditions. In addition, for the spatial analysis of the three pollution groups, backward stepwise approach of discriminant analysis was identified to provide data reduction (87.5%) to two parameters resulting in 85.2% correct assignment. The principal component

analysis/factor analysis identified ten parameters accounting for 81.9% of total variation. However, data reduction was not significant. The factors that were latent and identified from the principal components' varimax rotation suggest that variation in water quality was caused mainly by domestic sewage. The outcomes show that the methods can be applied to evaluate the river water quality variation using three monitoring sites and ten parameters: total nitrogen, total suspended solids, total ammonia, chemical oxygen demand, nitrite, total phosphorus, phosphate, nitrate, biological oxygen demand and electrical conductivity. This, in consequence, requires lesser cost and effort and hence paves way for more affordable, regular water quality evaluation of Little Akaki River.

For *index analysis*, twelve water quality parameters from twenty-seven sampling sites in the Dry season (January/February, 2017) and Wet season (October/November, 2015) were used for index determination. Results show that, all sampling sites except one site in the upstream were under *poor* water quality category. Afterwards, the neural network model was trained and validated, for twelve inputs and one output, using several combinations of hidden layers (2-20), number of neurons in the hidden layers (5, 10, 15, 20, 25), transfer, training and learning functions. The most optimal model architecture was obtained with eight hidden layers, fifteen hidden neurons that resulted in R^2 value of 0.93. This shows a good agreement between the calculated and predicted index values suggesting that artificial neural network can be successfully applied for modeling Little Akaki River's water quality index.

One of the ongoing efforts by the Addis Ababa Water and Sewerage Authority to control and minimize impact from pollution source is channeling a portion of domestic discharges to central treatment plants. However, the nature of these discharges is variable in time and space. Accordingly, impact on the river varies. Under limited treatment plant capacity, the river's health can be maximized either by putting in place strict environmental control or by preferentially channeling the streams having significant impact on the river. Identification of these streams can be done through routine field sampling and laboratory analysis and decision making afterwards. However, this requires high financial, time and human resource. In this regard, *water quality simulation* can help to understand the interaction between pollution sources and the river. For this, monitoring data can be used to predict pollution contribution of the various sources on the river and agencies can apply this approach for environmental decision-making. In this study,

QUAL2Kw was used to predict the river water quality. The model was calibrated and validated using data collected during dry (January/February, 2017) and wet seasons (October/November, 2015), respectively. The results from the calibrated model indicate that the model was able to reasonably predict the pollution of the river with R^2 values of 0.91, 0.90, 0.81 and 0.89, respectively for dissolved oxygen, biological oxygen demand, total nitrogen and total phosphorus. Moreover, sensitivity analysis showed that dissolved oxygen, biological oxygen demand, total nitrogen and total nitrogen predictions are highly sensitive to point source flow and Manning's n . Therefore, this model may be applied as an option for water quality management of the Little Akaki River.

For *optimal waste-load allocation*, a simulation-optimization model was developed through integration of a water quality model - QUAL2Kw and genetic algorithm - PIKAIA. Afterwards, cost-performance and cost-performance-equity models were applied on water quality data set. The model resulted in pareto-optimal curves for conflicting objectives such as treatment cost and equity versus water quality performance. These curves offer convenient means for informed decision-making during environmental planning and implementation. Especially, the strategy is helpful in finding compromised solutions for pollution problems with conflicting objectives. In general, the study results indicate that significant waste load reduction is required for an improved water quality condition of the Little Akaki River.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the almighty God for all the blessings during this thesis work. I would also like to express my deep and sincere gratitude to my main advisor Dr. Ing. Zebene Kiflie for his supervision, insightful guidance, valuable comments, follow-up and encouragement throughout this Ph.D work. I would like to extend my heartfelt gratitude to my co-advisors Doz. Dr. Andreas Windsperger and Dr. Nebiyeleul Gessese for the valuable advice and encouragement.

I am very grateful to Greg Pelletier for his brief but important email response about PIKAIA- genetic algorithm code integration with QUAL2Kw-VBA code for the purpose of optimal multi-objective waste-load allocation modeling.

I am extremely grateful to my parents for their love and prayers since day one. I am very much thankful to my dear wife Tenaye Abera and my son Kaleb Mulugeta for their love, patience and understanding during this research work. Also I am thankful to my sisters, brothers, sister in law and brothers in law for their encouragement. My special thanks go to my longtime friends Andinet Abebe and Endashaw Kinfu and their beloved wives for our invaluable friendship and encouragement during this thesis work.

I would also like to thank Ato Zenebe Bitere and Dr. Tatek Temesgen for our friendship and encouragement. It is worth appreciating my friends and Ph.D colleagues for encouragements made through sharing the progresses each of us made in our respective researches and the concern we showed to each other.

Last but not least, I would like to express my gratitude to the staffs at the School of Chemical and Bio Engineering of Addis Ababa Institute of Technology, all the people who have supported me to complete this research work directly or indirectly.

LIST OF PUBLICATIONS

1. Mulugeta Yilma, Zebene Kiflie, Andreas Windsperger, Nebiyeleul Gessese. *Application of artificial neural network in water quality index prediction: a case study in Little Akaki River, Addis Ababa, Ethiopia. Journal of Modeling Earth Systems and Environment.* 2018; 4(1): 175-187. <https://doi.org/10.1007/s40808-018-0437-x>
2. M. Yilma, Z. Kiflie, A. Windsperger, N. Gessese. *Assessment and interpretation of river water quality in Little Akaki River using multivariate statistical techniques. International Journal of Environmental Science and Technology.* 2018; 15(.). <https://doi.org/10.1007/s13762-018-2000-8>
3. Mulugeta Yilma, Zebene Kiflie, Andreas Windsperger, and Nebiyeleul Gessese. *Optimal Multi-Objective Waste-Load Allocation Modeling using QUAL2Kw and Genetic Algorithm: The Case of Little Akaki River: Under review (Journal of Environmental Modeling & Assessment)*
4. Mulugeta Yilma, Zebene Kiflie, Andreas Windsperger, and Nebiyeleul Gessese. *Water quality simulation of the Little Akakai River using QUAL2Kw: Ready for submission.*

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iv
LIST OF PUBLICATIONS	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	x
LIST OF TABLES	xi
LIST OF ACRONYMS	xii
1. INTRODUCTION	1
1.1. Background	1
1.2. Problem Statement	2
1.3. Study Objectives	5
1.3.1. General objective	5
1.3.2. Specific objectives	5
1.4. Research Application	5
1.5. Outline of the Research Approach	6
1.6. Assumptions, Limitations and Scope of the Study	6
1.6.1. Research Assumptions	6
1.6.2. Research Hypothesis	7
1.6.3. Research Limitations	7
1.6.4. Research Scope	7
2. LITERATURE REVIEW	8
2.1. General	8
2.2. Urban Rivers	8
2.3. Water Quality Assessment: Statistical Approach.....	9

2.4.	Water Quality Assessment: Water Quality Index Approach	10
2.4.1.	Water Quality Index.....	10
2.4.2.	Artificial Neural Network Application in Water Quality Index Prediction.....	11
2.5.	Water Quality Modeling and Simulation	12
2.5.1.	General.....	12
2.5.2.	Water Quality Model Equations	14
2.6.	Water Quality Modeling and Simulation Software.....	17
2.6.1.	Historical Development of Water Quality Models	17
2.6.2.	Water Quality Softwares Description	18
2.6.3.	Water Quality Models Application.....	21
2.6.1.	Evaluation and Selection of Water Quality Model.....	24
2.7.	Q2Kw Description.....	26
2.7.1.	General.....	26
2.7.2.	Simulation Equations	28
2.8.	Multi-objective Optimization.....	34
2.8.1.	General.....	34
2.8.2.	Genetic Algorithm (GA).....	35
2.9.	Simulation-Optimization Models.....	37
2.10.	Waste Load Allocation	37
2.10.1.	General Description	37
2.10.2.	WLA Problem Formulations.....	38
2.10.3.	WLA Models Applicability	40
3.	MATERIALS AND METHODS.....	41
3.1.	Study Area.....	41
3.2.	Little Akaki River.....	42

3.2.1.	Land Use	44
3.2.2.	Population	45
3.2.3.	Climate	47
3.2.4.	Geology.....	47
3.3.	Environmental Policies and Regulations.....	48
3.4.	Sampling and Analysis.....	48
3.5.	Multivariate Statistical Analysis	51
3.5.1.	Cluster Analysis	52
3.5.2.	Discriminant Analysis.....	53
3.5.3.	Principal Component Analysis	54
3.5.4.	Factor Analysis	57
3.6.	Water Quality Index Analysis.....	59
3.6.1.	CCME-Water Quality Index.....	60
3.6.2.	Artificial Neural Network.....	62
3.7.	Applications of Q2Kw	68
3.7.1.	River Discretization	68
3.7.2.	Data Entry	69
3.7.3.	Calibration and Validation.....	75
3.7.4.	Uncertainty and Sensitivity Analysis.....	77
3.8.	Simulation-Optimization based WLA Modeling	78
3.8.1.	Q2Kw Codes.....	78
3.8.2.	Genetic Algorithm Codes	79
3.8.3.	Q2Kw-PIKAIA Integration for Optimization.....	81
3.8.4.	Proposed Model	82
4.	RESULTS AND DISCUSSION.....	86

4.1.	Descriptive and Multivariate Analysis.....	86
4.1.1.	Descriptive statistics	86
4.1.2.	Multivariate Analysis.....	90
4.2.	Water Quality Index.....	106
4.2.1.	CCME-WQI.....	106
4.2.2.	Water Quality Index based on Artificial Neural Network.....	110
4.3.	Q2Kw Simulation Model.....	112
4.3.1.	Calibration and Verification	112
4.3.2.	Sensitivity analysis.....	118
4.4.	Multi-objective Waste-load Allocation Model	120
4.5.	Cost-performance model.....	120
4.5.1.	Model’s Applicability to the Local Condition.....	124
5.	CONCLUSIONS AND RECOMMENDATIONS	125
5.1.	Conclusions.....	125
5.2.	Significance of the Work for Rehabilitation of the LAR.....	127
5.3.	Recommendations for Future Work.....	128
6.	REFERENCES	129

APPENDICES

APPENDIX I: METEOROLOGICAL DATA.....	145
APPENDIX II: MAXIMUM and MINIMUM VALUES of RAW WQ DATA	149
APPENDIX III: PSEUDO-CODE FOR Q2Kw-PIKAIA-GA based WLA MODELING. .	150
APPENDIX IV: DATA-TABLES FOR WLA.....	155

LIST OF FIGURES

Figure 2-1 Mass balance for constituents in reach segment “i”	28
Figure 2-2: Classification of search and optimization methods	34
Figure 3-1 Map of Addis Ababa	42
Figure 3-2 Awash river basin.....	43
Figure 3-3 Partial view of irrigated land along LAR near Mekanisa.....	44
Figure 3-4 Sampling sites from Little Akaki River (LAR).....	49
Figure 3-5 Typical BP Network.....	63
Figure 3-6 ANN Model Framework for CCME-WQI.....	67
Figure 3-7 Little Akaki River Segmentation and Pollution Sources	69
Figure 3-8 Partial list of Q2Kw data entry sheets	70
Figure 3-9 Trapezoidal channel	71
Figure 3-10 Reach flow balance	73
Figure 3-11 A partial view of modular VBA code organization of Q2Kw model	79
Figure 3-12 Optimal MO-WLA modeling framework	82
Figure 4-1 Map of AA city and LAR.....	89
Figure 4-2 Dendrogram showing clustering of sampling sites of LAR.....	94
Figure 4-3 Box plot of the most discriminating parameters	101
Figure 4-4 Water Quality Index based on CCME for LAR.....	107
Figure 4-5 LAR BOD, CCME-WQI and Population Density relationship	109
Figure 4-6 Performance of CCME-WQI prediction using ANN model for LAR	111
Figure 4-7 Model calibration	116
Figure 4-8 Model verification	117
Figure 4-9 Cost-performance trade-off graphs	121
Figure 4-10 Cost-equity-performance trade-off.....	122
Figure 4-11 Equity and performance versus WL reduction.....	124

LIST OF TABLES

Table 2-1 Concentration of pollutants in urban runoff	9
Table 2-2 Countries of application of some WQI approaches	11
Table 2-3 Historical developmental stages of watershed, river and stream models	18
Table 2-4 Characteristics of common watershed and WQ models	20
Table 2-5 A survey of model based WQ and hydrology studies in Ethiopia.....	23
Table 2-6: Model state variables	27
Table 3-1 Addis Ababa City Land Use Condition, 2011	44
Table 3-2 Partial list of industries in the sub-cities of LAR catchment	45
Table 3-3 Population and settlement density in the AA Sub-Cities.....	46
Table 3-4 Sample collection sites from LAR.....	50
Table 3-5 Parameters and analytical methods.....	51
Table 3-6 CCME-WQI categories (CCME 2001).....	62
Table 3-7 River Reach based Manning’s roughness, n values of the LAR	75
Table 4-1: Basic Statistical analysis of WQ data of LAR.....	88
Table 4-2 Basic statistical analysis of water quality data of LAR (Wet Season).....	91
Table 4-3 Cluster pollution rank	97
Table 4-4 Classification matrix for DA of spatial variation in LAR	99
Table 4-5 PCs loading for parameters.....	103
Table 4-6 Rotated component matrix.....	105
Table 4-7: ANN model prediction performance	112
Table 4-8 List of Rate Parameters and respective Calibration Values.....	112
Table 4-9 Sensitivity analysis results	119

LIST OF ACRONYMS

A	Arada
AA	Addis Ababa
AAWSA	Addis Ababa Water and Sewerage Authority
ADE	Advection-Diffusion Equation
AK	Addis-Ketema
AKK	Akaki-Kality
AMSL	Above Mean Sea Level
ANN	Artificial Neural Network
B	Bole
BMSL	Below Mean Sea Level
BAR	Big Akaki River
BOD	Biological Oxygen demand
CA	Cluster Analysis
CBOD	Carbonaceous Biological Oxygen Demand
CCME	Canadian Council of Ministers of the Environment
CSA	Central Statistics Agency
CSO	Combined Sewer Overflows
DA	Discriminant Analysis
Dlink	Distance of Linkage
DO	Dissolved Oxygen
FA	Factor Analysis
FDRE	Federal Democratic Republic of Ethiopia
FDREMT	Federal Democratic Republic of Ethiopia, Ministry of Transport
Fig.	Figure
G	Gulele
GA	Genetic Algorithm
ISS	Inorganic suspended solids
K	Kirkos
KK	Kolfe-Keranio
km ²	Kilometer Square

KMO	Kaisere Meyere Olkin
K-S	Kolmogorove-Smirnov
L	Lideta
LAR	Little Akaki River
L/d	Liters/Day
M	Meter
M ³	Meter Cube
Mg/L	Milligram/Liter
MO	Multi-Objective
MoWIE	Ministry of Water, Irrigation and Electricity
MSM	Multivariate Statistical Methods
n-D	n-dimensional:
NL	Nefas silk-Lafto
PCA	Principal Component Analysis
Q2E	QUAL2E
Q2K	QUAL2
Q2Kw	QUAL2Kw
SO	Simulation-Optimization
T ⁰	Temperature
TMDL	Total Maximum Daily Load
VBA	Visual Basic for Applications
VFs	Varifactors
WACC	Waste Assimilative Capacity Concept
WBG	World Bank Group
WL	Waste-Load
WLA	Waste-Load Allocation
WQ	Water Quality
WQI	Water Quality Index
Y	Yeka

1. INTRODUCTION

1.1. Background

Water is essential to the existence of life on earth. Moreover, water contributes enormously to social and economic aspects of a nation (Gerbens-Leenes *et al.* 2009; Pereira *et al.* 2009). Shortage of freshwater due to manmade problems such as pollution has posed big challenge around the world, particularly in developing countries (Kivaisi 2001). Generally, pollution of rivers and streams comes from industrial waste, agricultural runoff and untreated municipal sewage (Pereira *et al.* 2009). This situation worsens through rapid urbanization, lack of sufficient wastewater treatment facilities and capacity shortfall in enforcing pollution control laws (Kivaisi 2001).

The daily wastewater generation in AA has been estimated at 398,985m³ (MoWIE 2017). In addition, the city struggles with critical wastewater handling and management problems. From the total wastewater produced only 7.5% is entertained and treated at the city's two treatment plants: Kalitiy and Kotebe wastewater treatment plants (MoWIE 2015). The remaining commercial and residential liquid waste is largely drained to sewer lines, which ultimately reaches BAR and LAR (UN-Habitat 2007; Van Rooijen & Taddesse 2009).

The catchment area of the Awash basin that includes the city of Addis Ababa (AA) is divided into two sub-catchments by approximately a north-south running surface water divide. These are the Big Akaki River (BAR) sub-catchment and the Little Akaki River (LAR) sub-catchment (Bekele, 1999). These Rivers are used for drinking, industrial, irrigation, and other purposes. Unfortunately, they are also used as an open waste recipient of the waste generated by

industries and other sources from AA and the surrounding areas (Bekele, 1999). Studies indicate that LAR is more polluted than BAR (Mohamed, 2002).

LAR is located in the Western part of AA and it is a combination of two main streams, namely Gefersa and Gulele. The former flows down from the Northwestern outskirts of the city while the latter begins as streams from Intoto Mountain in Gulele Sub-city. As the river advances to the central and Southern part of the city, it is exposed to wastewater, to a varying degree, from domestic, commercial, agricultural and industrial sectors (Mengistu & Alemayehu 2013).

In this context, water quality assessment of LAR is an important part in the effort to minimize its pollution level and negative impacts.

1.2. Problem Statement

Water is essential to life on earth. It has a very high social and economic value and its availability and use play an important role in the development of our societies (Hinkelmann 2006). Developing countries face problems associated with degraded water quality (WQ) (Saremi *et al.* 2010). The solutions to such problems require the application of scientific principles (Simonovic 2009b).

Annually AA generates about 49 million m³ of grey water, of which 87% is domestic, while the remaining 13% is of industrial origin (Beles 2014). The drainage and sewer system infrastructure was dysfunctional, although it had a coverage capacity of only 3% of the city in 2010 (UN-Habitat 2010). This is partly due to illegal encroachment of sewer trunks by squatter settlements (Beles 2014). This absence of a functional sewer system was described as the cause for record level pollution due to wastewater ending up in the Akaki River almost entirely (UN-Habitat 2007; Van Rooijen *et al.* 2010).

LAR is used for irrigation and domestic purposes, despite its bad odour, black colour and toxicity. It serves as an open waste recipient of industrial and domestic wastes. In this regard, the major sources of industrial pollution include tanneries, breweries, wineries, distilleries and a slaughter facility (Mebratu and Redda 2005; Bekele 1999). As a result, LAR faces WQ deterioration and management challenges (Melaku *et al.* 2007; Van Rooijen *et al.* 2010). However, there are inadequate comprehensive studies that are based on regular and continuous data collection approach particularly due to budgetary constraints. Because of this, regulatory and technical solutions that are often taken to address the pollution problem face limitations. Multivariate statistical methods are useful methods for WQ data reduction and can be applied for continuous assessments.

Some of the reports on LAR studied predominantly heavy metals and other parameters concentration impact on health and the River's usability for irrigation purpose (Van Rooijen *et al.* 2010; Mekonnen *et al.* 2012; Weldegebriel *et al.* 2012). The likes of: Melaku *et al.* (2007) and Worku *et al.* (2017) investigated the general health of the River. These and other similar studies mainly used concentration measurements of certain constituents and compared them against local and international standards. The multitude of objectives and differences in selected constituents in the various reports require scientific knowledge to understand the River water quality status. This may hinder policy makers and the public from knowing the extent of the River's pollution.

To control and minimize impact of pollution from domestic, industrial and commercial sources, one of the ongoing efforts by the Addis Ababa Water and Sewerage Authority (AAWSA) is to channel domestic discharges to central treatment plants. The nature of these discharges is variable in time and space. Accordingly, impact on the river varies. Under limited

treatment plant capacity, the river's health can be maximized either by putting in place strict environmental control or by preferentially channeling the streams having significant impact on the river. Identification of these streams can be done through routine field sampling and laboratory analysis and decision making afterwards. However, this requires high financial, time and human resource. In this regard, water quality simulation can help to understand the interaction between pollution sources and the river. For this, it can use monitoring data to predict pollution contribution of the various sources on the river. Thus, agencies can apply this approach for environmental decision-making.

A number of studies conducted on the pollution problem and ongoing efforts to enforce existing environmental regulations have not effectively restored LAR WQ condition to acceptable levels. Consequently, its pollution problem has increasingly worsened. Impaired rivers like LAR require alternative environmental decision-making mechanisms that can truly work under their circumstance. In this context, assimilative capacity estimation of a river or its reaches is a very important step towards restoring WQ of a waterbody. It reveals the most useful information on loading capacity of the river and pollution sources level of impact. With this consideration, to restore the impaired portion of the LAR, assimilative capacity of different reaches of the river needs to be identified. And these sections of the river should not be loaded beyond their assimilative capacity. This can be approached by using multi-objective (MO) simulation-optimization (SO) waste-load allocation (WLA) modeling. In this regard, optimal WLA based on a river's assimilative capacity enables estimation of discharge loads for an improved river water quality management.

1.3. Study Objectives

1.3.1. General objective

The main aim of this research is to undertake *water quality assessment using optimal multi-objective waste-load allocation modeling in order* to come-up with an environmental decision-making tool for addressing pollution problem of the Little Akaki River.

1.3.2. Specific objectives

The specific objectives of the research shall be to:

- Identify parameters of significance in spatial WQ variation of LAR, latent factors of pollution, spatial pollution condition using multivariate statistical approaches.
- Summarize the spatial pollution condition of LAR using water quality index approach.
- Do WQ simulation using QUAL2Kw (Q2Kw)-WQ model using calibration and validation data.
- Do a MO optimization of equity, waste load reduction and treatment cost using coupled genetic algorithm (GA) and Q2Kw-WQ model for better River WQ condition..

1.4. Research Application

The research can find application in the following areas:

- For economical and regular WQ assessment
- For devising responsive strategies by policy-makers for pollution problems.
- For environmental-decision making by environmental agencies and industries.
- For research purposes in environmental studies.

1.5. Outline of the Research Approach

The research follows classical approach and is outlined as follows:

Chapter 1 provides general information on the study area, the research need, research objective, research application, assumptions, limitations and scope.

Chapter 2 details literature review on the following topics: urban river characteristics, water pollution classification, pollutant constituents, water quality index (WQI), WQ regulatory tools, WQ modeling equations and software primarily Q2Kw, optimal MO *WLA* and its application.

Chapter 3 deals with research methodology and model formulation. The chapter provides approaches used in this study including data sources and collection, multivariate statistical analysis, water quality index analysis, applications of Q2Kw, and *SO* based *WLA* modeling.

Chapter 4 presents study results and discussion. The study outputs described in this section include: statistical analysis; water index of the river; calibration, validation and sensitivity analysis using Q2Kw; optimal *WLA* based on cost-equity and cost-performance-equity model formulations and its importance.

Chapter 5 summarizes conclusions drawn from this study and gives recommendations for future works.

1.6. Assumptions, Limitations and Scope of the Study

1.6.1. Research Assumptions

This research proposal is based on the following hypotheses:

- During January-February the river was under Low-flow conditions.
- During the low-flow condition, the pollution of LAR was at its highest level.

- During the data collection period, the conditions were under a fairly steady-state condition.

1.6.2. Research Hypothesis

The results of the model will provide insight into the pollution problem and provide understanding on how to restore and maintain the health of LAR, help develop additional management tools to restore and maintain impaired local rivers using statistical tools, water quality indexing, WQ modeling and waste load allocation modeling approaches.

1.6.3. Research Limitations

The research limitations are:

- There was no organized historical water and hydrological data on the River to support the modeling activity.
- Access to the River was mostly limited by settlements especially in Sub-Cities with high population density.
- The data collected was fairly enough to meet the research objectives. However, additional data could have improved the modeling results.

1.6.4. Research Scope

The scope of this study is limited to LAR. The study covers the pollution condition of the River and impact of pollution sources on the River. The study aims to find approaches which can be used during WQ assessment and environmental decision-making. In this regard, the study uses statistical multivariate analysis, water quality index analysis, Q2Kw based WQ simulation and optimal WLA modeling using coupled MO Q2Kw-GA.

2. LITERATURE REVIEW

2.1. General

Fresh water is a fundamental resource and it sustains life but it is scarce. That is only 2.53 per cent of the earth's water potential is freshwater, while the remainder is saltwater. Some two-thirds of this freshwater is locked up in glaciers and permanent snow cover. In addition, its systems are imperiled, and this threatens both human well-being and the health of ecological systems (Simonovic 2009a). The main causes are population growth, agricultural activities, industrialization, urbanization and climatic factors (USEPA 1996; Benedini & George 2013). In this context, WQ contamination results in water shortage and increases the costs of treating water for use (Carpenter *et al.* 1998).

Rivers are the primary sources supplying water for municipal, agricultural and industrial uses. Consequently, there is public concern on WQ pollution of Rivers. Pollutant release is the primary reason for pollution of Rivers. Moreover, the management of pollution sources is often challenging and requires techniques that are useful for addressing the problems. Improved WQ prediction and protection techniques are important for successful management of pollution problems in these waters (Tung & Hathhorn 1989; Mannina 2010; Mostafavi & Afshar 2011).

2.2. Urban Rivers

The water quality situation in urban areas is represented with parameters such as biological oxygen demand (BOD), ammonia, suspended solids (SS) and chemical oxygen demand (COD) (von Sperling & de Lemos Chernicharo 2002). **Table 2-1** below shows typical concentration of parameters in urban areas (Price & Vojinovi 2011). These values are usually violated particularly in developing countries (Asit *et al.* 2006). Some examples of urban rivers around the world

include: Bagmati River in Nepal with values of BOD (185mg/L) (Kannel *et al.* 2007a) and Bangladesh with values of BOD (82mg/L) (Ahammed *et al.* 2016).

Table 2-1 Concentration of pollutants in urban runoff

Parameters (mg/L)	Highway runoff	Residential areas	Commercial areas	Industrial areas
BOD5	12-32	7-56	5-17	8-12
Ammonia	0.02-2.1	0.3-3.3	0.03-5.1	0.2-1.2
SS	28-1178	112-1104	230-1894	34-374
COD	128-171	37-120	74-160	40-70

The major challenges of waste generation in urban areas are direct discharge and/or inadequate treatment of domestic and industrial wastewater. Particularly, management of streams and rivers becomes more complex under amplified urbanization and industrialization conditions. This problem is usually greater in developing countries than in developed countries. Some of the issues in developing countries include: (1) Lack of sufficient resource and inadequate public awareness. (2) urban drainage still dominated by 19th-century sanitary philosophy that disregards impact on the downstream. (3) contamination of streams by sediment, foul sewage and garbage. (4) lack of adequate experience in drainage management and (5) problems with enforcing standards(Silveira 2002, von Sperling, 2002).

2.3. Water Quality Assessment: Statistical Approach

Data collection and analysis through development of monitoring network that effectively accounts for spatial WQ differences in Rivers is essential for effective and lasting management (Ouyang *et al.* 2006; Qadir *et al.* 2008). In this regard, different multivariate statistical methods (MSM) such as cluster analysis (CA), discriminant analysis (DA), principal component analysis

(PCA) and factor analysis (FA) can be employed to improve knowledge on WQ status, reduce WQ data, identify the highly significant factors and interpret WQ data easily (Chapman 1992b; Vega *et al.* 1998; Singh *et al.* 2005; Bhuiyan *et al.* 2011; Bouguerne *et al.* 2017).

In this context, descriptive statistics and MSM application was observed in different reports. Exploratory statistics were applied for weighing river WQ condition (Chapman 1992b; Vega *et al.* 1998). CA was applied to sort sampling sites spatially based on WQ similarity (Vega *et al.* 1998; Bouguerne *et al.* 2017). DA was employed to identify variables that are significant enough for discriminating the different pollution groups (Gholizadeh *et al.* 2016; Sabri *et al.* 2016). PCA was applied to downsize WQ parameters dimensionality (Ayeni & Soneye 2013; Khan *et al.* 2017); while FA was employed to understand unobservable factors that cause WQ variation (Voutsas *et al.* 2001; Singh *et al.* 2005; Juahir *et al.* 2011).

In this study, the application of exploratory statistics and multivariate statistical approaches are considered for WQ data dimensional reduction and WQ assessment purposes.

2.4. Water Quality Assessment: Water Quality Index Approach

2.4.1. Water Quality Index

The public concern on the pollution of rivers and streams (Tunstall *et al.* 2000; Said *et al.* 2003) necessitates statements with minimum technical knowledge requirement (CCME 2001). One such method that gives more understandable summary of the WQ data is WQI (Brown *et al.* 1972).

WQI was introduced in 1965 in the United States (Horton 1965). It uses complex data and yields a single number that describes WQ status of a water body. It also evaluates WQ trends and allows preliminary classification of River water for the purpose of various designated uses

such as recreation, agriculture and others (Cude 2001; Kannel *et al.* 2008). However, there are various approaches of WQI. Examples are Canadian Council of Ministers of the Environment (CCME) WQI, National Sanitation Foundation *WQI* (NSF), Oregon WQI etc (Abbasi & Abbasi 2012). So far, there is no worldwide-accepted method among the many approaches. Consequently, there is still difference in global application of the WQIs (**Table 2-2**) (Sutadian *et al.* 2016).

Table 2-2 Countries of application of some WQI approaches

WQI Approach	Countries of Application	Reference
CCME	Canada, India, Albania, Chile, Egypt, Iran, Spain, Turkey and Poland	(Sutadian <i>et al.</i> 2016)
NSF	USA, Brazil, India and Iran	
Oregon WQI	USA	
Bascarón index	Spain, Argentina, Brazil, Korea and India	
Malaysian index	Malaysia	
Status and Sustainability index	France	

In this study, CCME is used for WQI determination of LAR due to its capacity to simplify the reporting of complex and technical WQ data (CCME 2001)

2.4.2. Artificial Neural Network Application in Water Quality Index Prediction

In many cases, WQI requires lengthy calculations and transformations. In addition, some of the transformations employ different formulae for the different values, or ranges of values, almost for every WQ input parameter. Alternatively, artificial neural network (ANN) can be used for modeling WQI of rivers (Gazzaz *et al.* 2012). The key characteristic of a neural network being its ability to learn (Sharma *et al.* 2015b), ANN can simplify and speed up WQI

determination. In fact, successful applications of ANN for WQI determination are reported in different papers (Khuan *et al.* 2002; Juahir *et al.* 2004; Mohammadpour *et al.* 2016; Sakizadeh 2016).

ANN is a computational process which attempts to represent and compute a mapping from multivariate data set as inputs neuron layers to one or several hidden layers and an output layer comprising one or several neurons (Lek *et al.* 1996; Mohammadpour *et al.* 2016). Hidden nodes in the hidden layer allow neural networks to detect and capture the pattern in the data. Logistic transfer functions are majorly used for hidden nodes, however for output nodes linear activation function are used in forecasting problems which involves continuous target values (Zhang *et al.* 1998). The method can handle non-linearity in WQ data without requiring assumptions about the underlying statistical properties of a data set (Lek *et al.* 1996). This study tries to develop ANN model for LAR and demonstrate its appropriateness for the prediction of CCME-WQI values.

2.5. Water Quality Modeling and Simulation

2.5.1. General

Modeling is any mathematical formulation describing a system and that can be used for prediction and management purposes (Aral 2010). Computational WQ model can also be defined as a replica of how we perceive the problem in view of knowledge, data, and scopes for the activity (Somlyódy & Varis 1992).

For water resource management purpose, monitoring and controlling are useful in pollution identification of rivers (Benedini & George 2013). The practice of modeling of water quality as a means of River water quality more common in developed countries than developing countries. Some examples are the application of Total Maximum Daily Load approaches

practices in Korea (OECD 2018), U.S.A after the Clean Water Act of 1972 (Lung 1993) and Europe after the Water Framework Directive (Lawson 2005). Here South Africa is taken as an example for the developing countries. South Africa, based on National Environmental Management Act (NEMA), recognizes beneficial use and ecological function of water bodies. However, practical application of the regulation is not yet fully realized (Hranova 2006b).

In this regard, WQ decision-making based on monitoring approach alone has its limitations, though it is essential part of WQ management. The approach is not cost-effective and falls short of showing pollution scenarios on a basin scale for planning and decision-making purposes. In addition, damage will be done at the time of taking corrective action for a pollution problem identified during monitoring. In this case, investigators will know the standard violations in the receiving water body during monitoring; however, the approach is not good enough to prevent this from happening beforehand (Gu & Dong 1998; Deksissa *et al.* 2004; Camargo *et al.* 2010; Gikas 2014).

On the other hand, WQ modeling is cost-effective and has the ability to establish cause–effect relations between pollution sources and receiving aquatic environments based on scientific knowledge. Moreover, it can enhance policy makers’ understanding of the assimilative capacity of water bodies. Thus, the approach can be used during the development stage of control measures and during assessment of their effectiveness (Gu & Dong 1998; Reichert *et al.* 2001; Bottino *et al.* 2010; Oliveira *et al.* 2012; Holguin-Gonzalez 2013). In this regard, impact pattern due to various pollution scenarios can be established for developing and implementing appropriate pollution control measures. Monitoring and modeling approaches, when used in combination, give useful information to the regulator for an effective environmental decision-making during planning and implementation (Reichert *et al.* 2001; Susilowati *et al.* 2004; Kannel

et al. 2011). Therefore, mathematical models have enormous potential to address WQ problems and give rational solutions (Benedini & George 2013).

2.5.2. Water Quality Model Equations

Mathematical models apply balance equations for momentum, mass and energy to describe chemical, physical and biological processes (Hinkelmann 2006). These processes are governed by basic transport equation that is applicable for all mechanistic WQ models (Reichert *et al.* 2001; Cox 2003).

In the mid-1800s Navier and Stokes developed the application of Newton's laws to hydrodynamics (Schütze *et al.* 2011). Flow of water in a river is a turbulent flow which can be described by the continuity (mass balance) and Navier-Stokes (momentum) equations given in **Equation 2-1**. The right-hand side (RHS) of the momentum equation is the forces that act on it: internal body forces, pressure, and stress (Foias *et al.* 2001):

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} &= \mathbf{f} - \frac{1}{\rho} \nabla p + \nu \Delta \mathbf{u} && \text{(momentum)} \\ \frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{u}) &= 0 && \text{(Continuity)} \end{aligned} \tag{2-1}$$

where p is pressure, \mathbf{u} is velocity, t is time, ∇ is Hamilton operator, \mathbf{f} is the field mass forces, ρ is density, ν is coefficient of viscosity and Δ is Laplace operator.

It is St. Venant who came up with the first basic physics explanation of water motion in open channels (Schütze *et al.* 2011). In this regard, assumptions made on description of turbulence determine the form of a hydrodynamic model. St. Venant equations are composed of the momentum and continuity (mass balance) equations. These equations for a 1-D channel with omitted eddy losses and wind shear become (Cox 2003; Holguin-Gonzalez 2013):

$$\frac{\partial A_x}{\partial t} + \frac{\partial Q}{\partial x} = q \quad (\text{conservation of mass}) \quad (2-2)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial\left(\frac{\alpha Q^2}{A}\right)}{\partial x} + gA \frac{\partial y}{\partial x} - gA(S_o - S_f) = 0$$

Local acceleration
Advective acceleration
Pressure force
Gravity force
Friction force
Kinematic wave

where t = time, A = wetted area, Q = discharge, q = lateral inflow per unit length, x = distance downstream, y = depth, g = acceleration due to gravity, S_0 = bed slope, α = momentum coefficient and S_f = the friction slope.

A 1-D model is often used in the modeling of shallow and narrow rivers. These rivers are known to have less than 1m and 20m depth and width respectively (Ji 2008). In this regard, very often, assumptions help to get a manageable equation. The lateral and transverse velocities are very much small in comparison with the longitudinal velocity, and can be neglected without causing too much error (Jayawardena 2014).

The ‘St. Venant’ equations require numerical methods for finding solutions. These methods necessitate small time steps to avoid problem of numerical instability. Consequently, computation of the full ‘St. Venant’ equations becomes complex and time consuming. With this consideration, for WQ and quantity purposes, simplifications are usually made on **Equation 2-2** and results in equations such as kinematic wave (ignores pressure gradient term that is essential for the description of backwater effects) and diffusion wave (ignores the acceleration terms) (Cox 2003; Holguin-Gonzalez 2013).

As indicated in ‘St. Venant’ equation, solute transport in rivers is often a subject of the advection and dispersion processes. *Advection* is bulk movement of solute at the velocity of the

liquid. Whereas *dispersion* is solute conveyance by mixing and molecular diffusion induced by concentration gradients (Nirmalakhandan 2002; Ji 2008).

Diffusive transport at the molecular level under steady state conditions can be described mathematically using Fick's First Law of diffusion. Accordingly, the molar rate of transport in the x -direction, $J_{x,i}$ (MT^{-1}), is directly proportional to the concentration gradient, dC_i/dx , and the area of the face $\Delta y\Delta z$ at location x is (Nirmalakhandan 2002):

$$J_{x,i} = -D_i A_x \left(\frac{dC_i}{dx} \right) \quad (2-3)$$

where D_i = molecular diffusion coefficient, negative sign indicates that the flux is positive in the x -direction.

Dispersive transport results from a combination of multiple mechanisms, such as molecular diffusion, turbulence, eddy currents, and velocity gradients. The quantification of concentration profiles and chemical fluxes by dispersive transport follows the same model as that for diffusive transport but uses a dispersion coefficient, E (L^2T^{-1}) (Nirmalakhandan 2002). While the magnitude of the molecular diffusion hangs mainly on the chemical and fluid properties, turbulent diffusion and dispersion depend primarily on the flow regime (Metcalf & Eddy 2005).

The dispersion process in rivers is often less important in the transport of pollutants, when compared to the advection, which predominates due to the relatively high water velocity (Benedini & George 2013). Molecular diffusion, which takes place to a microscopic scale, is small compared with turbulent diffusion and therefore can be ignored (Jayawardena 2014).

In shallow rivers, two processes control longitudinal dispersion: transverse homogenization by turbulent diffusion and longitudinal stretching due to the horizontal depth

averaged velocity shear (Fischer et al. 1979). In this regard, the diffusive transport rate under time-dependent, *unsteady* state condition can be quantified using Fick's Second Law (Nirmalakhandan 2002; Benedini & George 2013):

$$\left(\frac{\partial C}{\partial t}\right)_{disp} = \varepsilon_i \frac{\partial^2 C}{\partial x^2} \quad (2-4)$$

where ε_i is dispersion coefficient.

Manning's n Equation: Those models such as QUAL2E (Q2E) (older version of Q2Kw-WQ model) with a steady-state approach drop the $\partial Q/\partial t$ term and then Manning equations may be used to obtain the depth and velocity in the channel (Cox 2003). In this regard, for trapezoidal channel, the Manning's equation becomes (Pelletier & Chapra 2008a):

$$Q = \frac{S_0^{1/2} A_c^{5/3}}{n P^{2/3}} \quad (2-5)$$

where Q is flow [m^3/s], n is Manning roughness coefficient, S_0 is bottom slope [m/m], P is wetted perimeter [m], and A_c is cross-sectional area [m^2].

2.6. Water Quality Modeling and Simulation Software

2.6.1. Historical Development of Water Quality Models

In the past few decades, WQ models development has made remarkable progress also encouraged by the advance in computer technology. River WQ modeling is considered to have started by Streeter and Phelps (1925) groundbreaking work. They found relationship between constant pollution loading of the Ohio River and the resulting sag in the Dissolved Oxygen (DO) concentration in the downstream water. This first-order kinetic model described variation of DO concentration due to organic carbon (i.e., characterized as BOD) decomposition because of

bacterial action. Models developed afterwards worked on improvement of this two-state-variable model in three major steps. In the first step, state variables such as carbonaceous biological oxygen demand (CBOD) and nitrogenous biological oxygen demand (NBOD) were distinguished. In addition, particulate matter settling rate, dissolved matter decay rate and sediment oxygen demand were introduced. The second step included ammonia, NO₃ and NO₂ from N cycle. The third stage introduced algae and P cycling, which resulted in state variables such as organic phosphorus (OP), ON, algae biomass (in terms of chlorophyll a) and dissolved phosphorus (DP) (Reichert *et al.* 2001). Similarly, Gao and Li (2014) divide historical development stages of WQ models into three. This is summarized in **Table 2-3**. A number of papers give historical development of these models (Somlyódy & Varis 1992).

Table 2-3 Historical developmental stages of watershed, river and stream models

Development stage	Year	Description
Primary stage	1925 – 1965	<ul style="list-style-type: none"> • In this initial period, models were focused on modifying and further developing of the Streeter-Pheleps models. • Models were limited to 1-D steady-state models for BOD-DO simulation
Improvement stage	1965 – 1995	<ul style="list-style-type: none"> • In this rapid model development period, models took further steps and included simulation of parameters such as N, P, T0, phytoplankton, biologic growth rate, sunlight, sediment among others. • Models capacity was expanded to 2-D and 3-Ds.
Broadening stage	1995 onwards	<ul style="list-style-type: none"> • New methods such as fuzzy logic, neural network, and GA were included in WQ models. • During this period, models have found application in environmental agencies for pollution control purpose

2.6.2. Water Quality Software Description

WQ models are developed for specified objectives and under varied degrees of complexity in response to complex environmental problems (Aral 2010). In line with this, data

requirement, mathematical computations (deterministic, stochastic), software requirement, dimensions of analysis (1-D, 2-D and 3-D), assumptions, etc. of models may vary (Bahadur *et al.* 2013). **Table 2-4** shows characteristics of some of the widely known models: Hydrological Simulation Program Fortran (HSPF), WQ Analysis Simulation Program (WASP), BASINS, Soil and Water Assessment Tool (SWAT), and Q2Kw (Benedini & George 2013). In this regard, 2-D and 3-D models require large amount of data. They are complex models and therefore usually applied in large rivers (Cox 2003).

Table 2-4 Characteristics of common watershed and WQ models

Model (Source, Year of release)	Model Overview	Model strengths	Model limitations	References
S-P (Streeter & Phelps, 1925)	<ul style="list-style-type: none"> • 1-D steady-state WQ model • Considers degrading BOD as the only “sink” for DO. • Applies advection (ideal PFR) as the only cause for downstream movement. 		<ul style="list-style-type: none"> • Considers Atm. Reaeration as the only “source” for DO • Disregards dispersion/Diffusion. 	(Streeter & Phelps 1925)
HSPF (US Environmental Protection Agency, 1980.)	<ul style="list-style-type: none"> • A nonproprietary, 1-D WQ and watershed hydrology model. • Simulates toxic organic and conventional pollutants such as pesticide fate and transport. 	<ul style="list-style-type: none"> • Simulates land and receiving water simultaneously. • Simulates point and nonpoint sources 	<ul style="list-style-type: none"> • Extensive calibration data and high level of expertise requirement • Limited to well-mixed rivers and reservoirs 	(SINGH & FREVERT 2006; Ambrose Jr <i>et al.</i> 2009; Gao & Li 2014)
WASP (US Environmental Protection Agency, 1983)	<ul style="list-style-type: none"> • A nonproprietary 1-D, 2-D and 3-D steady an unsteady state surface WQ model. • Dynamically simulates pollutants such as ammonia, DO, BOD, bacteria and eutrophication. • Mechanistic process description with deterministic data requirement. 	<ul style="list-style-type: none"> • Flexible in dimension terms • Applies user-defined timestep • Simulates point and nonpoint source 	<ul style="list-style-type: none"> • Complex and experiences unstability. • It requires an external hydrodynamic model to provide flow file for solving advection 	(Di Toro <i>et al.</i> 1983; Ambrose Jr <i>et al.</i> 2009; Sharma & Kansal 2013; Gao & Li 2014)
BASINS (US Environmental Protection Agency, 1996)	<ul style="list-style-type: none"> • 1-D watershed model and WQ model. 	<ul style="list-style-type: none"> • Consists of ArcView-GIS interface that links to environmental databases • Can model basin scale 	<ul style="list-style-type: none"> • Requires a large set of GIS data, • Requires ArcView 3.x, and the Spatial Analyst 	(Ambrose Jr <i>et al.</i> 2009)
SWAT (by Dr. Jeff Arnold for the USDA Agricultural Research Service (ARS), 1998)	<ul style="list-style-type: none"> • 1 - D river basin, or watershed hydrology and WQ model • Simulates WQ variables such as TSS, ON, NO₃/NO₂, NH₃/NH₄, OP, pesticides, metals etc. 	<ul style="list-style-type: none"> • Suitable for study of watersheds from small to very large sizes • It considers land use and soils data • Simulates point and nonpoint sources 	<ul style="list-style-type: none"> • does not simulate sub-daily events such as a single storm event and • diurnal changes of DO in a water body. 	(Douglas-Mankin <i>et al.</i> 2010; Vanderkruk <i>et al.</i> 2010)
Q2Kw	<ul style="list-style-type: none"> • 1-D steady-flow WQ model • Simulates conventional Pollutants such as N, P, DO, BOD, pH, Periphyton bottom algae and Pathogens. • Mechanistic process description with deterministic data requirement. • Applies numerical method for solving Advective-diffusion equations of pollutant transport. 	<ul style="list-style-type: none"> • It can simulate water exchange between surface water column and hyporheic zone and sediment pore-WQ. 	<ul style="list-style-type: none"> • simulates only the main stem of a river but not branches of the river system. 	(Pelletier & Chapra 2005; Kannel <i>et al.</i> 2011; Sharma & Kansal 2013; Gao & Li 2014)

2.6.3. Water Quality Models Application

2.6.3.1. General Condition: Application of Water Quality Models

In many countries, river WQ assessment solely relies on physicochemical standards (Holguin-Gonzalez 2013). Nevertheless, the need for integrated approach in river management was observed in international legislations such as the Clean Water Act of 1972 (Law, 1972) the WQ Act of 1987 (USEPA 2011) and the Water Framework Directive (WFD) (EU 2000). In this regard, US federal laws have a two-tier system in which sets effluent limitations based on available technology. In cases where this first approach becomes insufficient to meet WQ standards for a particular river, a further reduction in waste-load (WL) is determined through a WLA approach (Reichert *et al.* 2001).

Integrated WQ assessment intends to evaluate physical, chemical and biological nature of water bodies in terms of their natural quality, intended use and human effects (Vaughan *et al.* 2009). Furthermore, the aforementioned legislations encourage the use of WQ models as decision support tools in which the legal emission limit values, pollution sources characteristics and recipient quality standards are considered (Holguin-Gonzalez 2013). The role of WQ models in environmental decision-making is regarded essential because the models allow watershed based management approach, predictive capacity and cost-effectiveness (Gao & Li 2014). Consequently, they are applied in developed nations to remediation decision-making. This approach uses accurate and rich historical database maintained by scientific establishments and government agencies (Ongley & Booty 1999).

2.6.3.1. Water Quality Models Application in Ethiopia

WQ models application in developing countries for environmental decision-making is not common. Their intended application is hindered due to unavailability of required model input

data in quantity and quality, in addition to lack of experience in model based WQ decision-making (Tyson *et al.* 1993; Gu & Dong 1998; Ongley & Booty 1999; Deksissa *et al.* 2004).

Furthermore,

in these countries, the monitoring approach is not implemented regularly mostly due to budgetary constraint. This situation limits the decision-making and implementation ability of regulators in environmental agencies. Therefore, there is the need for environmental agencies in developing countries such as Ethiopia to begin modeling exercises as part of their management programs based on economical WQ monitoring approaches. Furthermore, these countries can workout future conditions for the application of modeling tools in the environmental decision making process. Some of hydrological and WQ based scholarly studies in Ethiopia are summarized in **Table 2-5**.

Although a number of literature works on watershed modelling about rivers, lakes and groundwater in Ethiopia were observed, they mostly deal with hydrology of the water bodies. The number of modeling works on WQ is very low even when compared to the not many hydrological modeling works on waterbodies in Ethiopia. Furthermore, specifically on the Little Akaki River (LAR), which is often regarded as one of the most polluted rivers in Ethiopia, no WQ modeling works were found during the literature survey that were conducted on LAR and use the conventional modeling tools such as SWAT, Q2Kw and the likes. However, in this regard, Tesfaye (2007) did a 1-D WQ (BOD and DO) model based on mass hydrodynamic and transport equations that was developed and numerically solved using a computer program in MATLAB platform.

Table 2-5 A survey of model based WQ and hydrology studies in Ethiopia

Waterbody	Studied waterbody	Objective	Model Used	Reference
River	Gilgel Abay River	Streamflow	SWAT	(Dile <i>et al.</i> 2013)
	Awash river basin	Water allocation	MODSIM	(Berhe <i>et al.</i> 2013)
	Blue Nile river basin	Sediment transport	SWAT	(Betrie <i>et al.</i> 2011)
	Didessa Sub-Basin	Water resources allocation	WEAP	(Adgolign <i>et al.</i> 2016)
	Maybar gauged watershed	Sediment transport	SWAT	(Yesuf <i>et al.</i> 2015)
	Blue Nile river basin	Sediment transport	SWAT	(Ayele <i>et al.</i> 2017)
	Omo-Geba river basin	Hydrological assessment	SWAT	(Chaemiso <i>et al.</i> 2016)
	Holetta river	Hydrological assessment	SWAT	(Tibebe <i>et al.</i> 2016)
	Tekeze dam watershed	Hydrological assessment	SWAT	(Welde & Gebremariam 2017)
	Gilgel Abay basin	Stream flow	TOPMODEL	(Gumindoga <i>et al.</i> 2014)
Anjeni-gauged watershed	Sediment transport	SWAT	(Setegn <i>et al.</i> 2010)	
Lake	Lake Tana	Sustainable water use	SWAT	(Dargahi & Setegn 2011)
	Lake Ziway watershed	Hydrological assessment	SWAT	(Desta & Lemma 2017)
Groundwater	Central main Ethiopian rift lakes basin	Groundwater-surface water flow interactions	MODFLOW	(Ayenew 2001)
	Lake Tana basin	Groundwater-surface water flow interactions	MODFLOW	(Chebud & Melesse 2009)
	Geba river basin	Groundwater resources assessment	MODFLOW	(Gebreyohannes <i>et al.</i> 2017)
	Tana sub-basin	Lake Tana and groundwater interaction	SWAT	(Abiy <i>et al.</i> 2016)

2.6.1. Evaluation and Selection of Water Quality Model

2.6.1.1. General

The description of a model relies on its constituent terms such as variables, constants, and parameters. In a crucial role, formulae connects variables to produce outputs for inputs. In this regard, models that yield only one output value that corresponds to a given input value is called deterministic model. While, model that is based on probabilistic relationship among the variables is called stochastic modeling (Benedini & George 2013).

An actual state of a river can be characterized more realistically using complex models such as 2-D or 3-D. However, these models are highly sophisticated and are commonly reserved for deep and wide rivers. These Rivers have complex mixing patterns and extensive data requirement (Cox, 2003). On the other hand, 1-D models are traditionally applied to river modeling and they are usually adequate for simulating hydrodynamic and WQ processes. For most small and shallow rivers, these 1-D models are commonly applied. In 1-D models, water surface elevation, velocity, and discharge vary only in the longitudinal (along - the - river) direction and are considered as constants in the lateral (across - the - river) direction. This approach provides a simplified mathematical description of river flows (Ji 2008).

2.6.1.2. Model Selection

In the model selection step, an appropriate, freely available and widely used WQ model for rivers and streams was sought. USEPA freely offers widely accepted and applied WQ models. For instance, some of the most appropriate models for rivers and streams include Q2E, QUAL2K (Q2K) and Q2Kw. These models have wide application in different parts of the world (Zhang *et al.* 2012).

- **Q2K**

Park and Lee (2002) developed Q2K after modification of Q2E (Brown & Barnwell 1987), which included the addition of new WQ interactions such as conversion of algal death to BOD, denitrification, and DO change caused by fixed plants. In addition, all interface operations are programmed in Microsoft *Excel*/ Visual Basic for Applications (*VBA*) environment. Moreover, excel serves as the graphical user interface. Furthermore, the model simulates attached bottom algae, pH, light extinction and pathogens (Chapra *et al.* 2007). However, Q2K is similar to Q2E in the following aspects: it is 1-D and has branching, steady state hydraulics, diel water-quality kinetics, diel heat budget and mass and heat inputs.

- **Q2Kw**

Pelletier *et al.* (2006) developed Q2Kw by modifying Q2K originally developed by Chapra and Pelletier (2003). Q2Kw is a public and well documented modeling framework with many new elements (Pelletier 2005). These new elements include conversion of algal death to CBOD, DO interaction with fixed plants, and reduction of CBOD due to de-nitrification. Additionally, it has auto calibration system and it is useful in data limited conditions (Pelletier *et al.* 2006; Kannel *et al.* 2007b; Sharma *et al.* 2015a).

Furthermore, Q2Kw can simulate WQ and hydrological conditions of shallow streams as well as systems with diffusive pollution loads. In this regard, its ability to simulate bottom algae and the hyporheic zone (area below the streambed where water percolates through spaces between the cobble and rocks) adds to its ability to simulate shallow streams (Pelletier *et al.* 2006). (Bottino *et al.* 2010).

Consequently, the model was successfully applied to rivers in different parts of the world: Cértima River in Portugal (Oliveira *et al.* 2012), river Yamuna in India (Sharma *et al.* 2015a), Bagmati River in Nepal (Kannel *et al.* 2007a), a small karstic watershed in Brazil (Camargo *et al.* 2010), Gharbia drain in the Nile Delta, Egypt (Allam *et al.* 2016) and plain of Chryssoupolis in Kavala, North Greece (Gikas 2014).

Q2Kw also had application in the determination of total maximum pollution load that a river can receive to meet WQ criteria (Kannel *et al.* 2007a). Furthermore, application of embedded Q2Kw in SO models for the purpose of optimal WLA modeling was observed (Cho & Lee 2014; Allam *et al.* 2016).. Therefore, in this study, Q2Kw is selected for simulation and WLA purpose of the LAR.

2.7. Q2Kw Description

2.7.1. General

Q2Kw represent rivers as a 1-D channel with steady and non-uniform flow. It simulates the fate and transport of conventional (i.e., non-toxic) pollutants (**Table 2-6**) caused by the impact both from non-point and point pollutant loadings (Pelletier *et al.* 2006; Pelletier & Chapra 2008a).

Table 2-6: Model state variables

Variable	Symbol	Units*
T ⁰		°C
Conductivity	s ₁ , s ₂	µmhos
Inorganic suspended solids (ISS)	m _{i,1} , m _{i,2}	mgD/L
DO	o ₁ , o ₂	mgO ₂ /L
Slow-reacting CBOD	c _{s,1} , c _{s,2}	mg O ₂ /L
Fast-reacting CBOD	c _{f,1} , c _{f,2}	mg O ₂ /L
ON	n _{o,1} , n _{o,2}	µgN/L
NH ₃ -N	n _{a,1} , n _{a,2}	µgN/L
NO ₃ -N	n _{n,1} , n _{n,2}	µgN/L
OP	p _{o,1} , p _{o,2}	µgP/L
Inorganic phosphorus	p _{i,1} , p _{i,2}	µgP/L
Phytoplankton	a _{p,1} , a _{p,2}	µgA/L
Detritus	m _{o,1} , m _{o,2}	mgD/L
Pathogen	x ₁ , x ₂	cfu/100 mL
Generic constituent	gen ₁ , gen ₂	user defined
Alkalinity	Alk ₁ , Alk ₂	mgCaCO ₃ /L
Total inorganic carbon (TOC)	c _{T,1} , c _{T,2}	mole/L
Bottom algae biomass	a _b , a _h	gD/m ²
Bottom algae N	IN _b	mgN/m ²
Bottom algae P	IP _b	mgP/m ²

* mg/L ≡ g/m³, D = Dry weight, A = chlorophyll a

2.7.2. Simulation Equations

- **Mass Balance**

Q2K solves governing 1-D advection-diffusion equation (ADE) in a steady-state condition for a constituent concentration (c_i) in the water column of a stream reach i . This gives a general mass balance equation (**Figure 2-1**) that can be expressed as (Pelletier *et al.* 2006; Pelletier & Chapra 2008a):

$$\frac{dc_i}{dt} = \frac{Q_{i-1}}{V_i} c_{i-1} - \frac{Q_i}{V_i} c_i - \frac{Q_{ab,i}}{V_i} c_i + \frac{E_{i-1}}{V_i} (c_{i-1} - c_i) + \frac{E_i}{V_i} (c_{i+1} - c_i) + \frac{W_i}{V_i} + S_i \quad (2-6)$$

where Q_i = flow [m^3/d , ab = abstraction], V_i = volume (m^3), E_i = the coefficient of bulk dispersion between reaches i and $i+1$ [m^3/d], S_i = sources and sinks of the constituent due to reactions and mass transfer mechanisms [$g/m^3/d$ or $mg/m^3/d$] and W_i = the external loading of the constituent to reach i [g/d or mg/d].

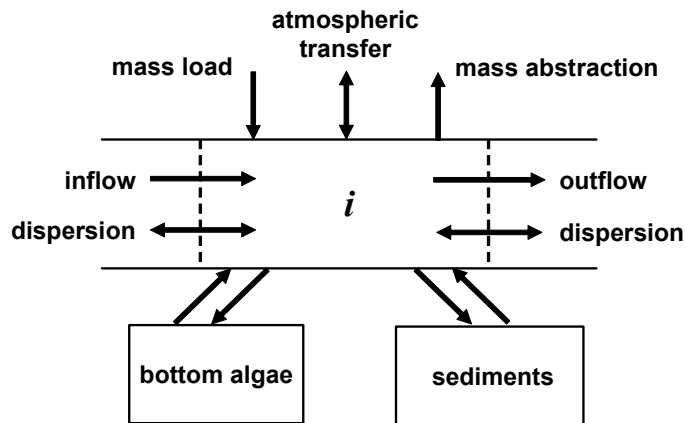


Figure 2-1 Mass balance for constituents in reach segment “i”

In this regard, Q2Kw assumes the deposition of settled ISS, phytoplankton, and detritus from the water column layer to the sediment diagenesis zone and do not enter the hyporheic pore water. The sediment diagenesis sub-model accounts for anaerobic metabolism of settled material

in the sediment. On the other hand, the hyporheic sub-model accounts for aerobic metabolism of heterotrophic bacteria in the hyporheic zone (Pelletier & Chapra 2008a).

Complete list and description of equations for constituents and general mass balance are available in Q2Kw theory and documentation report (Pelletier & Chapra 2008b). In this work, equations for some of the parameters are given (Sharma *et al.* 2015a):

$$\begin{aligned} \frac{dc_{2,i}}{dx} &= S_{2,i} + \frac{E'_{hyp,i}}{V_{2,i}}(c_i - c_{2,i}); \frac{dab,i}{dx} = S_{b,i}; \frac{dIN_b}{dt} = S_{bN,i}; \frac{dIP_b}{dt} \\ &= S_{bp,i}; \frac{da_{h,i}}{dt} = S_{ah,i} \end{aligned} \quad (2-7)$$

DO

$$\begin{aligned} S_0 &= r_{oa}(\text{PhytoPhoto} - \text{PhytoResp}) \\ &+ r_{od}(\text{BotAlgPhoto} - \text{BotAlgResp}) \frac{A_{st,i}}{V_i} - r_{oc}\text{FASTCOxid} \\ &- r_{oc}\text{SlowCOxid} - r_{on}\text{NH}_4\text{Nitr} + \text{Reaeration} - \text{COD}_{oxid} \\ &- \text{SOD} \frac{A_{st,i}}{V_i} \\ &= r_{oa} \cdot (k_{gp}\phi_{Np}\phi_{Lp}a_p - F_{oxp} \cdot k_{rp} \cdot a_p) \\ &+ r_{od}[\text{BotAlgPhoto} \\ &- F_{oxb}(k_{rb1} \cdot a_b - k_{rb2}\text{BotAlgPhoto})] \frac{A_{st,i}}{V_i} - r_{oc} \\ &\cdot (F_{oxcf} \cdot k_{dc} \cdot C_f) - r_{oc} \cdot (F_{oxcf} \cdot k_{dcs} \cdot C_s) - r_{on}(F_{oxnd}k_n n_a) \\ &+ \text{Reaeration} - k_{COD}[\text{COD}] - \text{SOD} \frac{A_{st,i}}{V_i} \end{aligned} \quad (2-8)$$

$$\text{Reaeration} = k_a[(e^{\ln 0.5}(1 - 0.0001148 \cdot \text{elev}) - 0)] \quad (2-9)$$

$$\text{InOs} = -139.3411 + \frac{1.575701 \times 10^5}{T_a} - \frac{6.642308 \times 10^7}{T_a^2} + \frac{1.2438 \times 10^{10}}{T_a^3} - \frac{8.621949 \times 10^{11}}{T_a^4} \quad (2-10)$$

CBODu

$$S_{cs} = r_{od} \text{DetrDiss} - \text{SlowChydr} - \text{SlowCOxid} - \text{SlowCHydr} = k_{hc}(T)C_s \quad (2-11)$$

$$S_{cf} = \text{SlowCHydr} - \text{FastCOxid} - r_{ondn} \text{Denitr} + J_{CH_4,d} \frac{A_{st,i}}{V_i} \quad (2-12)$$

$$\text{FastCOxid} = F_{oxcf} k_{dc}(T) f \text{SlowCOxid} = F_{oxcf} k_{dcs}(T) C_s \quad (2-13)$$

$$S_{DO} = r_{na} \cdot (k_{dp} \cdot a_p) + qN \left(\frac{k_{db} a_b}{H} \right) - k_{hn} \cdot n_o - \frac{v_{on}}{H} \cdot n_o \quad (2-14)$$

Nitrogen

$$S_{na} = k_{hn} \cdot n_o - F_{oxna} \cdot k_{na} + r_{na} (F_{oxp} k_{rp}) a_p - r_{na} P_{ap} (k_{gp} \phi_{Np} \phi_{Lp}) a_p + [J_{NH_4} - P_{ab} \text{BotAlgUptakeN}(\text{NU}_p \text{WCfrac})] \frac{A_{st,i}}{V_i} \quad (2-15)$$

$$P_{ap} = \frac{n_a n_n}{(k_{hnxp} + n_a)(k_{hnxp} + n_n)} + \frac{n_a k_{hnxp}}{(n_a + n_b)(k_{hnxp} + n_n)} \quad (2-16)$$

$$P_{ab} = \frac{n_a n_n}{(k_{hnxb} + n_a)(k_{hnxb} + n_n)} + \frac{n_a b}{(n_a + n_b)(k_{hnxb} + n_n)} \quad (2-17)$$

$$\begin{aligned}
S_{nn} = & F_{oxna} \cdot k_{mn} \cdot n_a - (1 - F_{oxdn})k_{dn}n_n - r_{na}(1 - P_{ap}) \\
& \cdot (k_{gp}\phi_{Np}\phi_{Lp})a_p \\
& + [J_{NO_3} - (1 - P_{ab})BotAlgUpTakeN(NU_pWCfrac)]
\end{aligned} \tag{2-18}$$

Notations: Ah = biofilm of attached heterotrophic bacteria in the hyporheic sediment zone; ap = phytoplankton concentration (mgA/m³); ab = bottom algae concentration (gD/m²); A_{st,i} = surface area of the reach (m²); BotAlgPhoto = bottom algae photosynthesis (gO₂/m²/day); BotAlgUptakeN = uptake rate for N in bottom algae (mgN/m²/day); BotAlgUptake P = uptake rate for phosphorous in bottom algae (mgP/m²/day); BotAlResp = bottom algae respiration (gO₂/m²/day); c_{2,i} = concentration in hyporheic sediment zone (mg/L); c_i = concentration in the surface water in reach i (mg/L); c_{i-1} = concentration in the u/s reach i - 1 (mg/L); COD_{oxid} = oxidation of non-carbonaceous non-nitrogenous chemical oxygen demand (gO₂/m²/day); Denitr = rate of denitrification [mgN/L/day]; E'_{i-1}, E'_i = bulk dispersion coefficients between reaches i - 1 and i and i and i + 1 (m³/day), respectively; E'_{hyp,j} = bulk dispersion coefficients between hyporheic zone and reach i (m³/day); FastCO_{oxid} = fast CBOD oxidation (gO₂/m²/day); F_{oxcf} = attenuation due to low oxygen [dimensionless]; F_{oxna} = attenuation due to low oxygen on ammonia nitrification (dimensionless); F_{oxp} = attenuation due to low oxygen on phytoplankton respiration; H_{2,i} = the thickness of the hyporheic zone (cm); IN_b = intracellular N concentration in bottom algae (mgN/m²); IP_b = intracellular P concentration in bottom algae (mgP/m²); J_{CH₄,d} = the sediment flux of dissolved methane in oxygen equivalents (gO₂/m²/day); J_{NH₄} = sediment flux of NH₃(mgN/ m²/day); J_{NO₃} = sediment flux of NO₃ (mgN/m²/day); J_{PO₄} = sediment flux of inorganic P (mgP/m²/day); k_a = reaeration rate (1/day); k_{dc}(T), k_{dcs}(T) = T⁰-dependent fast CBOD oxidation rate [day]; k_{dp} = phytoplankton death rate (/day); k_{gp} = maximum photosynthesis rate at T⁰ (day); k_{hc}(T) = T⁰-dependent slow CBOD hydrolysis rate [day]; k_{hn} = ON hydrolysis rate

(1/day); k_{hnxb} = preference coefficient of bottom algae for ammonium (mgN/m^3); k_{hnxp} = preference coefficient of phytoplankton for ammonium (mgN/m^3); k_{na} = nitrification rate for $\text{NH}_3\text{-N}$ (1/day); k_{nn} = T^0 -dependent nitrification rate for $\text{NH}_3\text{-N}$ (1/day); k_{rp} = phytoplankton respiration rate (1/day); NH_4nitr = ammonium nitrification ($\text{gO}_2/\text{m}^2/\text{day}$); NUpWCfrac = fraction of N uptake from the water column by bottom plants; Os = saturation concentration of $\text{DO}(\text{mgO}_2/\text{L})$; Pab = preferences for ammonium as a N source for bottom algae; Pap = preferences for ammonium as a N source for phytoplankton; Phytophoto = phytoplankton photosynthesis ($\text{gO}_2/\text{m}^2/\text{day}$); PhytoResp = Phytoplankton respiration ($\text{gO}_2/\text{m}^2/\text{day}$); PUpWCfrac = fraction of P uptake from the water column by bottom plants; $\text{Q}_{ab,i}$ = total flow abstractions from the reach i ; Q_i = outflow from reach i ; Q_{i-1} = inflow from the u/s reach $i - 1$; qN = actual cell quotas of N (mgN/gD); qP = actual cell quotas of phosphorous (mgP/gD); Reaeration = ($\text{gO}_2/\text{m}^2/\text{day}$); r_{na} = ratio of N to chlorophyll a (mgN/mgA); r_{od} = ratio of oxygen consumed to detritus (mgO_2/mgD) during bottom algae respiration; r_{on} = ratio of oxygen consumed to N during nitrification (mgO_2/mgN); r_{oa} = ratio of oxygen generated to phytoplankton growth (mgO_2/mgA); r_{oc} = ratio of oxygen consumed during CBOD oxidation ($\text{mgO}_2/\text{mgO}_2$); r_{pa} = ratio of P to chlorophyll a (mgP/mgA); $\text{S}_{2,i}$ = sources and sinks of the constituent in the hyporheic sediment zone due to reactions; $\text{S}_{ah,i}$ = sources and sinks of heterotrophic bacteria in the hyporheic sediment zone due to reactions ($\text{gD}/\text{m}^2/\text{days}$); $\text{S}_{b,i}$ = sources and sinks of bottom algae biomass due to reactions ($\text{gD}/\text{m}^2/\text{day}$); $\text{S}_{bN,i}$ = sources and sinks of bottom algae N due to reactions ($\text{mgN}/\text{m}^2/\text{day}$); $\text{S}_{bp,i}$ = sources and sinks of bottom algae P due to reactions ($\text{mgP}/\text{m}^2/\text{day}$); S_i = sources and sinks of constituent i due to reactions and mass transfers ($\text{mg}/\text{L}/\text{day}$); SlowCOxid = slow CBOD oxidation ($\text{gO}_2/\text{m}^2/\text{day}$); Ta = absolute T_0 ($^\circ\text{K}$), elev = elevation of the area (m); $\text{V}_{2,i} = (\phi_{s,i}A_{st,i}H_{2,i}/100)$ volume of pore water in the hyporheic

sediment zone (m^3); V_i = volumes of reach i (m^3), t = time (day); v_i = ISS' settling velocity (m/day); W_i = external loading of the constituent to reach i (mg/day); ϕ_{Lp} = phytoplankton light attenuation factor (dimensionless); ϕ_{Np} = phytoplankton nutrient attenuation factor (dimensionless); and $\phi_{s,i}$ = porosity of the hyporheic sediment zone.

2.7.2.1. Hydraulic Characteristics

- Manning's n Equation

Q2Kw allows three ways for velocity and depth relationships: rating curves, weirs, and Manning equations. For steady flow, the Manning equation is represented as (Pelletier & Chapra 2008a):

$$Q = \frac{S_0^{1/2}}{n} \frac{A_c^{5/3}}{P^{2/3}} \quad (2.19)$$

where Q = flow [m^3/s], n = the Manning roughness coefficient, S_0 = bottom slope [m/m], P = the wetted perimeter [m], and A_c = the cross-sectional area [m^2].

- Longitudinal Dispersion

Q2Kw computes longitudinal dispersion using channel's hydraulics (Fischer et al. 1979).

$$E_{p,i} = 0.011 \frac{U_i^2 B_i^2}{H_i U_i^*} \quad (2.20)$$

where $E_{p,i}$ = the longitudinal dispersion between reaches i and $i + 1$ [m^2/s], U_i = velocity [m/s], B_i = width [m], H_i = mean depth [m], and U_i^* = shear velocity [m/s].

2.8. Multi-objective Optimization

2.8.1. General

In Multi-objective (MO) optimization, a set of tradeoff solutions are required to arrive at a specific and compromised decisions. Thus MO optimization final solution relies on the practice of optimization and decision making (Tan *et al.* 2005). Schütze *et al.* (2011) provides classification of local and global optimization methods shown in **Figure 2-2**.

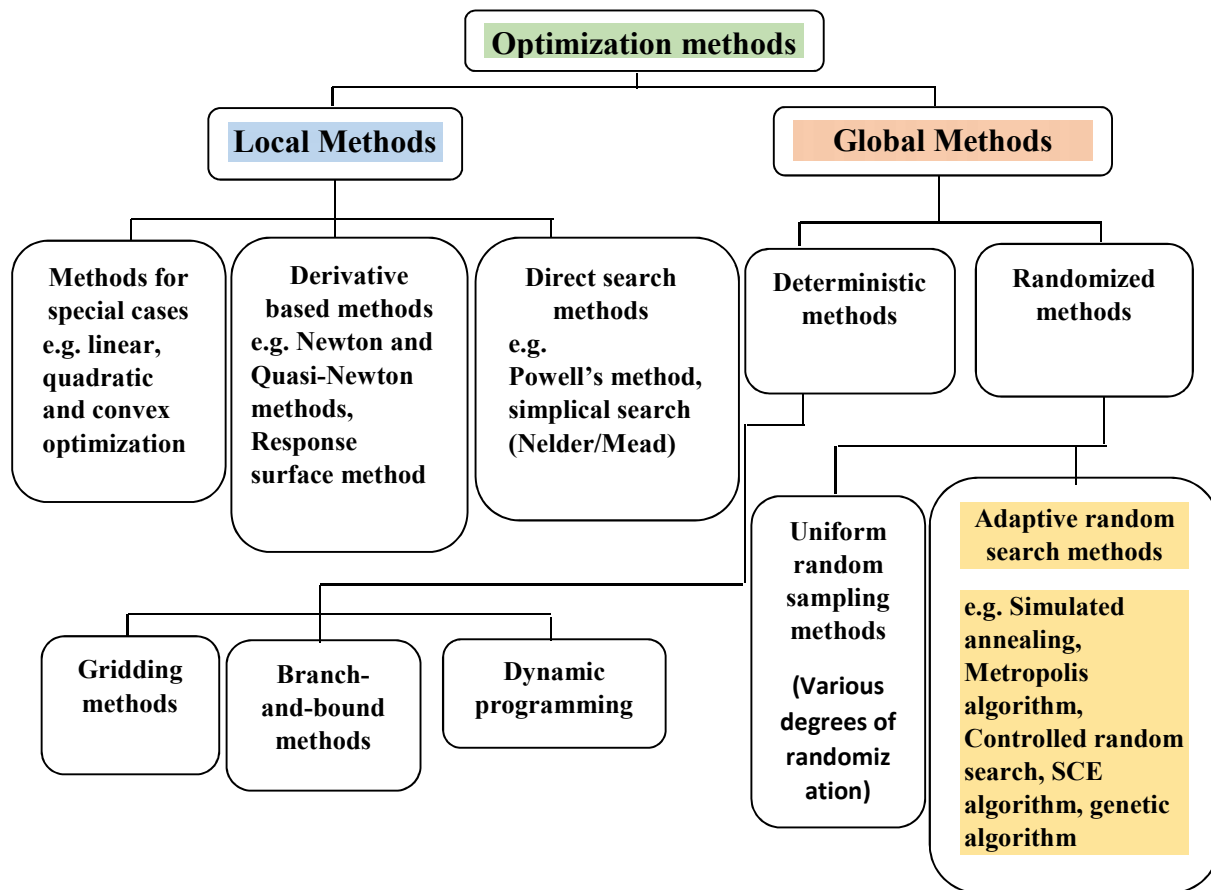


Figure 2-2: Classification of search and optimization methods

Local optimization methods do not ensure the attainment of a global minimum. On the other hand, Global optimization techniques do not get stuck in a local minimum. However, these

methods, usually require a higher number of function evaluations necessary to obtain a solution (Schütze *et al.* 2011).

There are several methods for solving MO optimization problems reported in the literature, such as GAs, variable neighborhood search, greedy randomized adaptive search procedure (GRASP), simulated annealing, tabu search, ant colony, and others (Andrade *et al.* 2013). GA has been selected as a solution method to solve the problem due to some of its effectiveness and inherent advantages over other MO solution methods (Simonovic 2009b). Consequently, it is the most widely used algorithm of the randomized global methods (Schütze *et al.* 2011).

2.8.2. Genetic Algorithm (GA)

GA is a search algorithm developed using the theory of breeding and natural selection (Wardlaw & Sharif 1999). It evaluates performances of candidate solutions at multiple points simultaneously. This is very suitable for solving MO optimization problems (Tan *et al.* 2005). In this context, GA readily facilitates the application of a simulation model to evaluate WQ response (Burn & Yulianti 2001). Consequently, its application for optimal WLA problems was observed in previous papers (Burn & Yulianti 2001; Mujumdar & Subbarao 2004; Murty *et al.* 2006; Allam *et al.* 2016).

The first task in applying GA is representing a potential solution of the optimization problem with a suitable coding structure, treated as a chromosome (Mujumdar & Subbarao 2004). Afterwards, each point in the "population" of trials is assigned a fitness value. The present population is subjected to a selection pressure relying on their respective computed fitness. Subsequently, the parameter values are modified with genetic operators of crossover and mutation in order to uncover new areas of parameter-space (Metcalf & Charbonneau 2003).

The crossover operation exchanges important building blocks between two strings that perform well. Applying this procedure, the GA aims to produce new strings that preserve the best material from two parent strings. Mutation allows for the random mutations of bits of information in individual genes. As a result, it introduces new genetic material to a population. In addition, elitism ensures retention of the best individual of a population between generations. This maintains a high and improving maximum fitness (Wardlaw & Sharif 1999).

Some examples of MO evolutionary algorithms are Vector Evaluated GA (VEGA) (Schaffer 1985), the non-dominated sorting GA (NSGA) (Srinivas & Deb 1994) and PIKAIA GA (PIKAIA-GA) (Charbonneau & Knapp 1995).

PIKAIA-GA is an optimization subroutine adaptable for use in a wide variety of modeling applications. It was written in Fortran 77 code (Charbonneau & Knapp 1995). Afterwards, it was translated to Excel VBA for use in Q2Kw WQ model (Pelletier *et al.* 2006). Internally, PIKAIA GA seeks to maximize a (user-defined) function $f(x)$ in a bounded n-dimensional (n-D) space as indicated in **Equation 2-21** (Charbonneau & Knapp 1995).

$$x \equiv (x_1, x_2 \dots x_n) \quad x_k \in [0.0, 1.0] \quad (2-21)$$

PIKAIA-GA carries out its maximization task on a population made up of fixed np individuals over a fixed, preset number of generations (Charbonneau & Knapp 1995). It has found application for optimization purpose in different works (Metcalf & Charbonneau 2003; Ghenaïet & Tayeb 2009).

In this paper, PIKAIA-GA with embedded Q2Kw is proposed for the MO optimal WLA problem.

2.9. Simulation-Optimization Models

Simulation models are mathematical models that represent the reality in descriptive way. In Optimization (or programming) models, it is important to identify an objective function to be optimized with the involved variables and in view of a predetermined criterion of operation. Generally, water related problems cannot be thoroughly examined by means of a single model, and several models are often necessary. Both simulation and optimization, that are interconnected are required in order to achieve a reliable result.(Benedini & George 2013)

In this regard, SO techniques are useful in WQ prediction and protection (Tong & Chen 2002; Santhi *et al.* 2006). These models provide the advantages of the two modules within the bounds of a single framework. As such, SO models provide modeled effect of pollution discharge for the purpose of optimal WL reduction. Thus, these models are convenient for WQ decision making (Mujumdar & Subbarao 2004).

2.10. Waste Load Allocation

2.10.1. General Description

The WLA problem is MO because it deals with problems with two or more objective functions (Cohon 1978b). The potential of a river to clean itself assignable to its assimilative capacity determines the amount of pollution it should receive. WLA modeling aims to distribute WL among polluters based on assimilative capacity of the receiving water body. In this regard, the approach tries to optimize WL reduction in an equitable manner. As a result, it differs from the traditional WQ management approach that seeks to maximize WL reduction from dischargers. This model is MO and includes WQ at monitoring sites, equity measure, WL removal fraction and performance measures. Equity measure makes possible the fair allocation of WL among polluters. Performance measure represents constraints such as number of

checkpoints with DO violations, magnitude of maximum DO violation, and total magnitude of DO violations at the checkpoints. The solution of an optimal MO-WLA model yields a trade-off curve, or surface, identifying a collection of points that define efficient solutions to the problem (Burn 1987; Tung 1992a; Burn & Yulianti 2001; Yandamuri *et al.* 2006).

This tradeoff curve can be used by a WQ agency to identify an appropriate solution for implementation under conflicting objectives inherent to the problem (Cho & Lee 2014). Furthermore, objective functions in WLA models shape decision makers perception and action on the pollution problem. Hence, for optimal solutions, models involving more objectives extending to social, economic, legal and environmental aspects should be explored (Cohon 1978a; Tung & Hathhorn 1989; de Andrade *et al.* 2013). Simultaneous application of several objectives is also recommended in the search for better WLA model outputs (Tung 1992a). For example, cost-equity formulations do not allow performance measures (i.e., DO standard) as a constraint in the formulation. This hinders decision maker desire to arrive at a cost-equity trade-off solution for a desired performance level (Yandamuri *et al.* 2006).

2.10.2. WLA Problem Formulations

The formulation of a WLA model depends on the objective function and the prevailing condition which also determines the constraints. It considers more than one objective, for example, maximizing economic efficiency and measures of WQ. Another objective commonly discussed in WLA is equity, for example, allocation of wastewater discharge fees in proportion to treatment costs (Andrade *et al.* 2013). According to (Tung 1992b), failure to include an equity measure in the WLA model could result in the allocation of large quantities of waste to the upstream users, whereas the downstream dischargers could be required to treat their effluent at levels of maximum possible efficiency (Andrade *et al.* 2013).

The minimization of treatment cost and maximization of equity among the polluters are conflicting objectives. Maximizing equity results in increased treatment cost, which may be undesirable from economic point of view whereas minimizing treatment cost leads to reduced equity among the polluters, which may not also be desirable for Pollution Control Agency (PCA). Hence, it would be appropriate to solve the *WLA* problem as a MO model so that tradeoff curve can be obtained between treatment cost and equity (Park 2010). From this tradeoff curve, a suitable solution point that satisfies both polluters and PCA could be selected. (Simonovic 2009b). The following are some of the works on WLA model formulations.

Two of the three GA based optimization models formulated in Burn and Yulianti (2001) are: cost versus quality model (i.e., minimize total treatment cost and minimize the sum of the violations of WQ standards) and cost versus equity model (i.e., cost minimization and equity in the prescribed treatment). Yandamuri *et al.* (2006) proposed total treatment cost, equity among waste dischargers, and DO violation performance measure (i.e., expressed as a weighted sum of the individual performance measures of the number of DO limit violations and the magnitude of the maximum and total DO limit violation) for an optimal MO WLA problem. de Andrade *et al.* (2013) formulated models using WL removal efficiencies, equity measure, limits for the concentration of DO and BOD. Cho *et al.* (2003) and Du *et al.* (2013) proposed WLA model using Total Nitrogen (TN), BOD, and TP as WQ indicators. Recently, Allam *et al.* (2016) constructed WLA model for a MO optimization problem, with the aim of minimizing WL abatement costs, while maximizing overall performance to comply with water reuse standards in irrigation systems.

The aforementioned works successfully applied WLA models for WQ management. In those papers, DO, BOD, TN and TP are the most common performance measure parameters.

Moreover, these parameters were used simultaneously and in different combinations. However, to the best of the authors' knowledge, no previous work was observed that treats DO, BOD, TN and TP as performance measures simultaneously, in an integrated Q2Kw and GA based optimal WLA model framework

2.10.3. WLA Models Applicability

The application of the WQ objective approach in the regulatory practice would rely heavily on the site conditions, and the specific geographic, climatic, economic and social factors. For example, for smaller rivers and streams that are ephemeral and are caused mainly by point effluent discharges, the natural assimilative capacity of the body cannot be considered. Rather, an effluent orientated approach would be feasible. In cases of larger surface water bodies, a WQ orientated approach for protection and pollution prevention, together with assessments of their assimilative capacity would be feasible .(Hranova 2006a).

3. MATERIALS AND METHODS

3.1. Study Area

Ethiopia is located in the horn of Africa and shares borders with Eritrea to the North, Sudan and South Sudan to the West, Kenya to the South and Somalia and Djibouti to the East. It covers a total area of 1.13 million square kilometers (Km²) with a topography of massive highlands, rugged terrain and low plains. The elevation ranges between two extremes that is from 125 m below mean sea level (bmsl) at Danakil depression to 4,620 meters (m) above mean sea level (amsl) at Ras-Dashen peak (Berhanu *et al.* 2014; Hailemariam 2017).

Administratively, Ethiopia is a Federal Democratic Republic composed of nine National Regional States and two Chartered Cities: Addis Ababa (AA) and Dire-Dawa (Hailemariam 2017). AA (**Figure 3-1**) is the capital city of Ethiopia, the seat of Head Offices of the African Union and United Nations Economic Commission for Africa. Besides, in the City, there are more than 100 embassies and many international Aid and Development organizations. Administratively, the City, with an area of 540 km², is subdivided into 10 sub-cities (namely: Gulele (G), Yeka (Y), Addis-Ketema (AK), Arada (A), Kolfe-Keranio (KK), Lideta (L), Kirkos (K), Bole (B), Nefas silk-Lafto (NL), and Akaki-Kality (AKK)) and 116 districts (Wereda) (FDREMT 2011).

AA is geographically located between 465000m and 485000m East and 980000m and 1005000m North. The city is on the southern flank of Intoto ridge (3,139 m amsl) having an average elevation of 2,600m amsl. The topography slopes down from the Intoto mountain in the North to the Southern border of the city, with a number of steep-sided valleys, rivers and streams. The center of AA lies on a rolling terrain with some flat land areas (Gizachew 2008).

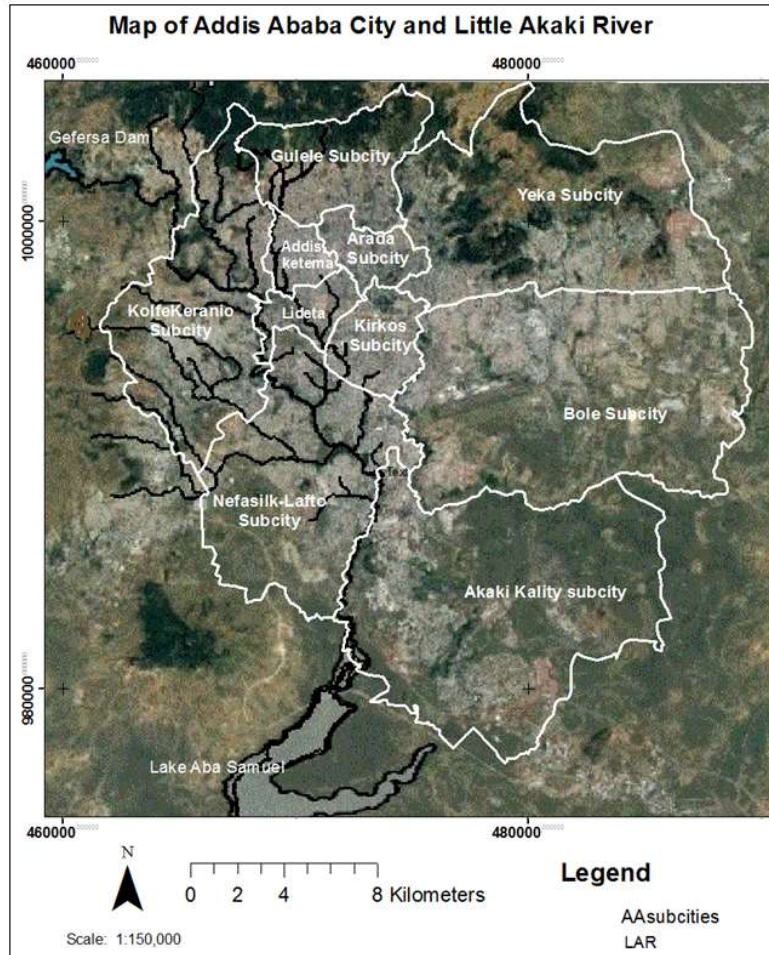
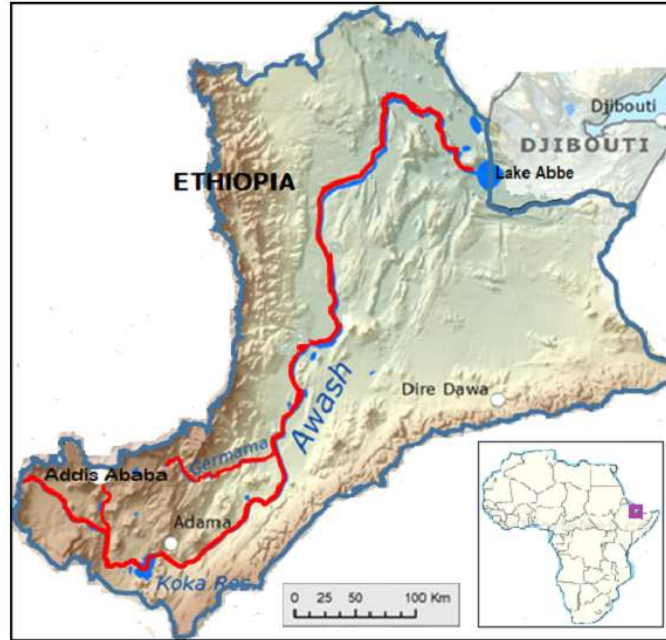


Figure 3-1 Map of Addis Ababa

3.2. Little Akaki River

The Akaki Rivers in AA are part of the greater Awash River Basin. The BAR drains the Eastern and Central part of the catchment area while the LAR drains the Western part of the catchment. These two rivers, with 1,462 km² catchment area, form one of the biggest tributaries of the Awash River basin (**Figure 3-2**). This basin, with a total coverage area of 110,000 km², begins near Ginchi Town at Dandi Wereda, 80km West of Addis Ababa. And, ends at Abbe Lake at the border with Ethiopia and Djibouti (Fasil *et al.* 2013).



Adopted from Wikipedia.org

Figure 3-2 Awash river basin

In the downward flow of the Rivers to the South direction, there is an elevation drop of about one kilometer (km) in a space of thirty km. The two rivers enter Abba Samuel Lake at the outskirts of AA before flowing down to Koka Dam (Oljira 2006).

Collection of domestic wastewater in Addis Ababa is very low that only 25-30% of households have wastewater collection system that uses either vacuum trucks or piped sewer line (WBG 2015). The LAR in Addis Ababa, due to deteriorating water quality condition, has not found services that are worth mentioning except urban irrigation (Figure 3-3). Even this service has raised public concern because of health hazards from vegetables irrigated with polluted water (Aschale 2015).



Figure 3-3 Partial view of irrigated land along LAR near Mekanisa

3.2.1. Land Use

Table 3-1 shows the general land use condition in Addis Ababa. The most upstream part of LAR, in Gulele sub-city, is covered with dense vegetation. Moreover, Western and Southern parts of Nefas-Silk Lafto sub-city; and Northern and Southern parts of Kolfe-Keranio sub-city have sizable vegetation coverage. Urban irrigation activities were observed along the LAR in Addis-Ketema, Lideta, Kolfe-Keranio, kirkos and Nefas-Silk Lafto sub-cities.

Table 3-1 Addis Ababa City Land Use Condition, 2011

Land cover classes	Area (%)
Forest	11
Urban built up	40
Agriculture and Vegetation	27
Bare land	22
Total	100

Source: (Arsiso et al. 2018)

All sub-cities in LAR catchment have a number of industries, however by proportion Nefas-Silk Lafto has the most of all. **Table 3-2** lists some examples of industries in the different sub-cities.

On the other hand, although all sub-cities in LAR catchment have mixed residential and commercial use, Addis Ketema sub-city is probably the most utilized in this regard.

Table 3-2 Partial list of industries in the sub-cities of LAR catchment

	Sub-Cities					
	Gullele	Addis Ketema	Kolfe Keranio	Lideta	Kirkos	Nefas-silk Lafto
Industries	Ethio-Marble	ABYA Mechanical Engenering PLC/ Metal	Tikur Abay shoe factory	St. George Brewery	Addis Ababa Abbatoirs	Addis Ababa Cement factory
	Gullele Soap & Detergent	Dire ind. Plc (leather)	Repi Soap Detergent	National Tobacco	Moha Soft Drinks	Addis Modjo Edible Oil
	Ediget garment	Dil Edible Oil Factory	Samjo Cosmotics	Awash Winery	Chora Gas & Chemicals	Batu Leather
			Gion industry chemical	East Africa Bottling	Tana enginering	Family Mil
			Natran plastic	Lejo Auluminium PLC	Misrak Flour Factory	Horizon Addis Tyre
			Addis Abeba bottle & Glass factory	Anbessa flour Factory	Shewa Bakery	National Alchol & Liquor
			K.O.J.J. food processing complex	National Alchol & Liquor Factory	Yekatit paper converting enterprise	Star Soap & Detergent
			Addis Abeba Tannery	Anbessa Shoes Factory	Moha Soft Drinks co2 Factory	Akyd Resin
				Artistic Printing	Nifas Silk Paints	Kadisco Paint

3.2.2. Population

The world population was 7.6 billion in 2017 and it is projected to reach 9.8 billion in 2050. Half of this population growth is expected to be concentrated in just nine countries: India, Nigeria, the Democratic Republic of the Congo, Pakistan, *Ethiopia*, the United Republic of Tanzania, the United States of America, Uganda and Indonesia (ordered by their expected contribution to total growth) (UN-DESA 2017).

Ethiopia is the second most populous country in Africa (Hailemariam 2017). Ethiopia's and AA's population in 2016 was 92.2 million and 3.35 million respectively (CSA 2013). Moreover, in 2016, the average population density in AA city was 6,208 per km² (Table 3-3).

Table 3-3 Population and settlement density in the AA Sub-Cities

Sr. No	Sub-City	Projected Population (2016)		Population Density/km ²
		Number	%	
1	Gulele	327,426	9.8	10,004
2	Kolfe Keranio	524,729	15.7	8,060
3	Addis Ketema	312,414	9.3	34,790
4	Arada	258,808	7.7	22,388
5	Lideta	246,805	7.4	19,904
6	Yeka	424,217	12.7	5,155
7	Kirkos	270,721	8.1	16,650
8	Nefas-Silk Lafto	387,017	11.5	6,086
9	Bole	378,104	11.3	3,127
10	Akaki Kality	221,759	6.6	1,759
Total Population		3,352,000	100	6,208

Source: (CSA 2013)

The three least densely populated sub-cities were Akaki-Kality, Bole and Yeka. However, Addis Ketema, Arada and Lideta sub-cities were the three most densely populated sub-cities (CSA 2013). These mostly populated sub-cities belong to the central and older part of the city and encompass the City slum areas of the City. Furthermore, AA faces water scarcity because of rapid urbanization and increased individual water demand and the per capita distribution is estimated to be around 40 liters/day (L/d), which is well below the city's goal of 110 L/d (WBG 2015).

Some of the factors that affect wastewater disposal management are urban population numbers, settlement conditions and water use patterns (UN-Habitat 2001). All human settlements use natural resources and subsequently generate waste. Its impact depends on availability of water supply, sanitation and solid waste management. Especially informal settlements and slums in urban areas exacerbate pollution problem (UN-Habitat 2002). For example, increase in urban population results in addition of organic wastes and overloading of sewer facilities. This ultimately pollutes local rivers and streams. Also, informal housing establishments without connection to municipal sewers exacerbate this problem (Mvungi *et al.* 2003; Das *et al.* 2010).

3.2.3. Climate

The Ethiopian climatic zones, based on altitude difference, are five: Alpine for altitude of 3,000 m and amsl, Temperate for 2,300m to 3,000 m amsl, sub-tropical for 1,500 m to 2,300 m amsl, Tropical for 800 m to 1,500 m amsl and Desert for less than 800 m (The National Atlas of Ethiopia (1981). Akaki River catchment is in the Temperate and sub-tropical climatic zones because its altitude ranges between 3,200m amsl at Intoto Mountain to 2,060m amsl at Aba Samuel reservoir (Ayenew *et al.* 2008). The three seasonal periods in the Akaki River catchment are: period of heavy rain during June to September, dry season during October to February and little rain during March to May (Feven 2007).

3.2.4. Geology

The Akaki River catchment geology is in the Miocene–Pleistocene volcanic successions category. It mainly includes basaltic lava flows, acidic and intermediate lava flows, and pyroclastic flows. These rocks form interlayered sequence with quaternary normal faults. The

principal soil types in the Southern and Northern parts of AA are vertisols and cambisols, respectively (Alemayehu *et al.* 2009).

3.3. Environmental Policies and Regulations

The Constitution of Ethiopia proclaims that all citizens shall have the right to live in a clean and healthy environment (Articles 43,44 and 92). The overall Environmental Policy goal of the Federal Democratic Republic of Ethiopia (FDRE) is to improve the health and quality of life of all Ethiopians and to promote sustainable social and economic development. In line with the Constitution adopted Policies, guideline ambient Environmental Standards for Ethiopia was prepared and adopted (EPA & UNIDO 2003).

3.4. Sampling and Analysis

Samples were collected using 250 ml polypropylene containers. Which were readied, preserved and transported as per standard methods (APHA *et al.* 1998). The sampling positions (**Figure 3-4, Table 3-4**) were identified both to evaluate LAR condition and impact of tributaries on the River. Accordingly, sixteen sampling sites were identified from tributaries before joining LAR; and eleven sites from the river before and after mixing with tributaries. The water samples from these sites were subjected to analysis of fifteen WQ parameters (**Table 3-5**). Measurements were done in situ for DO, T⁰, electrical conductivity (EC) and pH while the remaining eleven parameters were analyzed at the AA City Environmental Protection Laboratory (AA-EPA). In addition, WQ data from point sources and main stem were determined from measurements. Likewise, channel properties were derived from field measurements. These approaches were observed in other papers (Turner *et al.* 2009).

Flow, water T⁰, pH, electrical conductivity (EC), DO, total suspended solids (TSS), total alkalinity as CaCO₃ (alkalinity), orthophosphates as phosphorus (PO₄P), TP, ammonium as N

(NH₄N), NO₃ as N (sum of NO₃N and NO₂N), 5 days biochemical oxygen demand as O₂ (CBOD or BOD) and chemical oxygen demand as O₂ (COD) were measured. Similar parameters were measured for TMDL determination of urban river using Q2Kw (Kannel et al. 2007b).

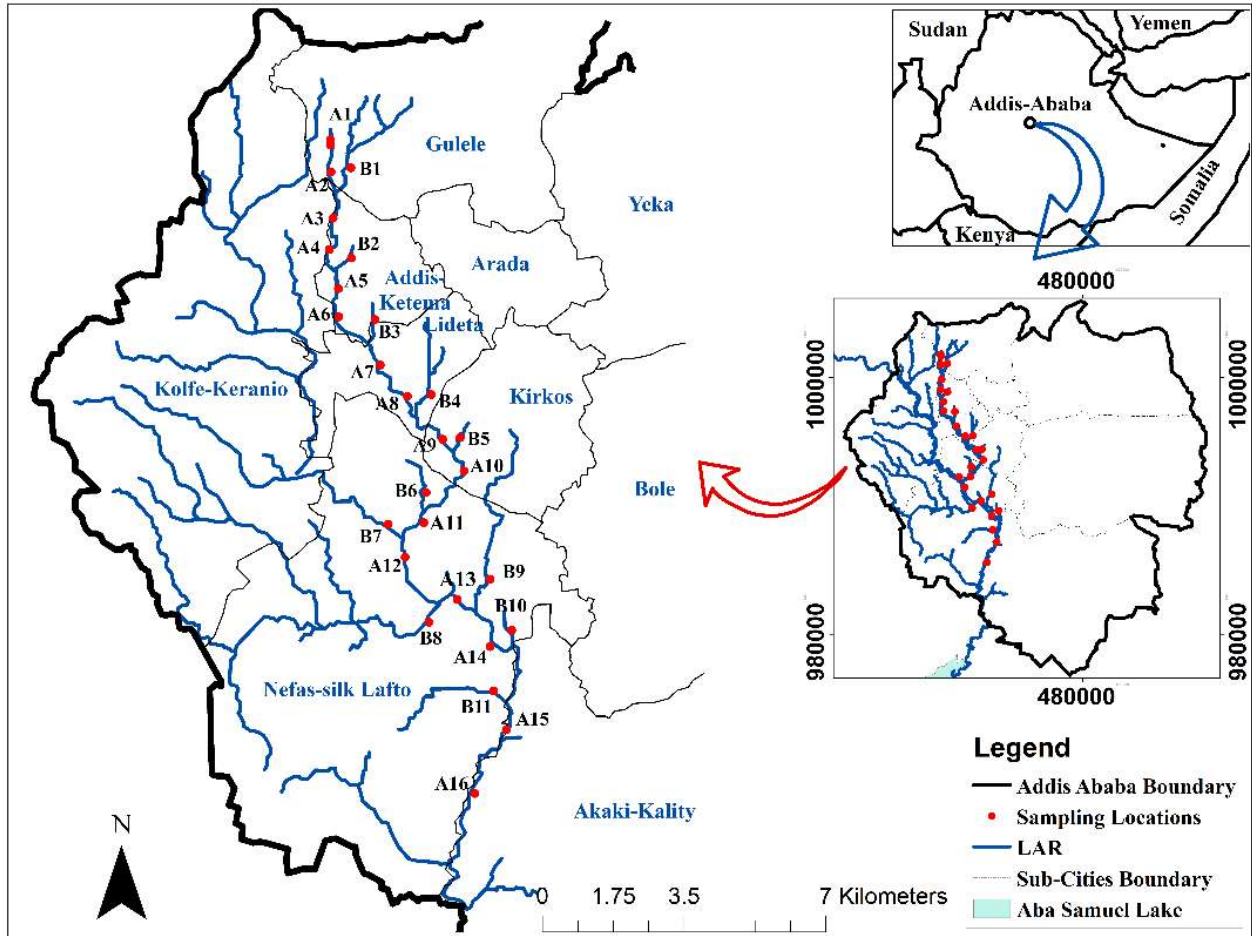


Figure 3-4 Sampling sites from Little Akaki River (LAR)

Table 3-4 Sample collection sites from LAR

No.	Site Code*	Station	Area
1	A1	LAR Upstream	Shegole
2	A2	LAR near Marble Factory	General Wingate Technical College
3	B1	Gulele stream near Finance-Medhanialem	Finance-Medhanialem Roundabout
4	A3	LAR near Sefere-Selam	Kolfe Atena-Tera akababi
5	A4	LAR at Eshete Dildiye	Chimad behind Ethio-Tebib hospital
6	B2	Eshete Dildiye	Chimad behind Ethio-Tebib Hospital
7	A5	LAR near Eyesus Gedam	Amanuel Total akababi
8	A6	Lar near Wndemamachoch	Kefteгна 4
9	B3	Fereja stream near Wendemamachoch	Kefetegn 4
10	A7	LAR - Tolossa-Sefer akababi	Tor-hayloch Mekoninoch Kibeb
11	A8	Lideta Tsebel	Mechare Meda
12	B4	Lideta - Buhe stream	Hostel near Mechare Meda
13	A9	LAR - Behind St. Michael Church	Bulgaria Akababi
14	B5	Mexico Stream near Bulgaria Akababi	Africa Union
15	A10	LAR near Kera	Kera
16	B6	Abo Mazoria Stream	Abo Mazoria
17	A11	LAR at Mekanisa	Mekanisa – National Alcohol
18	B7	Mekanisa- Gefersa Stream	Mekanisa – National Alcohol
19	A12	LAR near German Adebabai	German Adebabai
20	B8	Lebu Stream	Mebrat Mazoria
21	A13	LAR Lafto-Gofa Condominium	Lafto
22	B9	Lafto-Jaja Stream	Lafto
23	A14	LAR Saris	Saris - Dama Hotel
24	B10	Saris- Teramaj Stream	Saris - Dama Hotel
25	B11	Chefe Stream	Hanna Mariam
26	A15	LAR near Batu Leather	Kality Masetegna
27	A16	LAR under New Toll Road Bridge	Mekanisa -Kotari

Table 3-5 Parameters and analytical methods

Parameter	Abbreviations	Units	Analytical Methods
pH	pH	pH unit	pH meter
Total Suspended Solid	TSS	mg/L	Photometric
Total Ammonia as N	T-NH ₃ as N	mg/L	Salcilate
TOC	TOC	mg/L	Oxidation
Nitrate as N	NO ₃ as N	mg/L	Cadmium reduction
Nitrite	NO ₂	mg/L	Diazotization
Total N	T-N	mg/L	Persulfate digestion
Total Phosphorus	T-P	mg/L	Potassium persulphate digestion then molybdate blue
Phosphate	PO ₄	mg/L	Molybdate blue
Biological Oxygen Demand	BOD	mg/L	Gravimetric
Total Alkalinity	T-Alk.	mg/L	Volumetric
DO	DO	mg/L	DO meter
Chemical Oxygen Demand	COD	mg/L	Reactor digestion method
Temperature	T ⁰	°C	Thermometer
Electric Conductivity	EC	mS/cm	Conductivity Meter

3.5. Multivariate Statistical Analysis

In this study, statistical analysis were computed using SPSS Statistics 17.0, JMP 13.0, STATISTICA 8.0 and Excel 2013 software packages. In this context, methods that were followed during cluster analysis, discriminant analysis, PCA and factor analysis are provided below.

3.5.1. Cluster Analysis

Clustering algorithm seeks and forms clusters by relative uniformity within a data (Larose 2014). Since CA tries to find observation vectors that are alike and collect them into clusters, many approaches use an index of similarity or closeness between units of observations. A proper sign of closeness is the distance or gap in the observations. Distance is an indicator of dissimilarity because it increases as two observations go further apart. Euclidean distance is a typical function for a distance found between vectors. The scale of variables measurement is a useful element while applying Euclidean distance. For this, many authors recommend standardization of each variable (Brümelis *et al.* 2000; Rencher 2002). The squared mode of Euclidean distance is given in **Equation 3-1**:

$$Distance (Q_i, Q_j) = \sum_{j=1}^n (X_{1i} - X_{2j})^2 \quad (3-1)$$

where Q_i is the i^{th} object and X_{ij} is the value of the j^{th} variable of the i^{th} object.

Spatial variability was determined using distance of linkage (dlink) for a specific case divided by the largest linkage distance (dmax). Then dlink is standardized by multiplying it with 100 (Alberto *et al.* 2001).

In this study, the most general hierarchical agglomerative CA approach having dendrogram as a visual summary (Alberto *et al.* 2001; McKenna 2003) was applied. For this, Ward's method (Ward 1963) with squared Euclidean distance as an indicator of similarity was used on a data set that was normalized to obtain groups of similar sampling sites for LAR. The approach was commonly applied for WQ evaluation and interpretation in different reports (Singh *et al.* 2005; Bouguerne *et al.* 2017).

3.5.2. Discriminant Analysis

DA was used for discriminating between two or more groups that occur naturally, by applying a set of discriminating variables. These variables measure characteristics on which the groups are expected to differ (Shrestha & Kazama 2007; Close *et al.* 2016). DA applies a linear equation of a regression analysis on raw data with pre-knowledge of membership of objects to particular groups or clusters and provides statistical classification of samples (Alberto *et al.* 2001; Singh *et al.* 2004; Wu & Kuo 2012). DA is expressed in **Equation 3-2**.

$$f(G_i) = k_i + \sum_{j=1}^n w_{ij} P_{ij} \quad (3-2)$$

where k_i is a constant specific to each particular group, i is the count of groups (G), n is count of parameters used in group classification, w_j is specific parameter (P_j) weight coefficient designated by DA.

Independent variables entry options in DA include simultaneous and stepwise (both backward and forward) approaches. Simultaneous entry calculates discriminant function by employing all independent variables at once. This approach is used for limited number of independent variables and in the interest of discovering how well certain variables perform as discriminators in the absence of others. However, stepwise method puts independent variables into the discriminant function (DF) one at a time. This step of entry is based upon variables' relative importance to group separation with greater discriminant weights entered first (Hair *et al.* 1995; Swanson & Holton 2005).

In this study, standard, forward as well as backward stepwise approaches of DA were applied on each matrix of the primary data. In the forward stepwise mode, Discriminant Function

Analysis (DFA) variables were added stepwise until no worthwhile change occurs, whereas, in the backward stepwise mode, variables were removed starting from least significant until significant change occurs. For this purpose, three groups obtained from CA were selected for spatial evaluations. Approach similar to this was followed in other reports (Brūmelis *et al.* 2000; Qadir *et al.* 2008).

3.5.3. Principal Component Analysis

PCA reduces dimensionality of variables in the primary data set by linearly transforming them into a considerably reduced set of uncorrelated variables that represent most of the information in the initial set of variables. These new variables remove potential problems resulted from multi-collinearity. Then, it becomes less complicated to interpret the data set and use it in further analysis than a bigger set of correlated variables (Dunteman 1989; Abdul-Wahab *et al.* 2005).

In this aspect, ahead of PCA/FA, the following assessments are required on the data set to confirm its suitability based on the underlying assumptions. Normality and completeness of data set were checked (O'Shea & Jankowski 2006). The data goodness of-fit to log-normal distribution with 95% or higher confidence is tested with the Kolmogorove-Smirnov (K-S) statistics. Kaisere Meyere Olkin (KMO) was assessed to evaluate sampling adequacy that indicates the degree of common variance probably caused by latent factors. KMO value close to one generally indicates the usefulness of PCA and FA. Likewise, Bartlett's sphericity test signals the lack of relation within variables in the event that the correlation matrix becomes identity matrix. A Bartlett's test having significance level of under 0.05 implies significant relationships within variables (Shrestha & Kazama 2007; Singh *et al.* 2009b; Bilgin & Konanç 2016).

$$C = \begin{bmatrix} 1 & c_{12} & \dots & c_{1p} \\ c_{21} & 1 & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & 1 \end{bmatrix} \quad (3-6)$$

where $C_{ji} = C_{ij}$, is the correlation between x_j and x_i . That is, the PCA is carried out on the matrix of correlation. In that case, summation of eigenvalues in the diagonal is designated as p , which is the number of x variables. Where variance of variable x_i is the diagonal element C_{ii} and covariance of x_i and x_j is the off-diagonal term C_{ij} . Eigenvalues $\lambda_1, \lambda_2 \dots \lambda_p$ represent variances of PCs and related eigenvectors $\mathbf{a}_1, \mathbf{a}_2 \dots \mathbf{a}_p$ denote the coefficients of the i th PC, y_i (**Equation 3-3**). Eigenvalues of the covariance matrix C are obtained using the characteristic equation (**Equation 3-7**):

$$\text{Det}(A - \lambda_i I) = 0 \quad (3-1)$$

where I is the identity matrix with $(P \times P)$ dimension, λ_i is eigenvalues and represent the diagonal values of the diagonal matrix $[D]$. With the assumption that eigenvalues are ordered as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, then λ_i corresponds to the i th PC. At this stage, PCs with eigenvalues less than 1 are disregarded.

In **Equation 3-3**, the loading vector (eigenvector) $a_i = \{a_{i1}, a_{i2}, \dots, a_{ip}\}$ ($i = 1, 2 \dots m$) is calculated by the corresponding eigenvalue *matrix* A using **Equation 3-8**:

$$A x = \lambda x \quad (3-8)$$

There can be up to n of these eigenvalues. **Equation 3-8** can be solved by arbitrarily setting $x_i = 1$, and the resulting vector of x values with transpose $\mathbf{x}^T = (1, x_2, x_3 \dots x_n)$, or any multiple of this vector, is called the i^{th} *eigenvector* of matrix \mathbf{A} .

The first PC is a linear combination with maximal variance; the objective is essentially to search for a dimension along which the observations are widely separated. Furthermore, the second PC is a linear combination having maximal variance in the orthogonal direction (i.e., at ninety degree) to the first PC, and so on (Rencher 2002).

3.5.4. Factor Analysis

Unlike PCA which aims to explain data variance using few uncorrelated components, FA identifies underlying, unobservable, latent factors for variables correlation (Chapman 1992b; Kucuksezgin 1996; Helena *et al.* 2000; Wu & Kuo 2012). FA comes after PCA and minimizes the contribution of less significant variables by varimax rotation of the axis defined by PCA to get new variables also called varifactors (VFs) (Chapman 1992b; Vega *et al.* 1998). This transformation attempts to extract a lower dimensional linear structure from the data set, hence eases the dataset interpretation process (Qadir *et al.* 2008). Different works elaborate description and calculation steps on FA (Afifi & Clark 1997; Manly 2004). FA tries to represent each of the variables as a linear grouping of *common factors* plus a *unique factor* as expressed in **Equation 3-9**:

$$\begin{aligned}
 x_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + e_1 \\
 x_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + e_2 \\
 &\vdots \\
 x_p &= l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + e_p
 \end{aligned}
 \tag{3-9}$$

where $F_1, F_2 \dots F_m$ are common factors, l_{ij} is the loading of the i^{th} variable on the j^{th} common factor, m is the count of common factors (typically this number is considerably smaller

than P), and $e_1, e_2 \dots e_P$ are distinct factors, each linking to one of the original variables. Factor loading l_{ij} signifies the correlation between x_i and F_i .

The factor model in **Equation 3-9** illustrates that each variable is influenced by the underlying common factors. The loadings show how every x_i independently depends on the *common factor*. Consequently, description for the common factor is inferred. In this regard, the factor model breaks the variance of x_i into two parts. Since x_i is standardized, its variance is equal to one and is composed of the following two parts (**Equation 3-10**):

1. *Communality* (h_i^2), i.e., variance in variables caused by common factors;
2. *Specificity* (u_i^2), i.e., variance in variables caused by unique factor e_i

$$\text{Var } x_i = 1 = h_i^2 + u_i^2 \quad (3-10)$$

The numerical aspects of *FA* are concerned with finding estimates of the factor loadings (l_{ij}) and the communalities (h_i^2). The factor loadings (l_{ij}) are obtained from eigenvalues in the varimax rotation. The communality of x_i is expressed in **Equation 3-11**:

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + \dots + l_{im}^2 \quad (3-11)$$

In any *FA*, the count m of common factors is required. Ideally, it should be known prior to the factor analysis. If it is not the case, most investigators use the count of factors as number of eigenvalues exceeding one. The extent of variance defined by the factors is determined as given in **Equation 3-12**:

$$\text{Variance due to Factors} = \frac{l_{1j}^2 + l_{2j}^2 + \dots + l_{pj}^2}{p} \quad (3-12)$$

In this study PCA and FA were applied on a standardized data set using correlation matrices, rather than covariance matrices, to define PCs because the results of analyses are better for direct comparison for different sets of arbitrary variables (Jolliffe 1986; Singh *et al.* 2009b). The variances (eigenvalue or latent roots) and PCs (eigenvectors) of the correlation matrix results were analyzed by examining the major components that are supposed to hold large share of the whole variation. An eigenvalue indicates factor's significance and these two are directly related. Significant are those eigenvalues having 1.0 or more (Kim & Mueller 1978). Also, factor loadings were organized as 'weak', 'moderate' and 'strong', relating to absolute loading values of 0.50–0.30, 0.75–0.50 and >0.75 respectively (Liu *et al.* 2003). In some applications, the PCs are an end in themselves and may be amenable to interpretation (Rencher 2002).

3.6. Water Quality Index Analysis

Public participation is considered as one of the most important components of river management (Von Korff *et al.* 2012). In this circumstance, using WQI, the public will be communicated about pollution problems with simple language that can convey rivers water quality condition without loss important information. This can create awareness and encourage the public to take part in pollution discussions, impact mitigation activities and water quality protection works (Chun 2012).

In this regard, no previous work was found on WQI-ANN analysis of the Little Akakai River. Although the approach used in the thesis is one of the most widely used, it can encourage other researchers to explore alternative indexing approaches in Ethiopia. In this regard, of all the available WQI approaches in the world, Ethiopia (Environmental Agencies) can select one appropriate approach to be consistently used to track water quality conditions. For that matter, a

new indexing approach may be developed based on the purpose and designated use of the different rivers in Ethiopia.

3.6.1. CCME-Water Quality Index

In this study, CCME-WQI was used for WQI determination of LAR due to its capacity to simplify the reporting of complex and technical WQ data. The CCME-WQI provides a mathematical framework for assessing ambient WQ conditions relative to WQ objectives. It is flexible with respect to the type and number of WQ variables, the period of application, and the type of water body (stream, river reach, lake, etc.) tested. Therefore, before the index is calculated, the water body, time period, variables and appropriate objectives need to be defined (CCME 2001).

CCME-WQI has three factors: Scope (Factor 1, F1), Frequency (Factor 2, F2) and Amplitude (Factor 3, F3). These factors are computed to get a single value (between 0 and 100) and to summarize WQ status of water body. For instance, values such as zero and one hundred indicate the worst possible and the best possible index scores respectively. The steps for determination of CCME-WQI are given hereunder (CCME 2001):

F1 (Scope), represents the percentage of variables that do not meet their objectives at least once during the time period under consideration (“failed variables”), relative to the total number of variables measured (**Equation 3-13**):

$$F_1 = \left(\frac{\text{Number of failed variables}}{\text{Total number of variables}} \right) \times 100 \quad (3-23)$$

F2 (Frequency), represents the percentage of individual tests that do not meet the objectives:

$$F_2 = \left(\frac{\text{Number of failed tests}}{\text{Total number of variables}} \right) \times 100 \quad (3-14)$$

F3 (Amplitude), represents the amount by which the failed test values do not meet their objectives. F3 is calculated in three steps:

Step 1: The number of times by which an individual concentration is greater than (or less than, when the objective is a minimum) the objective is termed an “excursion” and is expressed as follows in **Equation 3-15** and **Equation 3-16**.

When the test value must not exceed the objective:

$$\text{excursion}_i = \left(\frac{\text{Failed Test Value}_i}{\text{Objective}_j} \right) - 1 \quad (3-15)$$

For the cases in which the test value must not fall below the objective:

$$\text{excursion}_i = \left(\frac{\text{Objective}_j}{\text{Failed Test Value}_i} \right) - 1 \quad (3-16)$$

Step 2: The collective amount by which the individual tests are out of compliance, referred to as the normalized sum of excursions (*nse*), is expressed in **Equation 3-17**.

$$\text{nse} = \frac{\sum_{i=1}^n \text{excursion}_i}{\text{number of tests}} \quad (3-17)$$

Step 3: *F3* is then calculated by an asymptotic function that scales the normalized sum of the excursions from objectives (*nse*) to yield a value between 0 and 100 (**Equation 3-18**).

$$F_3 = \left(\frac{\text{nse}}{0.01\text{nse} + 0.01} \right) \quad (3-18)$$

The CCME-WQI is finally calculated as expressed in **Equation 3-19**.

$$\text{CCME WQI} = 100 - \left(\frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732} \right) \quad (3-19)$$

Table 3-6 CCME-WQI categories (CCME 2001)

Rank	CCME WQI values	Interpretation
Excellent	(95–100)	WQ is protected with a virtual absence of threat or impairment; conditions are very close to natural or pristine levels
Good	80–94	WQ is protected with only a minor degree of threat or impairment; conditions rarely depart from natural or desirable levels
Fair	60–79	WQ is usually protected, but occasionally threatened or impaired; conditions sometimes depart from natural or desirable levels
Marginal	45–59	WQ is frequently threatened or impaired; conditions often depart from natural or desirable levels
Poor	0–44	WQ is almost always threatened or impaired; conditions usually depart from natural or desirable levels

After the CCME-WQI value was determined for LAR, WQ was ranked by relating it to one of the given categories in **Table 3-6**. In this case, the Ethiopian surface water guideline was used in the WQI determination process.

3.6.2. Artificial Neural Network

In this study, ANN model was used to predict CCME-WQI for LAR. The input signals in the input layer are processed in forward through to the hidden layer. The input signal can be a single signal or an array of signals, whereas the output signal is typically single. To each

connection between two successive layers of neurons, a modifiable weight (W_i) and bias (b_i) are associated in each iteration. A back propagation (BP) neural network with three inputs, one hidden layer and one output layer (**Figure 3-5**) is explained here as an illustration. (Lek *et al.* 1996; Javan *et al.* 2015; Parsaie & Haghiabi 2015; Mohammadpour *et al.* 2016). Assuming the input signal by a vector \mathbf{x} (x_1, x_2, \dots, x_n), the corresponding weights to unit “ j ” by \mathbf{w}_j ($w_{1j}, w_{ij}, \dots, w_{mj}$), and the bias $w_{j0}(=b)$, the net input to the unit “ j ” is given by **Equation 3-20**. (Karunanithi *et al.* 1994; Khuan *et al.* 2002; Kamruzzaman *et al.* 2006; Sakizadeh 2016):

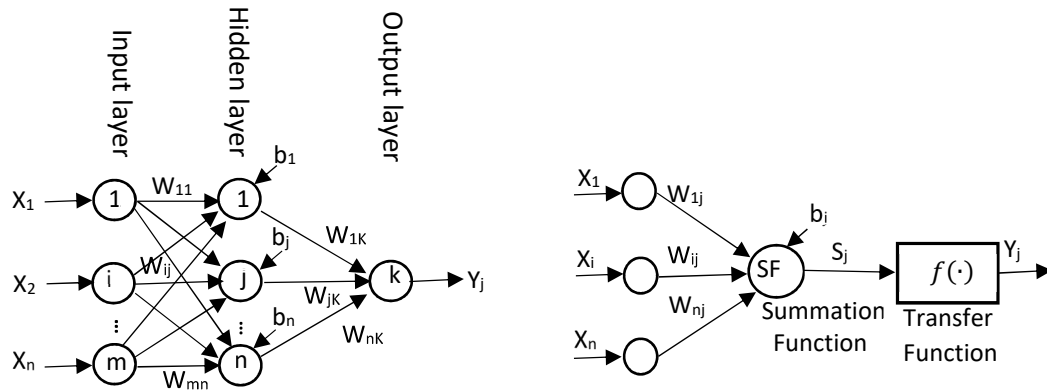


Figure 3-5 Typical BP Network

$$net_j = \sum_n w_{jn}X_n + w_{j0} = w_jX + b \tag{3-20}$$

A sum of weighted inputs are then transformed to an output value by applying an activation function, i.e., $f(\cdot)$. For example, the output of a single-layer perceptron neuron is given in **Equation 3-21**. A simple Perceptron has a set of input neurons and a single output neuron to which all input neurons are connected (Karunanithi *et al.* 1994; Sterratt *et al.* 2011; Magoulès & Zhao 2016).

$$y_i = f_o \left(\sum_i w_i X_i + b \right) \quad (3-21)$$

Also, the explicit expression for an output value of a three layered multi-layer perceptron (MLP) **Error! Reference source not found.** is given in **Equation 3-22** (Nourani *et al.* 2013):

$$y_j = f_o \left[\sum_{j=1}^{M_N} w_{jk} \cdot f_h \left(\sum_{i=1}^{N_N} w_{ij} x_i + w_{jo} \right) + w_{ko} \right] \quad (3-22)$$

where w_{ij} is a weight in the hidden layer connecting the i th neuron in the input layer and the j th neuron in the hidden layer, w_{jo} is the bias for the j th hidden neuron, f_h is the activation function of the hidden neuron, w_{jk} is a weight in the output layer connecting the j th neuron in the hidden layer and the k th neuron in the output layer, w_{ko} is the bias for the k th output neuron, f_o is the activation function for the output neuron, x_i is i th input variable for input layer and y_j and y are computed and observed output variables, respectively. N_N and M_N are the number of the neurons in the input and hidden layers, respectively.

The network is trained using one of the three popular algorithms: error back propagation, cascade correlation and conjugate gradient. Basically the objective of training is to reduce the global error, E , defined below (Deo & Thirumalaiah 2000).

$$E = \frac{1}{P} \sum_{P=1}^P E_P \quad (3-23)$$

where P is total number of training patterns and E_p is error for training pattern, p , which is given by:

$$E_P = \frac{1}{2} \sum_{k=0}^N (o_k - t_k)^2 \quad (3-24)$$

where, N is total number of output nodes, O_k is network output at the k^{th} output node and t_k is target output at the k^{th} output node. In this context, the goal of every training algorithm is to reduce this global error by adjusting the weights and biases (Deo & Thirumalaiah 2000; Kamruzzaman *et al.* 2006).

The three major steps in ANN configuration, in standard commercial software, are neural network type choice, design choice, training and testing of the network (Khuan *et al.* 2002; Georgilakis 2009). Commonly ANN modelling follows the following steps: database collection, analysis and preprocessing of the data, and training of the neural network. The latter includes the choice of architecture, training functions, training algorithms and parameters of the network. The final step is testing of the trained network and using the trained ANN model for simulation and prediction. The relative mean squared error (RMSE), coefficient of determination (R^2) and mean squared error (MSE) were used for testing purpose (Dawson & Wilby 2001; Dogan *et al.* 2009; Heddam 2016; Wagh *et al.* 2016). The ANN model framework for CCME-WQI is given in **Figure 3-6**.

The choice of network configuration depends on the scale and type of problem. For instance, nonlinear computation requires nonlinear network. Use of a nonlinear transfer function makes a network capable of storing nonlinear relationships between input and output. The sigmoid (logistic) and hyperbolic tangent are most commonly used transfer functions. However, instead of using these functions in the output layer one can use a linear function. The linear output layer lets the network produce values outside the range -1 to $+1$. In addition, except for purely linear networks, the more neurons in a hidden layer, the more powerful the network becomes (Demuth & Beale 1993; Karunanithi *et al.* 1994; Govindaraju & Rao 2013).

In this study, input data was standardized to accommodate the difference in variables unit and to minimize impact on ANN algorithm (Dawson & Wilby 2001; Singh *et al.* 2009a). To this end, the data were transformed using input data (X_{in}), the minimum data (X_{min}), and the maximum data (X_{max}) as shown in the following relationship (Zhang *et al.* 1998; Barzegar &

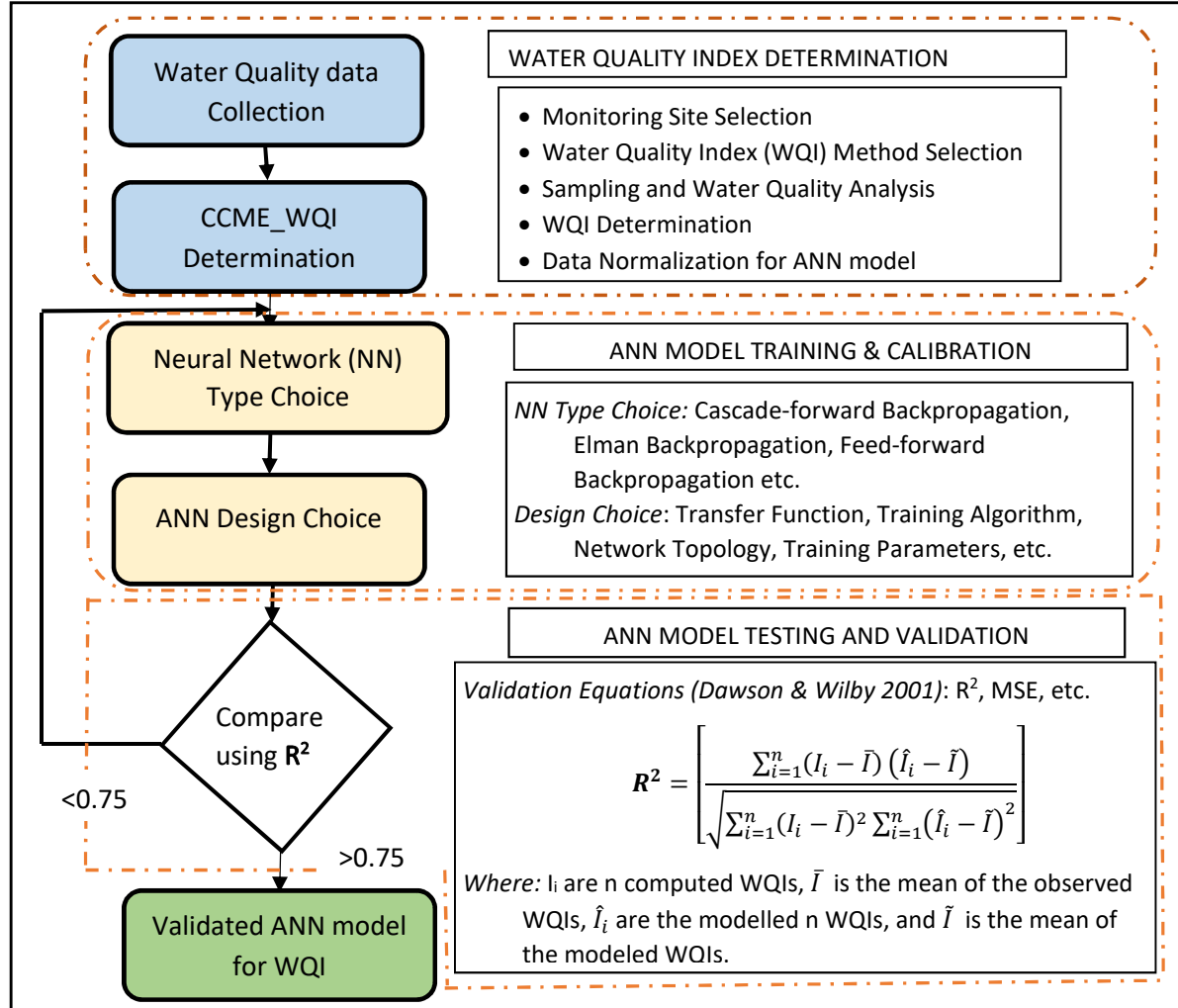


Figure 3-6 ANN Model Framework for CCME-WQI

$$Z_{stn} = \frac{X_{in} - X_{min}}{X_{max} - X_{min}} \quad (3-25)$$

In addition, for CCME-WQI prediction, the commonly applied feed forward (FF) neural network with back propagation (BP) training algorithm was used (Zhang *et al.* 1998; Brion & Lingireddy 2000; Ehteshami *et al.* 2016). For twelve inputs and one output, prediction was conducted using several combinations of hidden layers (2-20), number of neurons in the hidden layer (5, 10, 15, 20, 25), transfer function (logsig-non linear, linear), training function (trainlm)

and learning function (learnngdm, learnngd). Afterwards, architectures with best correlation coefficient (R) were retained. Suggestion on approaches were observed in different reports (Zhang 1998; Juahir 2004 and Singh 2009). All computations were performed using the EXCEL 2016 and MATLAB.

3.7. Applications of Q2Kw

3.7.1. River Discretization

For the purpose of this study, a data collection (i.e., for discharge and WQ variables) network was designed to capture the variability of the headwater (also called the upstream boundary condition), point sources (i.e., tributary inflows) in the sixteen tributaries and eleven sampled sites on the river LAR. Similar approach was proposed in other papers (Pelletier & Chapra 2008a; Hobson *et al.* 2014).

To implement the model, the studied 24 km length of LAR was discretized into ten reaches with different reach lengths. **Figure 3-7** shows the river system segmentation along with the locations of pollution loads and abstraction. The river segmentation is based on making divisions at points of major changes, such as confluence with major tributaries. Furthermore, each of these reaches was assumed to have similar hydraulic characteristics such as slope of reaches, width, manning roughness coefficient etc. The WQ at uppermost Gulele locality was considered as upstream boundary. The downstream boundary was not prescribed considering absence of effects in modeling. Similar approach was used in different reports (Kannel *et al.* 2007b; Gikas 2014; Sharma *et al.* 2015a).

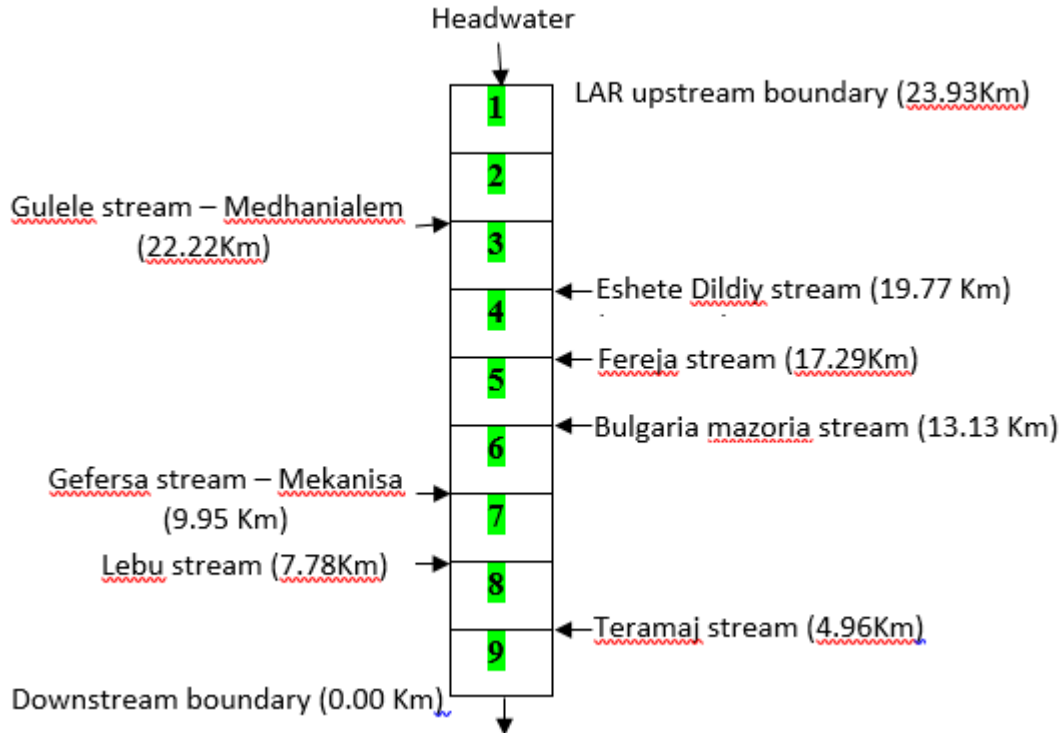


Figure 3-7 Little Akaki River Segmentation and Pollution Sources

3.7.2. Data Entry

For simulation purpose, in addition to WQ data, Q2Kw requires site-specific information that is important to characterize the stream and its surroundings. This includes geometric properties of the river (such as channel slope, channel width, side slope, and manning roughness coefficient), hydraulic properties (such as travel time, stream flow, velocity), pollutant loads, and meteorological parameters (such as air T^0 , dew point T^0 , wind speed, cloud cover, solar radiation) (Hobson *et al.* 2014; Sharma *et al.* 2015a). In addition, for data entry purpose, Q2Kw provides convenient numerous Microsoft Excel sheets (**Figure 3-8**).

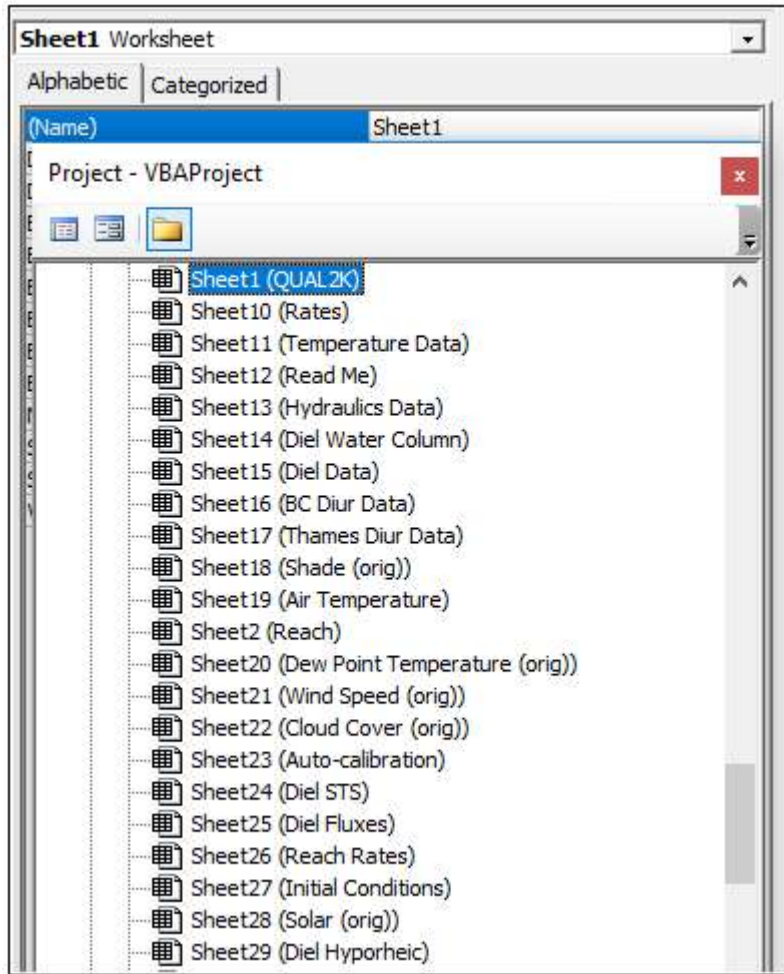


Figure 3-8 Partial list of Q2Kw data entry sheets

- **Geometric Properties**

In this study, Manning’s roughness coefficient was adjusted to the following characteristics: depth and mean velocity as a function of the stream’s bottom declivity, width, and slopes, with each reach idealized as a trapezoidal channel (**Figure 3-9**).

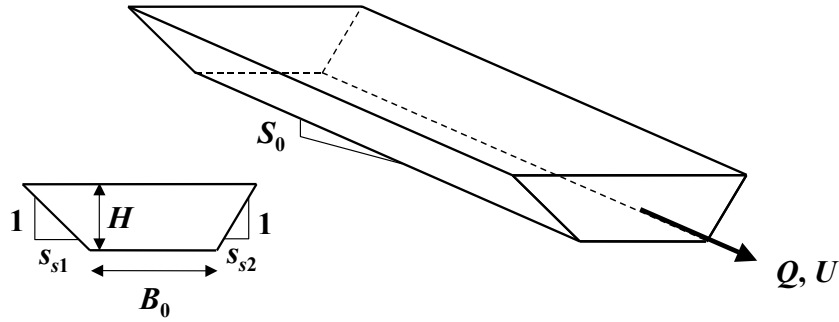


Figure 3-9 Trapezoidal channel

For steady flow, the Manning and other relevant equations are provided in **Equations 3-26 to Equations 3-29** (Pelletier & Chapra 2008a):

$$Q = \frac{S_0^{1/2} A_c^{5/3}}{n P^{2/3}} \quad (3-26)$$

where Q = flow [m³/s], S_0 = bottom slope [m/m], n = the Manning roughness coefficient, A_c = the cross-sectional area [m²], and P = the wetted perimeter [m].

The cross-sectional area of a trapezoidal channel is computed as:

$$A_c = [B_0 + 0.5(S_{s1} + S_{s2})H]H \quad (3-27)$$

where B_0 = bottom width [m], s_{s1} and s_{s2} = the two side slopes [m/m], and H = reach depth [m].

The wetted perimeter is computed as:

$$P = B_0 + H\sqrt{S_{s1}^2 + 1} + H\sqrt{S_{s2}^2 + 1} \quad (3-28)$$

After substituting **Equations 3-27** and **3-28**, **Equation 3-26** can be solved iteratively for depth:

$$H_k = \frac{(Qn)^{3/5} \left(B_0 + H_{k-1}\sqrt{S_{s1}^2 + 1} + H_{k-1}\sqrt{S_{s2}^2 + 1} \right)^{2/5}}{S^{3/10} [B_0 + 0.5(S_{s1} + S_{s2})H_{k-1}]} \quad (3-29)$$

where $k = 1, 2, \dots, n$, $n =$ the number of iterations. Here, an initial guess of $H_0 = 0$ is employed and the method is terminated when the estimated error falls below a specified value of 0.001%. The estimated error is calculated as:

$$\varepsilon_a = \left| \frac{H_{k+1} - H_k}{H_{k+1}} \right| \times 100\% \quad (3-30)$$

The average reach width, B [m], can be computed as:

$$B = \frac{A_c}{H} \quad (3-31)$$

- **Hydraulic Properties**

Flow was measured using float method determining cross-sections and measuring velocities at three to five points along river cross-sections. Similar approach was observed in other reports (Kannel *et al.* 2007a).

The cross-sectional area can be determined with **Equation 3-27** and the velocity can then be determined from the continuity equation:

$$U = \frac{Q}{A_c} \quad (3-32)$$

Residence time of any given reach k is determined as:

$$\tau_k = \frac{V_k}{Q_k} \quad (3-33)$$

where $\tau_k =$ the residence time of the k^{th} reach [d], $V_k =$ the volume of the k^{th} reach [m^3] = $A_{c,k}\Delta x_k$, and $\Delta x_k =$ the length of the k^{th} reach [m].

A river reach forms one part of a river system. Each reach is also divided into a number of computational elements or sub-reaches. Q2Kw implements steady-state flow (i.e., $dQ/dt = 0$) balance for each model reach. Therefore the balance for a computational element (**Figure 3-10**) can be written as:

$$Q_i = Q_{i-1} + Q_{in,i} - Q_{ab,i} \quad (3-34)$$

where Q_i = outflow from reach i into the downstream reach $i + 1$ [m^3/s], Q_{i-1} = inflow from the upstream reach $i - 1$ [m^3/s], $Q_{in,i}$ is the total inflow into the reach from point and nonpoint sources [m^3/s], and $Q_{ab,i}$ is the total outflow from the reach due to point and nonpoint abstractions [m^3/s].

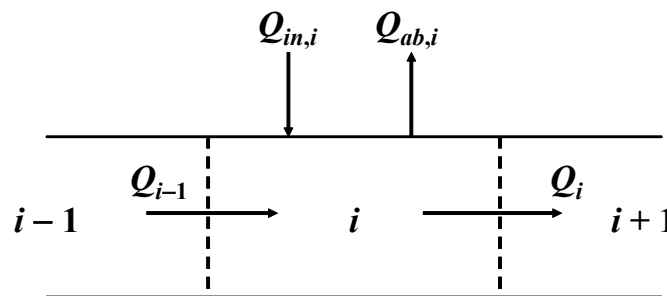


Figure 3-10 Reach flow balance

- **Meteorological Parameters**

For the study area, air T^0 , relative humidity, wind speed, sun hours and precipitation data were obtained from the National Meteorology Agency of Ethiopia (**APPENDIX I**).

- **Model Implementation Data**

In model implementation stage, time step for the calculations was set at 5.625 minutes to avoid instability in the model. For integration and pH modeling purposes, Euler's method and Newton-Raphson method were applied, respectively. Euler's method is the default setting

because it commonly attains sufficiently accurate results at a moderate computational time (Pelletier & Chapra 2008b).

Similar approaches were observed in other reports (Camargo *et al.* 2010; Hossain *et al.* 2014; Sharma *et al.* 2015a).

To simulate exchanges in the hyporheic zone, level I was chosen, because it included zero-order and first-order oxidation of the fast-reacting dissolved components of BOD. The sediment/hyporheic zone thickness, sediment porosity and hyporheic exchange flow were assumed to be 10 cm, 40%, and 5%, respectively. The algae cover and bottom-sediment oxygen demand were both assumed to be 50 and 100% respectively. Moreover, to obtain the best adjustment, the modeling system assigns standard weights to various parameters in order to minimize the error between the observed and simulated data. In this regard, a weight of fifty for DO, because it is the most influential parameter, versus weights of five for BOD₅, TN and TP were used. Similar approach was observed in other reports (Camargo *et al.* 2010; Kannel *et al.* 2011; Hossain *et al.* 2014; Sharma *et al.* 2015a).

The number of model runs over which to perform the optimization (i.e., based on GA) of the parameter set was selected for model runs equal to number populations times number of generations. The population, $np = 200$ with generations, $ngen = 100$ has been used for GA resulting in the generation of a total of 20,000 different values and out of which the value which gives maximum fitness function has been chosen for the period and kept constant for the validation purpose. The model was run until the system parameters were appropriately adjusted and reasonable agreement between model results and field measurements were achieved. Similar

approach was observed in different works (Pelletier & Chapra 2005; Kannel *et al.* 2007a; Camargo *et al.* 2010; Hobson *et al.* 2014; Sharma *et al.* 2015a).

3.7.3. Calibration and Validation

- **Model Calibration**

Q2Kw was auto-calibrated and validated for DO, BOD, TN, TP and eleven sampled sites on LAR. These parameters were commonly used for model calibration and validation purposes in different reports (Kannel *et al.* 2007a; Sharma *et al.* 2015a)

For model calibration, the usual practice is to establish that certain constituents are predicted correctly before moving onto the more interconnected mechanisms associated with nutrient cycling. In this study also discharge was calibrated first, using Manning’s roughness coefficient *n* as calibration variable that was varied within suggested limits for natural streams; the final values ranged between 0.02 and 0.04 (**Table 3-7**). Any remaining difference in discharge between observed and predicted could be attributed to sources such as nonpoint sources, inaccessible sources etc. Similar approach was suggested and observed in different reports (Hobson, 2014; Gikas, 2014; Pelletier, 2008).

Table 3-7 River Reach based Manning’s roughness, *n* values of the LAR

Reaches	Downstream location (km)		Manning, <i>n</i>
	1	2	
1	23.93	23.30	0.030
2	23.30	22.22	0.040

3	22.22	19.77	0.020
4	19.77	17.29	0.032
5	17.29	13.13	0.038
6	13.13	9.95	0.033
7	9.95	7.78	0.033
8	7.78	4.96	0.040
9	4.96	0.00	0.039

Furthermore, the numerical integration of the model is compiled as a Fortran Program for fast performance. In this regard, for automatic calibration, the PIKAIA-GA orders the compiled program to run each time the fitness function is evaluated for maximization purpose. Consequently, the VBA accounts for only a small fraction of the computational time when the GA is run. The automatic calibration in Q2Kw calculates the goodness of fit by comparing the simulated result with measured data based on fitness value of root mean square error (RMSE) inverse indicated in **Equation 3-35** (Pelletier *et al.* 2006; Pelletier & Chapra 2008b).

$$f(x) = \left[\sum_{i=1}^q w_i \right] \left[\sum_{i=1}^q \frac{1}{w_i} \left[\frac{\frac{\sum_{j=1}^m O_{i,j}}{m}}{\left[\frac{\sum_{j=1}^m (P_{i,j} - O_{i,j})^2}{m} \right]^{1/2}} \right] \right] \quad (3-35)$$

Where $O_{i,j}$ = observed value, $P_{i,j}$ = predicted value, m = number of pairs of predicted and observed values, w_i = weighting factor, and q = number of different state variables included in the reciprocal of the weighting normalized RMSE.

In this study, the default kinetic parameters in the “Rate” sheet were taken for the auto-calibration purpose (Pelletier & Chapra 2008b).

- **Model Verification**

To test the model's ability to predict the WQ conditions during different hydrological periods, Q2Kw was run for the wet period, without changing the parameters that were calibrated for the dry period (Camargo *et al.* 2010). In this regard, the verification was used to compare different sets of field data and meteorological data, while for the input variables of the model (e.g., rates of nitrification, denitrification, etc.) obtained from the calibration of the model were maintained (Gikas 2014).

3.7.4. Uncertainty and Sensitivity Analysis

In this study, a simulation Add-In (YASAIw) was integrated to Q2Kw to convert the existing deterministic Q2Kw model into a probabilistic model for sensitivity analysis. Eckstein and Riedmueller (2002) originally developed YASAIw as a generalized Microsoft Excel Add-In. It was then enhanced and adapted to suit Q2K by the State of Washington (Pelletier 2009). Consequently, the YASAIw allows for any number of Q2Kw inputs to be converted to stochastic parameters defined by probability distributions. The software uses Monte Carlo sampling to sample from the input distributions for a prescribed number of model iterations and provides probabilistic output for key parameters (Sharma *et al.* 2015a).

In this study, the dry season data was used to understand the effect of input variables on DO, BOD, TN, and TP. For this purpose, rate parameters and flowrate were varied for sensitivity analysis. This analysis was conducted by running the Q2Kw model in a loop that repeats a specified number of times. Each time the model run is repeated, the program generates a new set of randomly varied input variables. The program records the input values and output values at the end of each run and then repeats the process (Sharma *et al.* 2015a).

3.8. Simulation-Optimization based WLA Modeling

3.8.1. Q2Kw Codes

The Q2Kw model uses open Excel VBA code organized in modular form (**Figure 3-11**). The availability of the VBA code of Q2Kw allows to design user specified WQ simulations and optimizations (Pelletier *et al.* 2006). As a result, the PIKAIA-GA code in the *Autocal_algorithm* module was modified for SO based WLA modeling purpose.

As specified in (Pelletier *et al.* 2006), the PIKAIA-GA maximizes the function $f(x)$ in a bounded n-Dspace for:

$$x = (x_1, x_2, \dots, x_n) \quad x_k \in [0.0, 1.0] \quad (3-36)$$

Where n is the number of parameters that are being optimized. The parameters (x) are bounded in the range of 0.0 to 1.0 in the GA. These optimization parameters for the model are scaled from the values of x according to a linear interpolation between the specified minimum and maximum value of each kinetic rate parameter that is being optimized. The fitness of a particular set of optimization parameters is represented by the maximization of the value of the function $f(x)$ (Pelletier *et al.* 2006).

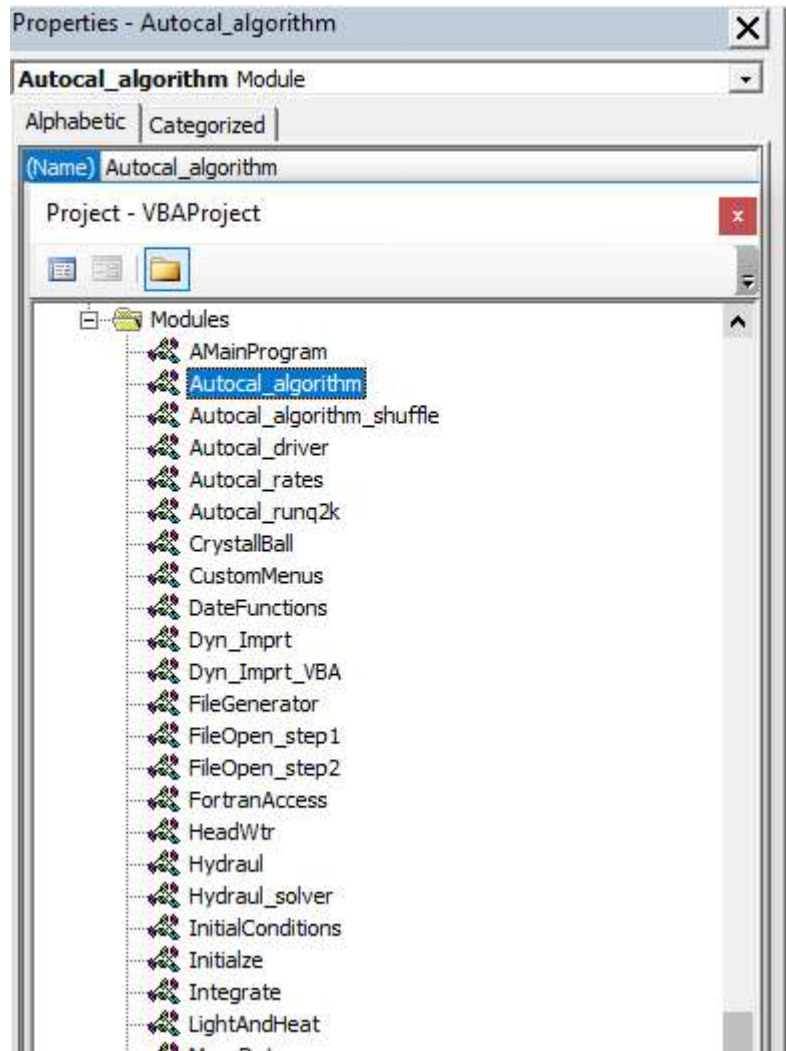


Figure 3-11 A partial view of modular VBA code organization of Q2Kw model

3.8.2. Genetic Algorithm Codes

In this study, PIKAIA-GA was used for the MO-WLA optimization problem. The algorithmic structure of PIKAIA-GA in pseudo-code form is shown below (Metcalf & Charbonneau 2003).

for $j := 1, N_G$ do	GENERATION LOOP	M
for $k := 1, N_P/2$ do	BREEDING	M
$k_1 := \text{SELECT}(\text{fit})$	<i>select first parent</i>	M
$k_2 := \text{SELECT}(\text{fit})$	<i>select second parent</i>	M
$g_1 := \text{ENCODE}(P_{\text{old}}(k_1, 1:n))$	<i>create first parent string</i>	M
$g_2 := \text{ENCODE}(P_{\text{old}}(k_2, 1:n))$	<i>create second parent string</i>	M
call CROSS (g_1, g_2)	<i>crossover operator</i>	M
call MUTATE (g_1)	<i>mutation operator</i>	M
call MUTATE (g_2)	<i>mutation operator</i>	M
$P_{\text{new}}(2k-1, 1:n) := \text{DECODE}(g_1)$	<i>decode and store</i>	M
$P_{\text{new}}(2k, 1:n) := \text{DECODE}(g_2)$	<i>offspring strings</i>	M
endfor		
for $k := 1, N_P$ do	FITNESS EVALUATION	S
$\text{fit}(k) := \text{FF}(P_{\text{new}}(k, 1:n))$	<i>user-supplied fitness function</i>	S
endfor		
for $k := 1, N_P$ do	REPRODUCTION PLAN	M
$P_{\text{old}}(k, 1:n) := P_{\text{new}}(k, 1:n)$	<i>full generational replacement</i>	M
endfor		
call ADJMUT (fit)	<i>adjust mutation rate</i>	M
endfor		

Referring to the above PIKAIA-GA pseudo-code, the number of model parameters being optimized is n , N_P is the (fixed) population size, and N_G is the number of (time-like) generations over which the evolution is executed. The tasks ENCODE and DECODE transform an n -D floating-point array to a string, and vice versa. The function SELECT picks a single individual (flagged by an integer, here k_1 or k_2), with a probability proportional to its fitness-based rank in

the current population. Selection is done using a rank-based version of the classical Roulette Wheel Algorithm. It is to be noted that the strings g_1 , g_2 are modified upon exit from CROSS and MUTATE, and since each breeding event produces two offsprings, the first inner loop only needs to repeat $N_P/2$ times to produce a new, full-size population. Function and operations calls designated as “M” are executed serially by the Master program, and those labeled “S” are carried out in parallel by the Slave processes. (Davis 1991; Metcalfe & Charbonneau 2003).

The GA algorithm seeks to find a population having a maximum value of the user-supplied fitness function (FF), which takes as input an n-D floating-point array $x(1 : n)$ containing a set of parameter values defining one occurrence of the model being optimized, and outputting a measure of goodness-of-fit. The code evolves a population of N_P trial points in the n-D search space, stored in the array $P_old(1:N_P, 1 : n)$, through a preset number of generations N_G . The population is usually initialized with random changes uniformly distributed in user-specified intervals, so that the evolutionary search remains bounded but otherwise entirely unbiased by the choice of initial conditions (Metcalfe & Charbonneau 2003).

3.8.3. Q2Kw-PIKAIA Integration for Optimization

Integrated Q2Kw – GA model is important in finding trade-off solutions for a decision maker faced with *WLA* problem among various dischargers (Allam *et al.* 2016). The proposed optimal WLA framework is shown below in **Figure 3-12**. It consists of the MO optimization model, with the WQ simulation model (Q2Kw) embedded in the optimization model (GA) (Yandamuri *et al.* 2006). In each output of new solution, the input data for Q2Kw model was kept constant and multiplied with random WL reduction factors supplied from the GA (Mostafavi & Afshar 2011). In this study, the Pseudo-code algorithm of Q2Kw-PIKAIA-GA is provided in **APPENDIX III**. In addition, data-tables that are relevant to the pseudo-code are given in **APPENDIX IV**.

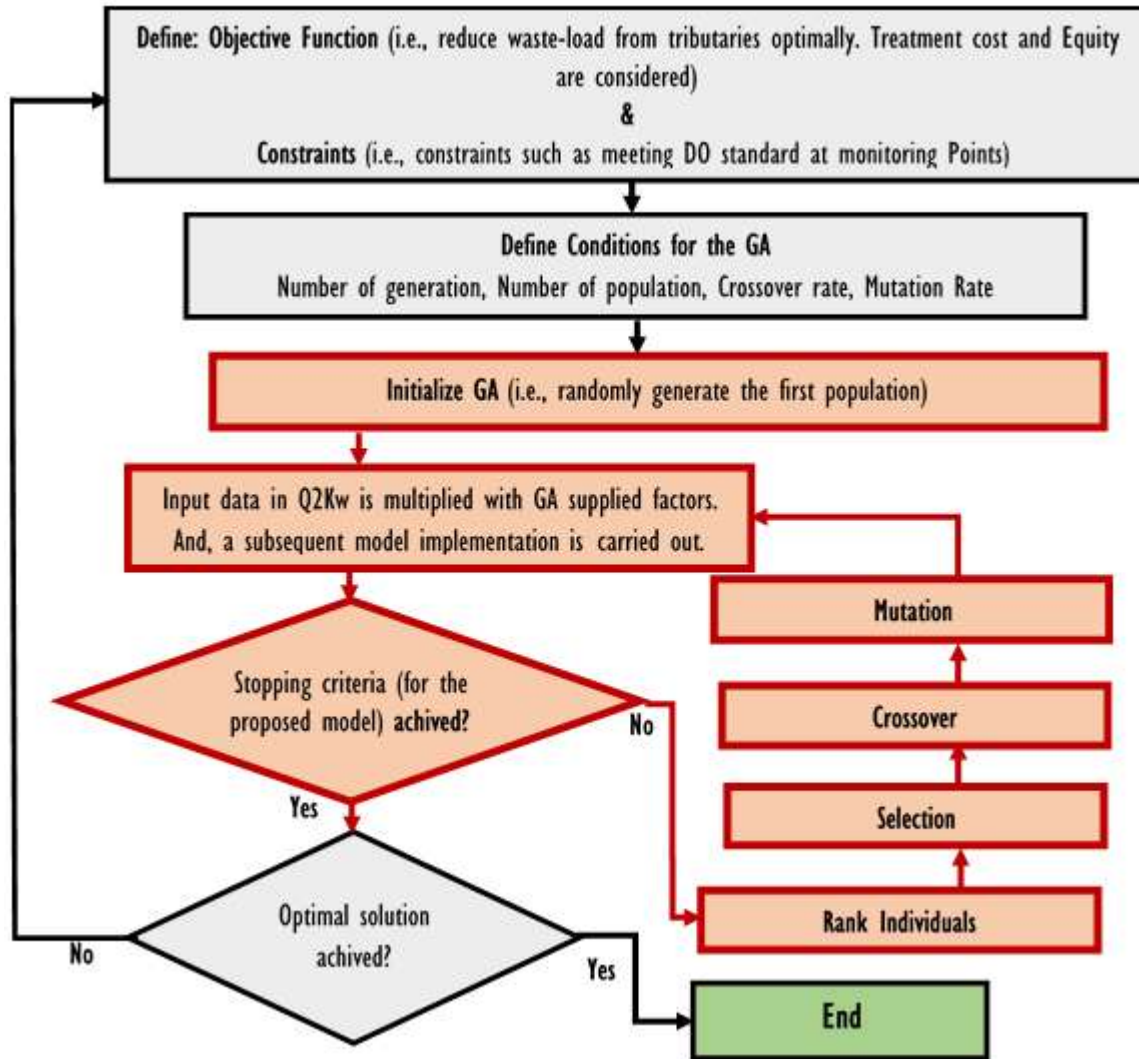


Figure 3-12 Optimal MO-WLA modeling framework

3.8.4. Proposed Model

The proposed optimal WLA model was formulated to achieve minimum *WL* abatement, minimum inequity measure and maximum performance measure (Tung & Hathorn 1989). The equations for the WLA formulations are described as follows. **Equation** Error! Reference source not found. is the objective function that defines the treatment cost at each source expressed by waste-removal efficiencies (Burn & Yulianti 2001; de Andrade *et al.* 2013). **Equation (3-37)** is

objective function that defines the equity measure among all the pollution sources measure (Burn & Yulianti 2001).

The proposed approach for the overall performance measure is shown in **Equations** from **3-39** to **3-40**. This is similar approach to Yandamuri *et al.* (2006) . In this regard, the three individual performance measures are expressed in terms of number of violations, total magnitude of violations, and magnitude of maximum violation. In addition, the individual performances are expressed as the ratio of the “difference between individual performance during no treatment and that corresponding to actual treatment” to “corresponding to no treatment.” **Equation** Error! Reference source not found. is a constraint equation that defines performance measure in terms of the violations of local DO and BOD standards (Burn & Yulianti 2001; Yandamuri *et al.* 2006; de Andrade *et al.* 2013). WQ monitoring parameters of TP and TN (Wang *et al.* 2008) were incorporated in **Equation**Error! Reference source not found. also as constraints.

Minimize

$$w_1 \sum_{i=1}^n x_i \quad (3-37)$$

Minimize

$$+w_2 \sum_{i=1}^n \left| \frac{x_i}{\bar{x}} - \frac{w_i}{\bar{w}} \right| \quad (3-38)$$

$$+ w_3 \left(\frac{N_o - N_a}{N_o} + \frac{V_o - V_a}{V_o} + \frac{TM_o - TM_0}{TM_o} \right)_{DO} + w_4 \left(\frac{N_o - N_a}{N_o} + \frac{V_o - V_a}{V_o} + \frac{TM_o - TM_0}{TM_o} \right)_{BOD} \quad (3-39)$$

$$+ w_5 \left(\frac{N_o - N_a}{N_o} + \frac{V_o - V_a}{V_o} + \frac{TM_o - TM_0}{TM_o} \right)_{TP} + w_6 \left(\frac{N_o - N_a}{N_o} + \frac{V_o - V_a}{V_o} + \frac{TM_o - TM_0}{TM_o} \right)_{TN} \quad (3-40)$$

Subject to

$$N_0 = f_1 \left((Q_j)_0, Q_{std} \right) \quad \text{and} \quad N_a = f_2 \left((Q_j)_a, Q_{std} \right) \quad (3-41)$$

Where

$$N_a = \sum_{j=1}^{NC} (y_j)_a \quad (3-42)$$

In which

$$N_a = \begin{cases} 1 & \text{if } Q_{std} > (Q_j)_a \\ 0 & \text{if } Q_{std} \leq (Q_j)_a \end{cases} \quad (3-43)$$

$$V_0 = f_1 \left((Q_j)_0, Q_{std} \right) \quad \text{and} \quad V_a = f_2 \left((Q_j)_a, Q_{std} \right) \quad (3-44)$$

Where

$$V_a = \left[(S_1)_a, (S_2)_a, \dots, (S_j)_a \right] \quad (3-45)$$

In which

$$(S_j)_a = \begin{cases} (Q_{std} - (Q_j)_a) & \text{if } Q_{std} > (Q_j)_a \\ 0 & \text{if } Q_{std} \leq (Q_j)_a \end{cases} \quad (3-46)$$

$$TM_0 = f_5 \left((Q_j)_0, Q_{std} \right) \quad \text{and} \quad TM_a = f_6 \left((Q_j)_a, Q_{std} \right) \quad (3-47)$$

$$TM_0 = f_5 \left((Q_j)_0, Q_{std} \right) \quad (3-48)$$

$$TM_a = f_6 \left((Q_j)_a, Q_{std} \right) \quad (3-49)$$

Where

$$TM_a = \sum_{j=1}^{NC} (S_j)_a \quad (3-50)$$

where in **Equations 3-37 to 3-38**, x_i = waste-removal level for the waste source i , W_i = mass flow for the waste source i . In **Equations 3-39 to 3-40**, N_0 = number of DO/BOD/TN/TP standard violations corresponding to no treatment, N_a = number of DO/BOD/TN/TP standard violations corresponding to actual treatment. V_0 = magnitude of maximum DO/BOD/TN/TP standard violations corresponding to no treatment, V_a = magnitude of maximum DO/BOD/TN/TP standard violations corresponding to actual treatment. TM_0 = total magnitude of maximum DO/BOD/TN/TP standard violations corresponding to no treatment, TM_a = total magnitude of maximum DO/BOD/TN/TP standard violations corresponding to actual treatment. In Eqs. (19-28), N_C = number of checkpoints. The index $(y_j)_a$ keeps the count of DO/BOD/TN/TP standard violation at a checkpoint j (zero-one) integer variable. Q_j = DO/BOD/TN/TP concentration at the checkpoint j , and subscripts 0 and a indicate the level of treatment corresponding to no treatment and actual treatment respectively. Q_{std} = specified DO/BOD/TN/TP standard for the river.

The weighted sum of **Equations 3-37 to 3-40** (de Andrade *et al.* 2013; Allam *et al.* 2016) is used as fitness function in this Q2Kw-GA MO optimal WLA problem. The WLA results presented here for LAR are based on ten reach divisions, 100 - 0 % pollutant removal levels, 100 populations and 50 generations. Other major conditions of this analysis are roulette wheel as a selection mechanism and mutation probability (pm) ranging between 0.005 and 0.25.

4. RESULTS AND DISCUSSION

4.1. Descriptive and Multivariate Analysis

4.1.1. Descriptive statistics

For the purpose of WQ assessment of the LAR, as suggested in other reports (Voutsas *et al.* 2001), the raw WQ data (wet season and dry season) was averaged and analyzed using descriptive statistics (mean, minimum, maximum, standard deviation) to get summary of the WQ condition of LAR (**Table 4-1**). The mean values of TSS, T-N, BOD, and DO far exceed the standard stipulated in the Ethiopian guideline for surface waters (EPA & UNIDO 2003). High TSS values reduce light penetration in water and thus limits photosynthesis and respiration in aquatic organisms (Barankiewicz *et al.* 2014). High BOD indicates organic pollution that results in river oxygen depletion (Inoue & Ebise 1991; Kumar *et al.* 2010). In addition, DO concentrations below 5 mg/L may adversely affect survival of biological communities and below 2 mg/L may lead to the death of most fish. The mean values of TOC (greater than 10mg/L) and COD (greater than 100mg/L) imply discharge of municipal and/or industrial wastes (Chapman 1992a).

Moreover, the spatial WQ variation is evident from the high standard deviation values. Similar result was observed in another report as well (Vega *et al.* 1998). For instance, a maximum 1,006 mg/L TSS value of Fereja stream (sampling site 9 in **Figure 4-1**) belongs to the most densely populated Addis Ketema sub-city (Error! Reference source not found.). Likewise, T-N had a maximum value of 196 mg/L in the third densely populated Kirkos sub-city. Unlike other parameters, of all sampled sites, none met criteria for BOD. Again, the degree of violation was greatest in the most densely populated sub-city. In addition, except few sampled sites in the

upstream, DO values fell short of 4mg/L in the sampled sites. Raw WQ data were subjected to descriptive statistics: range, standard deviation.

Table 4-1: Basic Statistical analysis of WQ data of LAR

Parameter	Mean	Minimum	Maximum	Standard Deviation	Local Standard* ¹
TSS, mg/L	264.4	12.5	1006.0	205.2	50
NO ₃ as N, mg/L	10.4	0.0	48.9	14.2	50
NO ₂ , mg/L	0.1	0.0	1.1	0.2	0.1
T-N, mg/L	65.4	0.1	196.0	46.1	1
T-Alk, mg/l.	484.0	138.9	835.5	156.0	Not stated
TOC, mg/l.	37.8	3.2	78.2	21.6	Not stated
COD, mg/l.	503.5	110.0	1026.0	318.6	Not stated
BOD, mg/l.	192.7	57.5	476.0	106.8	≤ 5
DO, mg/l.	1.5	0.0	6.2	2.0	> 4
T ⁰ , °C	16.2	9.0	19.3	2.7	Should not result in >3 ⁰ C variation
EC, mS/cm	0.8	0.3	3.0	0.5	1 at 20 ⁰ C
pH	8.0	7.1	8.5	0.3	6-9

*¹ EPA and UNIDO, 2003. (EPA & UNIDO 2003) *² un-impacted concentration at upstream was 1.8mg/L

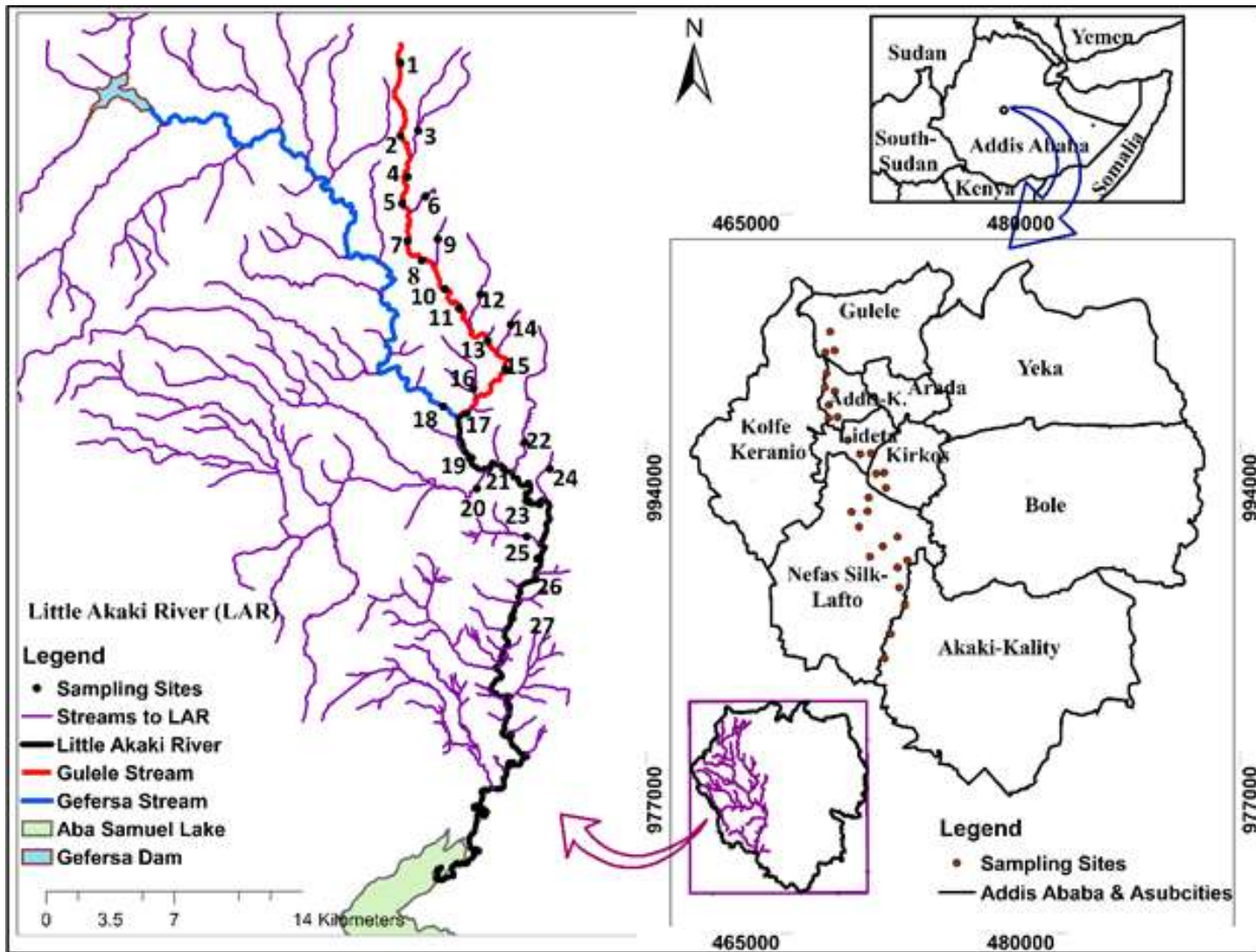


Figure 4-1 Map of AA city and LAR

and mean to assess the risks of pollution and summarize data sets in easier and more understandable form.

Chen *et al.* (2016) described direct correlation of higher population density and higher organic contaminant discharge, such as food waste, and human sewage. However, EC was highest in the more industrialized NL sub-city (Florescu *et al.* 2011).

4.1.2. Multivariate Analysis

- **Data Treatment**

For the purpose of multivariate statistical analysis, raw wet season WQ data of LAR were subjected to descriptive statistics: range, standard deviation and mean to assess the risks of pollution and summarize data sets in easier and more understandable form. Other reports that suggested this approach include (Chapman 1992b; Voutsas *et al.* 2001; Bhuiyan *et al.* 2011; Khan *et al.* 2015; Taoufik *et al.* 2017). The summarized data is given in **Table 4-2**.

Table 4-2 Basic statistical analysis of water quality data of LAR (Wet Season)

Parameter	Minimum	Maximum	Mean	Standard Deviation	Local Standard*¹
TSS, mg/L	20.0	1062.0	238.4	205.5	50
T-NH₃ as N, mg/L	0.4	83.0	29.7	20.0	0.2
NO₃ as N, mg/L	0.0	12.1	2.5	3.1	50
NO₂, mg/L	0.0	0.7	0.1	0.2	0.1
T-N, mg/L	0.0	97.0	24.3	26.8	1
T-P, mg/L	0.0	18.4	5.3	4.7	0.005
PO₄, mg/L	1.8	105.4	37.0* ²	25.3	Should not be impacted by > 15%
T-Alk., mg/L	7.5	975.0	439.1	199.7	Not Stated
TOC, mg/L	2.1	21.4	16.3	4.6	Not Stated
COD, mg/L	25.0	1580.0	368.3	314.9	Not Stated
BOD, mg/L	20.0	720.0	164.9	156.7	≤ 5
DO, mg/L	0.0	9.5	2.2	2.6	> 4
T°, °C	9.8	23.4	19.4	3.5	Should not result in >3 ⁰ C variation
EC, mS/cm	0.3	5.1	1.1	8	1 at 20 ⁰ C
pH	6.7	9.0	8.1	0.5	6-9

*¹ EPA and UNIDO, 2003. (EPA & UNIDO 2003)

*² un-impacted concentration at upstream was 1.8mg/L

The high standard deviation in the data set indicates WQ spatial variation. Such results were mentioned elsewhere (Vega *et al.* 1998). The means of the parameters were compared with the surface WQ standards set by the EPA of Ethiopia: ‘Guideline Ambient Environment Standards for Ethiopia’(EPA & UNIDO 2003). Results show NO₃, NO₂, T⁰ and pH were within the allowable limits compared with local surface water standards. However, parameters such as TSS, T-NH₃, T-N, T-P, PO₄, BOD and DO violated the recommended limits.

A mean pH value of 8.1 showed that LAR was slightly alkaline although it still meets the allowable range. The maximum pH value of 8.9 was observed at sampling site 24 (Teramaj stream) in **Figure 4-1**. This site is located in one of the most industrialized sub-cities: NL sub-city. The EC of LAR and its tributaries range from 310µS/cm to 1,340µS/cm and 750µS/cm - 5,130µS/cm, respectively. This indicates that parts of the river and tributaries violated the 1000.0 µS/cm threshold value. Teramaj stream’s maximum EC value of 5,130 µS/cm suggests discharge of industrial (Florescu *et al.* 2011) waste in addition to the domestic one. T⁰, as expected, increased gradually from 10⁰C at the upstream of the river to 23.4⁰C downstream with little irregularity.

The DO concentrations of LAR and its tributaries were found to be in the range from 0.02mg/L to 7.27mg/L and from 0.11mg/L to 9.53mg/L, respectively with 89% of the sampled sites (24 from the 27 sites) showing DO concentrations (under 4mg/L) below the standard. In this regard, DO concentrations under 5 mg/L may negatively impact the survival and functioning of biological population and under 2 mg/L may result in the loss of most fish (Chapman 1992b).

TOC with mean concentration of 16.4mg/L (greater than 10mg/L) implies the discharge of municipal and/or industrial waste. The mean COD concentration of 386.3mg/L is greater than 20mg/L or less for clean waters to greater than 200 mg/L in waters getting effluents. In addition,

industrial wastewaters may have COD values ranging from 100 mg/L to 60,000 mg/L O₂. This indicates the susceptibility of waste matter in LAR to oxidation. Furthermore, BOD of a raw or untreated sewage is around 600 mg/L, while treated sewage effluents have BOD values ranging from 20 to 100 mg/L subject to the level of treatment applied (Chapman 1992b).

The range of BOD concentration of LAR were found to be from 20mg/L to 242mg/L and from 27mg/L to 720mg/L for LAR and tributaries, respectively. The mean BOD concentration of 164.9mg/L (local standard of 5mg/L maximum) and maximum concentration of 720mg/L at sampling site 6 near Mesalemiya in Addis Ketema sub-city implies that the tributary was loaded with high level of sewage effluent. The 37mg/L phosphate mean concentration of LAR was 20 times higher than the concentration at the upstream of the river. This violates the local standard recommending concentrations not to exceed the un-impacted condition by 15%.

The raw data, in this study, was $\log_{10}(X)$ transformed to get normal distribution for the data and then was transformed by Z-score to get dimensionless data and hence reduce the difference in units and achieve comparable ranges for variables. The Kaiser-Meyer-Olkin (KMO) value of 0.726 (> 0.5) suggests that the data set is acceptable for PCA analysis. In addition, the Bartlett's sphericity test with a significance value of 0.000 indicates a sufficient correlation within parameters hence fit for a PCA. Similar approach was noticed in other reports (Aruga *et al.* 1995; Brūmelis *et al.* 2000; Liu *et al.* 2003; Singh *et al.* 2005; Shrestha & Kazama 2007; LI *et al.* 2009; Wu & Kuo 2012; Kim *et al.* 2015; Monica & Choi 2016; Powley *et al.* 2016; Ling *et al.* 2017).

- **Cluster Analysis (CA)**

The log-transformed and z-transformed data set was used for CA (Singh *et al.* 2009b). The vertical dendrogram (**Error! Reference source not found.****Error! Reference source not**

found.) obtained from Ward’s method CA on normalized data set implies that the twenty-seven sampling sites, based on their similarity, may be grouped stepwise into three at $(D_{link}/D_{max}) \times 100 < 40$. Similar method was used in other reports (Alberto *et al.* 2001; Shrestha & Kazama 2007; Singh *et al.* 2009b).

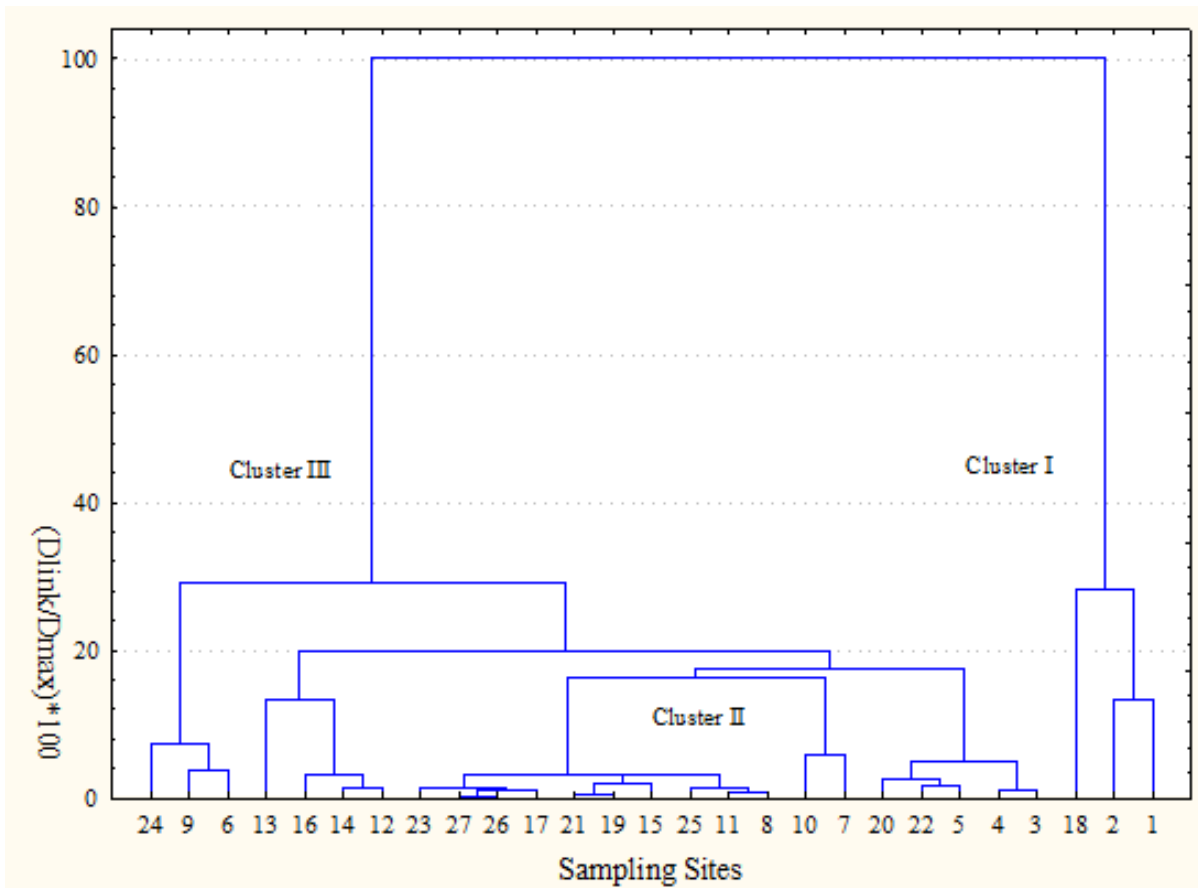


Figure 4-2 Dendrogram showing clustering of sampling sites of LAR

In this study, an assessment based on the local standard generally categorizes the river WQ in the three clusters as polluted. However, the degree of pollution identified in all sampled sites was not uniform. To account for the difference in relative pollution levels among the sampled sites, the three clusters were given pollution rank based on methods observed in different papers and ranges specified in .

Singh *et al.* (2004) also obtained three clusters having similar characteristic attribute and natural background source types. Singh *et al.* (2004) then categorized the clusters based on land use condition (such as the upstream as low pollution (LP), areas with various waste drains and polluted tributaries as high pollution (HP), downstream part demonstrating WQ recovery due to self-purification as medium pollution (MP). Similarly, Shrestha and Kazama (2007) categorized three sites in the upstream and one site in the most downstream as relatively less polluted (LP). They also categorized the sites with pollution from domestic sources and industries as high pollution (HP). A similar approach was also noticed in another report (Wang *et al.* 2014).

In this study also, the pollution rank was specified in **Table 4-3** based on LAR appropriateness for a particular use. Its use is affected by the WQ status that is commonly caused by the land use condition. LAR application includes irrigation activity. Because of this, pollution rank was defined based on the river appropriateness for recreational purpose only (as medium pollution) and for irrigation activity or no use at all (as high pollution).

Consequently, Cluster I is formed by three sampling sites (1, 2 and 18) representing the most upstream part in G sub-city and a stream from Gefersa water supply dam sampled at Mekanisa. This Cluster, based on **Table 4-3** is designated as relatively Less Pollution (LP). Likewise, Cluster II involves seventeen sampling sites (3, 4, 5, 7, 8, 10, 11, 15, 17, 19, 20, 21, 22, 23, 25, 26 and 27) and cluster III consists of seven sampling sites (6, 9, 12, 13, 14, 16 and 24). These Clusters are designated as Medium Pollution (MP) and High Pollution (HP) respectively. In contrast, 85% of the sampling sites in this group (that is six out of seven sampling sites in the cluster) are from the tributaries mostly in the densely populated sub-cities: AK, L and K. Higher concentrations of TSS, T-NH₃, T-N, T-P and PO₄ were observed in group

III as compared to Clusters I and II indicating the presence of increased sewage and detergent discharge to parts of the river (Kundu *et al.* 2015).

Table 4-3 Cluster pollution rank

Clusters	Average values for parameters (mg/L)											
	TSS	T-NH ₃	NO ₃	NO ₂	T-N	T-P	PO ₄	TOC	COD	BOD	DO	EC*
I	50.3	7.9	2.5	0.4	1.3	0.3	5.5	6.5	49	24	7.9	0.5
II	190.4	24.6	3.4	0.1	20.1	3.8	31.6	17.8	293.9	120.1	1.5	1
III	435.6	51.7	0.5	0	44.1	11	63.7	16.8	685.8	334	1.4	1.6

Data Assessment Approach:

Direction:	High local standard violation	Relative pollution rank based on parameters average value:
	Medium local standard violation	Cluster I Less pollution
	Within standard	Cluster II Medium pollution
	No local standard stated	Cluster III High pollution

Parameter	Local standard	Medium pollution	High pollution	References
TSS, mg/L	50	150 – 300	> 300	(DOE 2007)
T-NH ₃ as N, mg/L	0.2	0.9-2	>2.7	(DOE 2007)
NO ₃ as N, mg/L	50	not applicable	not applicable	
NO ₂ , mg/L	0.1	0.15 - 1	>1	(NEPA 1988)
T-N, mg/L	1	1.5 -2	>2	
T-P, mg/L	0.005	0.1-0.2	>0.2	(NEPA 1988)
PO ₄ , mg/L	-	not applicable	not applicable	
BOD, mg/L	≤ 5	10-30	>30	(USEPA 2012)
DO, mg/L	> 4	2 - 4	<2	(DOE 2007)
EC, mS/cm	1 at 20 ⁰ C	1-2	>2	(FAO 1976)

*unit: in mS/cm

This clustering result allows data interpretation and monitoring of LAR using three sampling sites to the minimum. Hence, cost of sampling and analysis can be minimized without significantly losing information on the WQ variation of the river. Similar applications of such approach to WQ analysis were observed in other reports (Singh *et al.* 2005; Shrestha & Kazama 2007; Singh *et al.* 2009b; Khan *et al.* 2015; Kundu *et al.* 2015).

- **Discriminant Analysis (DA)**

DA was used to know the most critical parameters in relation with variation between the clusters. Primary data and three spatial groups obtained from CA were used in the DA. In this case, WQ parameters were treated as independent while pollution groups were considered as dependent variables. Classification matrixes (CM) in **Table 4-4** show that 100%, 96.3 % and 85.2 % of the data points were correctly classified in the standard, forward as well as backward stepwise modes, respectively. The standard DA approach constructs discriminant functions (DFs) using fifteen parameters while it was only two and six parameters for backward stepwise and forward stepwise modes, respectively.

Table 4-4 Classification matrix for DA of spatial variation in LAR

Monitoring region	% Correct	Region assigned by DA		
		C1	C2	C3
Standard DA mode				
C1	100.0	3	0	0
C2	100.0	0	17	0
C3	100.0	0	0	7
Total	100.0	3	17	7
Forward stepwise DA mode				
C1	100.0	3	0	0
C2	100.0	0	17	0
C3	85.7	0	1	6
Total	96.3	3	18	6
Backward stepwise DA mode				
C1	100.0	3	0	0
C2	94.1	1	16	0
C3	57.1	0	3	4
Total	85.2	4	19	4

^{C1} includes stations (1, 2 and 18)

^{C2} includes stations (3,4,5,7,8,10,11,15,17,19,20,21,22,23,25,26 and 27)

^{C3} includes stations (6, 9,12,13,14, 16 and 24)

In forward stepwise mode, pollutants such as NO₃, NO₂, T-P, T-Alk., TOC and T⁰ were found as insignificant variables indicating that these have less variation in terms of their spatial distribution. Backward stepwise mode deleted least significant and discriminated two significant variables (T-P and DO). These two parameters, with 85.2% correct assignation, are the critical parameters useful to make distinction within the three pollution groups. This implies that expected spatial WQ variation can be explained sufficiently using T-P and DO. Wilks' lambda shows discriminatory capacity of functions for the partition of cases into groups. p-level < 0.05, in this case shows that both functions are statistically significant. Similar approach was used for WQ assessment in different reports (Alberto *et al.* 2001; Singh *et al.* 2005; Qadir *et al.* 2008; Juahir *et al.* 2011; Zheng *et al.* 2016; Bouguerne *et al.* 2017; Hajigholizadeh & Melesse 2017; Taoufik *et al.* 2017).

Box and whisker plots of the discriminating parameters obtained from DA (backward stepwise approach) are used to visually evaluate spatial WQ variation patterns. Similar trends of spatial variations with inverse relationship were observed for DO and T-P in **Figure 4-3**. For instance, cluster 3 has the highest mean TP value resulting in less than 1mg/L DO concentration. Which may be because of consumption of DO resulting from increase in biological activity as a result of high nutrient availability in this section of the River. Box plots with long whiskers at the top of the box (as shown for DO and T-P in cluster 2) imply that underlying distribution is skewed toward high concentration. On the other hand, box plots with large spread such as T-P in cluster 3 imply WQ spatial variation.

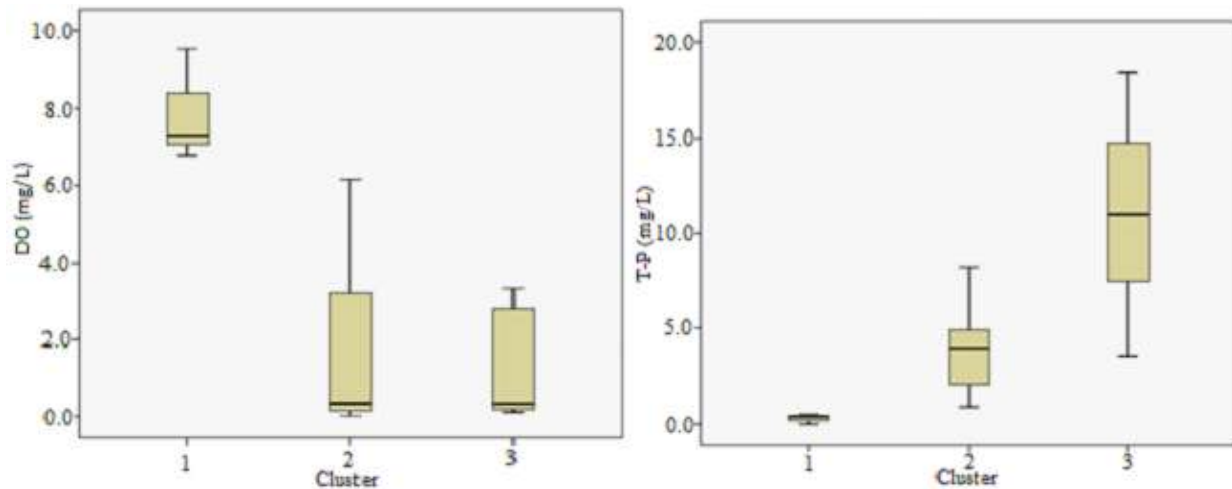


Figure 4-3 Box plot of the most discriminating parameters

In this regard, DO and TP had the highest mean value in cluster 3 and cluster 1, respectively due to the inverse correlation effect. Unlike DO, higher concentration of T-P was observed in cluster 3 in the downstream due to accumulated sanitary sewage, industrial discharge and agricultural runoff. Generally, from the plots it can be realized that WQ deterioration increased from cluster 1 to cluster 3 (Vega *et al.* 1998; Bilgin & Konanç 2016).

- **Principal Component Analysis**

The correlation extent among parameters is important in data reduction process in PCA. Component loadings are the linear combinations in each PC and explain the correlation between the primary variables and the obtained components (Ouyang *et al.* 2006; Bilgin & Konanç 2016). In this study, the correlation among parameters was not predominantly strong, especially NO_3 and NO_2 had relatively weak correlation with the rest of the parameters having coefficients less than 0.326 (absolute value). While DO resulted in medium inverse correlation with PO_4 (0.590), TOC (0.505), BOD (0.622) and EC (0.507) due to partial oxidization of organic matter by oxygen, it did not have strong correlation with any of the parameters (Vega *et al.* 1998). Some

parameters had strong positive correlation. Coefficients of parameters that had strong correlation with T-NH₃ are T-P (0.915), PO₄ (0.885), COD (0.844), EC (0.771) and pH (0.768). In addition, coefficients of parameters that had strong correlation with PO₄ are TSS (0.887), T-NH₃ (0.885), T-P (0.891), COD (0.914), BOD (0.877) and EC (0.763). Moreover, coefficients of parameters that had strong correlation with COD are: TSS (0.876), T-NH₃ (0.844), T-P (0.808), PO₄ (0.914) and BOD (0.873) which is due to organic contamination related to domestic sewage (Vega *et al.* 1998; Armah *et al.* 2010).

Eigenvalue higher than unity means that a PC contains more information than one original variable, offering opportunity for reduction of data dimensionality. The PCA analysis on a normalized data resulted in four PCs (**Table 4-5**) with Eigenvalues >1 accounting for 81.9% of the total WQ variation (Helena *et al.* 2000). 54.5%, 10.1%, 9.4% and 7.8% of total variance is expressed by PC1, PC2, PC3 and PC4, respectively with corresponding eigenvalues of 8.18, 1.51, 1.41 and 1.17. As for PC1, the parameters demonstrating strong positive loading are TSS, T-NH₃ as N, T-N, T-P, PO₄, COD, BOD and EC. While PC2 was with no parameters demonstrating strong loading, PC3 has NO₃ with strong positive loading. Similarly, PC4 shows NO₂ as a strong positive loading. Therefore, ten parameters are required to describe 81.9% of WQ variations. The data reduction from the initial 15 parameters is just 33% and is not considered significant. Varimax rotation was conducted on PCA components to minimize the participation of less significant variables.

Table 4-5 PCs loading for parameters

Particular	Prin.1	Prin.2	Prin.3	Prin.4
TSS	0.87756	-0.23460	-0.15756	-0.00529
T-NH ₃	0.92129	0.26991	-0.02188	-0.00410
NO ₃	-0.29332	-0.08140	0.84705	-0.25759
NO ₂	-0.19001	0.10444	0.49776	0.77267
T-N	0.74857	-0.38446	0.19392	0.37448
T-P	0.90498	0.06342	0.00875	0.13243
PO ₄	0.97929	-0.06615	-0.03687	0.03723
T-Alk	0.31016	0.01330	0.41989	-0.54444
TOC	0.66729	-0.50649	0.17160	0.06687
COD	0.92741	-0.08079	-0.10469	-0.03067
BOD	0.87004	-0.20496	-0.20533	-0.12188
DO	-0.58450	0.33965	-0.19899	0.13554
T ^o	0.72509	0.50381	0.00797	0.08372
EC	0.80070	0.27678	0.27386	-0.07978
pH	0.66017	0.67718	0.11227	-0.02399

Strong loading for parameters such as TSS, T-NH₃, T-N, T-P, PO₄, BOD, NO₃-N, and NO₂ in PC1 (**Table 4-5**) can be mainly due to sewage and detergent discharge from domestic, food and beverage processing industries and agricultural runoff from the peri-urban section of LAR watershed. Consequently, odor is felt in many parts of the LAR due to the release of NH₃

as an implication of organic matter decomposition. This makes the water harmful to aquatic life and ill-suited for potable use (Wu & Kuo 2012; Kundu *et al.* 2015; Powley *et al.* 2016).

- **Factor Analysis**

Varimax rotation was done on PCA components to degrade the influence of variables having limited significance. After varimax rotation, the number of parameters (**Table 4-6**) with strong loading remained 10 in number maintaining the same data reduction capacity as PCA. However, the loadings of parameters were changed after rotation. For example, EC loading changed from 0.800 to 0.427 after rotation. PCA/FA was implemented to downsize data dimensionality in different reports (Brūmelis *et al.* 2000; Alberto *et al.* 2001; Qadir *et al.* 2008; Singh *et al.* 2009b; Bhuiyan *et al.* 2011; Jung *et al.* 2016).

LAR WQ is affected due to pollution from industrial plus domestic sources. Therefore, this study takes these sources as the *common pollution factors*. In this regard, for Factor 1, component matrix (**Table 4-6**) dictates that, T-N, TOC, TSS, PO₄, BOD, COD, T-P and DO belong to VF1 which can be due to the presence of sewage and detergents. Consequently, it can be labeled as *domestic pollution factor*. The biodegradation of these organic matters causes concentrations of waterborne BOD, DO, and T-NH₃ (Chapman 1992b; Vega *et al.* 1998; Shrestha & Kazama 2007; Wu & Kuo 2012). Likewise, DO utilization for the disintegration of organic materials is shown by the negative DO loading in VF1 (Singh *et al.* 2004). Also, T-NH₃ (caused by the decomposition of organic matter) and TN as the significantly high loading variables indicate domestic sewage discharge in the river (Shrestha & Kazama 2007; Wu & Kuo 2012).

Table 4-6 Rotated component matrix

Particulars	Component			
	VF1	VF2	VF3	VF4
T-N	.873			
TOC	.847			
TSS	.794	.381		
PO ₄	.776	.584		
BOD	.750	.395	-.351	
COD	.732	.534		
T-P	.653	.637		
DO	-.652			
pH		.949		
T ^o		.854		
T-NH ₃	.510	.801		
EC	.427	.747		
NO ₂			.942	
NO ₃				.864
T-Alk				.636

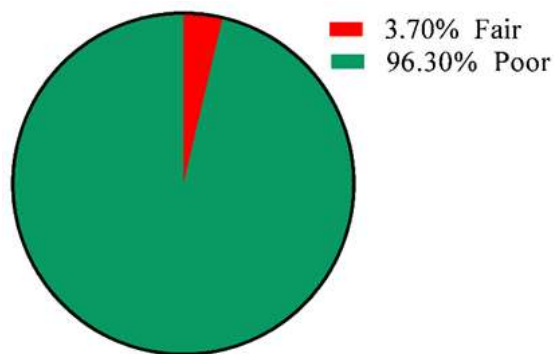
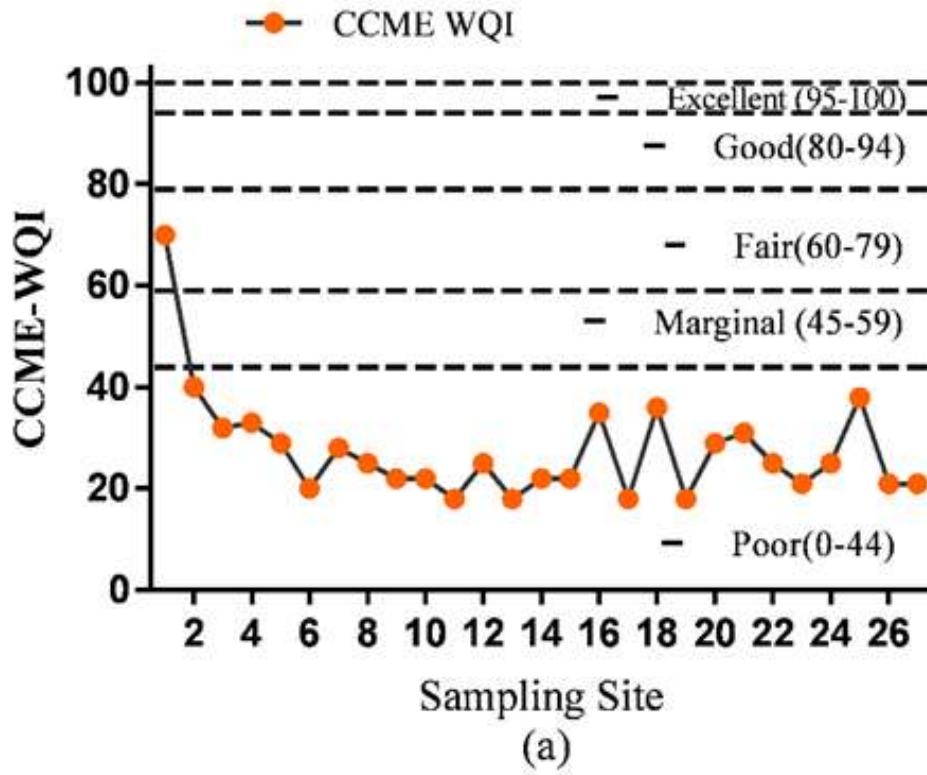
EC, pH and T^o loads heavily on Factor 2 with low loading on Factor 1. Moreover, correlation between pH and EC (0.755) and EC and T^o (0.644) is strong (greater than 0.5). In this aspect, electrical conductivity measurements attributable to the influx of salts serve as a useful

indicator of industrial pollution (Webster 1995; Florescu *et al.* 2011). Therefore, Factor 2 can be labeled as *industrial pollution factor*.

4.2. Water Quality Index

4.2.1. CCME-WQI

Raw data of LAR was used for CCME-WQI determination. The indices obtained are presented in **Figure 4-4**. It shows that, except one in the river headwater, the rest (96.3%) fall entirely in the *poor* WQ category. Consequently, the WQ is considered as impaired and that usually departs from desirable levels. For the purpose of comparison, no previous WQI analysis on LAR was observed. However, this result is in line with other reports indicating LAR WQ deterioration (Van Rooijen *et al.* 2010; Weldegebriel *et al.* 2012). The headwater, on the other hand, was of *fair* quality but degraded due to streams coming from residential sources. Lack of strict increment in WQ degradation in the downstream direction could be the result of abstraction for urban irrigation, pollution sources with different flowrates and variation in concentration in the different sources (Whitehead *et al.* 2009).



Total=27

(b)

Figure 4-4 Water Quality Index based on CCME for LAR

Figure 4-5 (a) generally shows inverse relationship between BOD and WQI. Also, the trend line equation and R^2 values in **Figure 4-5 (b)** indicate weak relationship between the two. However, the graphs indicate the importance of BOD as an important WQ parameter. Its use as an instrument for the control of dischargers was observed in different papers (Burn & Lence 1992; Carmichael & Strzepek 2000).

Population density is related to pollution of rivers and streams in urban areas (Chen *et al.* 2016). **Figure 4-5 (c)** shows the impact of population density on WQI of LAR. The figure reveals opposite trend between WQI and population density. This implies the likelihood of WQ deterioration because of population growth. However, the inverse relationship between WQI and population density is not very strong as can be seen from **Figure 4-5 (d)** and an R^2 value of 0.022. In addition, stronger inverse relationship is observed between BOD and WQI than population density with WQI.

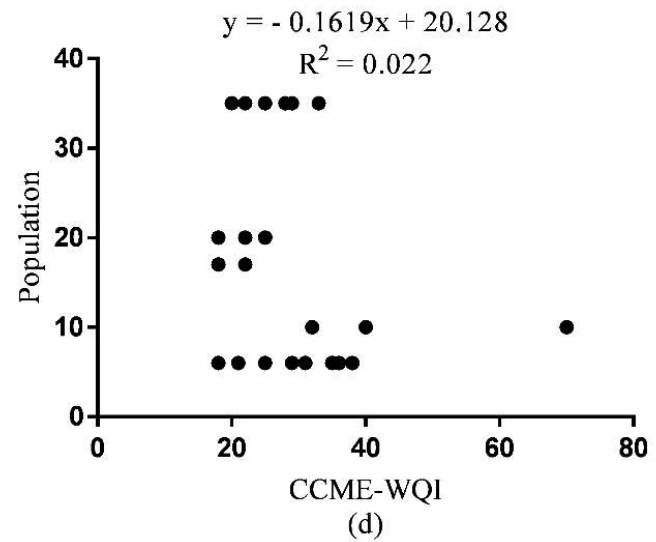
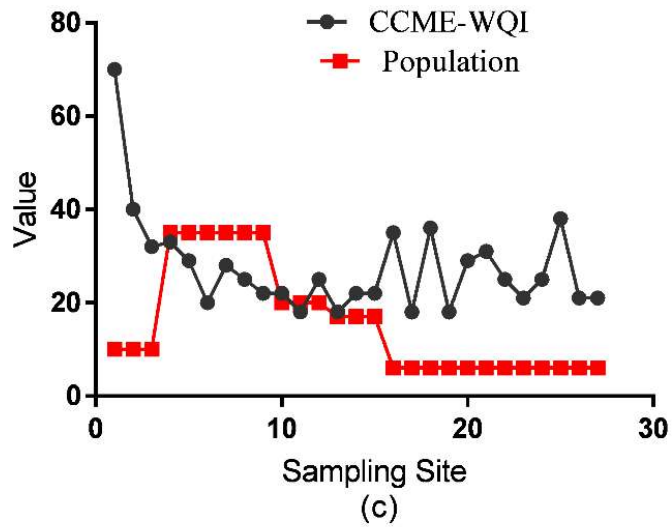
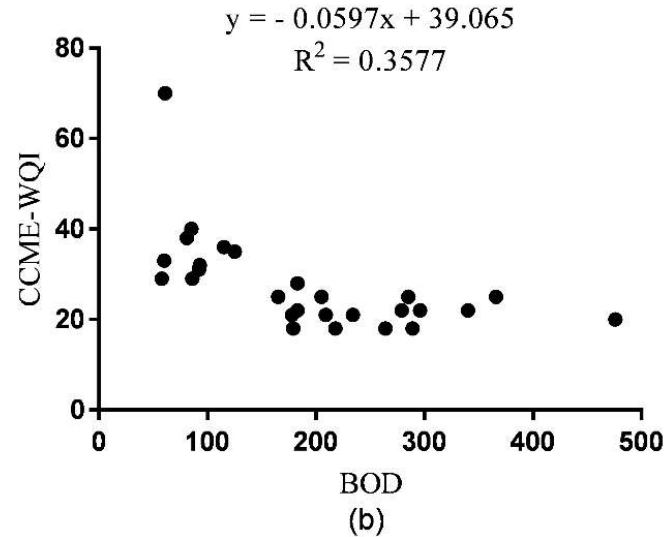
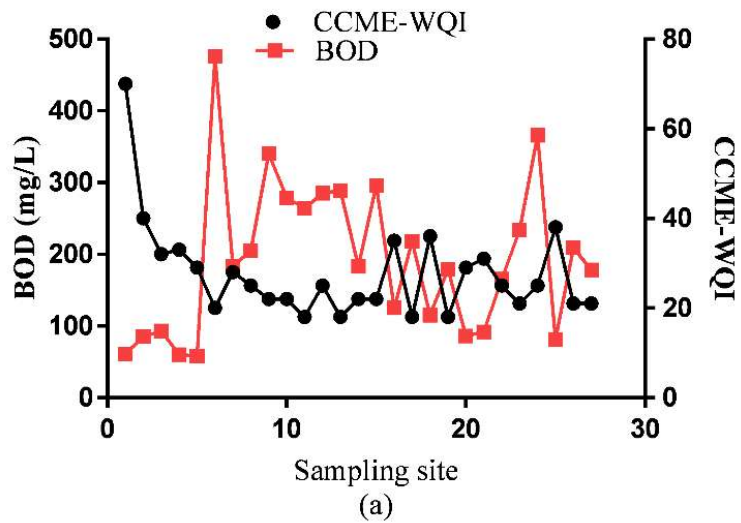


Figure 4-5 LAR BOD, CCME-WQI and Population Density relationship

4.2.2. Water Quality Index based on Artificial Neural Network

The model results improved when run using logsig activation function for the hidden layer with purelin at the output layer. **Figure 4-6 (a-d)** show the performance of the model against layers and neurons number in the hidden layers. In this case, although there is no clear pattern in the relationship, performance of ANN generally improves with lower number of hidden layers. Moreover, number of neurons that is a little greater than the number of layers improves the performance. In this regard, eight layers and fifteen neurons were found as the most optimal combinations.

The performance of the modeling was analyzed using Root Mean Square Error (RMSE), Mean Square Error (MSE) and R^2 . **Table 4-7** shows some of the of the model runs with good performance. From this, ANN3 model architecture (Hidden layers: 8 and Hidden neurons: 15) with values of RMSE: 1.97, MSE: 0.07 and R^2 : 0.93 provided the best result. The values generally indicate that ANN can successfully model CCME-WQI data. Similar approach was observed in other reports (Singh *et al.* 2009a).

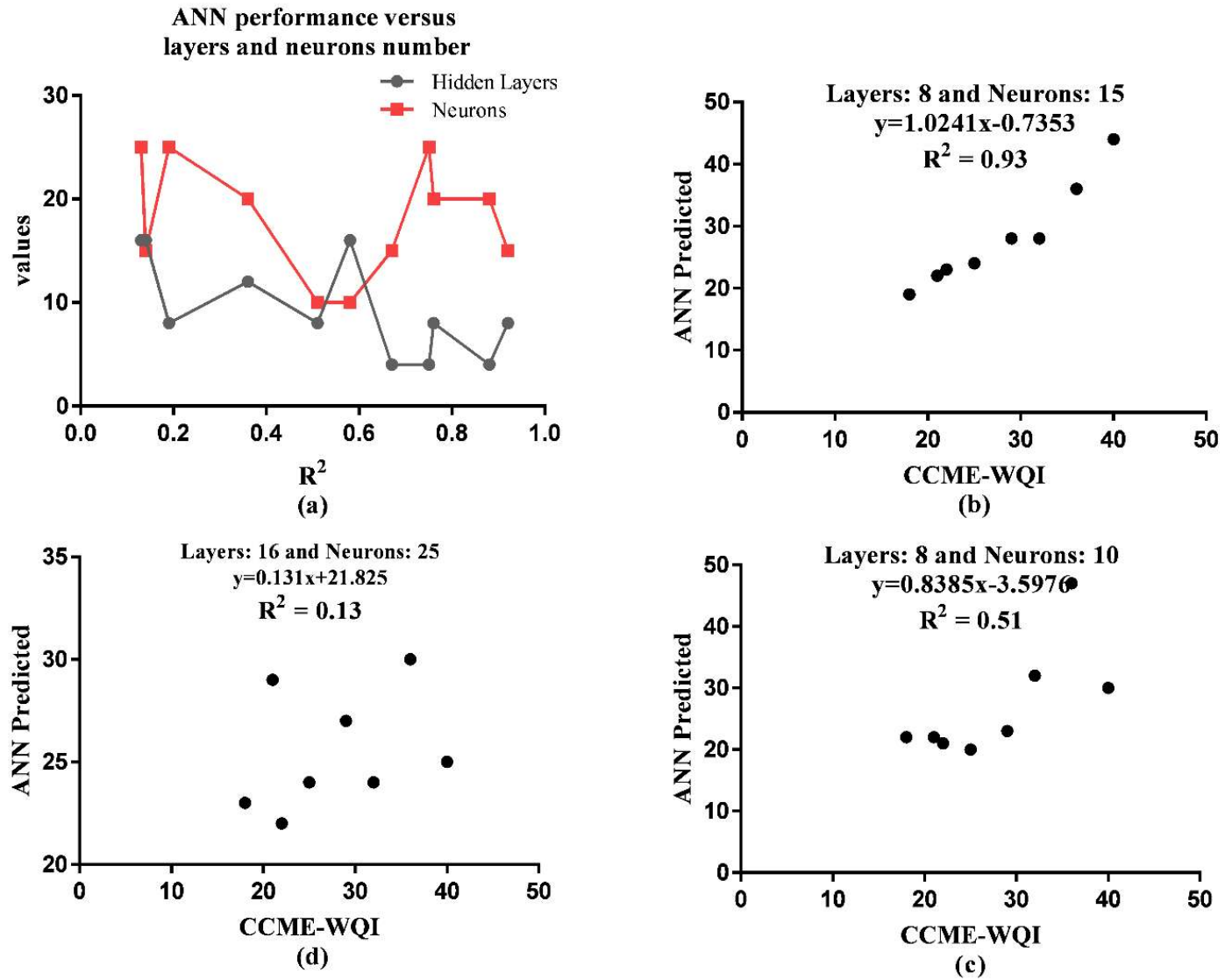


Figure 4-6 Performance of CCME-WQI prediction using ANN model for LAR

Table 4-7: ANN model prediction performance

NN Predictions	RMSE	MSE	R ²	ANN			
				R	Layers	Neurons	Learning Function
ANN1	4.62	1.77	0.62	0.9528	4	10	Learngdm
ANN2	4.70	4.15	0.60	0.9969	3	15	Learngdm
ANN3	1.97	0.07	0.93	0.9993	8	15	Learngdm
ANN4	3.63	0.01	0.87	0.9986	4	20	Learngdm
ANN5	4.64	4.15	0.61	0.9949	12	15	Learngd
ANN6	4.59	4.15	0.75	0.9917	14	25	Learngd

4.3. Q2Kw Simulation Model

4.3.1. Calibration and Verification

The calibrated parameter values in the model are presented in **Error! Reference source not found.** These ranges of model rate parameters, required by Q2Kw, were obtained from various literatures (Kannel *et al.* 2007b) including: Environment Protection Agency (EPA) guidance document (USEPA, 1985b), Q2Kw user manual (Pelletier & Chapra 2005) and documentation for the enhanced stream WQ model Q2E and Q2E-UNCAS (Brown & Barnwell 1987).

Table 4-8 List of Rate Parameters and respective Calibration Values

Parameter	Value	Units	Min value	Max value
<i>Inorganic suspended solids:</i>				
Settling velocity	0.382	m/d	0	2
<i>Oxygen:</i>				
User reaeration model parameter A	5.262		3	6
User reaeration model parameter B	0.9245		0.5	1
User reaeration model parameter C	-1.78735		-1.85	-1.5
<i>Slow CBOD:</i>				
Hydrolysis rate	0.0982	/d	0.05	0.25
Temp correction	1.047		1	1.07
Oxidation rate	3.265	/d	0	5
Temp correction	1.047		1	1.07
<i>Fast CBOD:</i>				
Oxidation rate	3.85	/d	0	5
Temp correction	1.047		1	1.07
<i>Organic N:</i>				
Hydrolysis	0.17975	/d	0.05	0.3
Temp correction	1.07		1	1.07
Settling velocity	0.6662	m/d	0.05	2
<i>Ammonium:</i>				
Nitrification	1.7079	/d	0.05	3
Temp correction	1.07		1	1.07
<i>Nitrate:</i>				
Denitrification	0.626	/d	0	2
Temp correction	1.07		1	1.07
Sed denitrification transfer coeff	0.936	m/d	0	1
Temp correction	1.07		1	1.07
<i>Organic P:</i>				
Hydrolysis	0.171	/d	0.05	0.3
Temp correction	1.07		1	1.07
Settling velocity	0.3503	m/d	0.05	2
<i>Inorganic P:</i>				
Settling velocity	1.292	m/d	0	2
Sed P oxygen attenuation half sat constant	0.73	mgO ₂ /L	0	2
<i>Detritus (POM):</i>				
Dissolution rate	0.4766	/d	0.05	0.5
Temp correction	1.07		1.07	1.07
Settling velocity	0.87485	m/d	0.05	2

The dissolved oxygen in the LAR is below 2mg/L nearly all of the twenty seven sampled sites except two in the upstream. Moreover, the River has foul odor and grey to dark color. It shows that oxygen supply, against the deoxygenation, is insufficient to keep the river under aerobic condition (Rowe, 1995). As a result of this, the condition of the river can be categorized as anaerobic. Under this condition hydrogen sulfide, methane and carbon dioxide are released to give the river unpleasant smell (Rao, 2002). Continuous sewage discharge into the River, biological oxygen demand and turbidity create septic condition (McKinney, 2004; Caduto, 1990 #791). In this consideration, due to significant anaerobic condition of the River, the phytoplankton and bottom algae rate parameters were disregarded among the calibration rate parameters in the QUAL2Kw model. In this regard, reaeration was considered as the main re-oxygenation mechanism that is used by LAR.

The calibration results based on four selected calibration-parameters (DO, BOD, TP and TN) at ten monitoring locations in ten reaches are shown in **Figure 4-7**. The horizontal axis of each graph indicates the distance (km) upstream of the mouth of the canal (distance 0km is downstream end). The model calibration results are in well agreement with the measured data, with some exceptions. R^2 values 0.91, 0.90, 0.81 and 0.89 of DO, BOD, TP and TN, respectively indicate the effectiveness and interrelationships between the observed and simulated values. For instance, Kannel *et al.* (2007a) and Sharma *et al.* (2015a) considered R^2 value that is greater than 0.5 as acceptable especially for the developing countries where the financial resources are often limited for higher accuracy data analysis and frequent monitoring campaigns.

Similar to calibration results, the calibrated model resulted in acceptable agreement between measured wet season data and model prediction (**Figure 4-8**). In this regard, the model was run under different WQ conditions, but without changing the calibrated parameters. And the

R^2 values obtained between the simulated and observed values of DO, BOD, TN and TP are 0.98, 0.82, 0.77 and 0.75, respectively. Attending to this good agreement, the model can be used to test

different scenarios in a general strategy to understand the river response to different pollution loads. Similar approach was also observed in other reports (Oliveira *et al.* 2012).

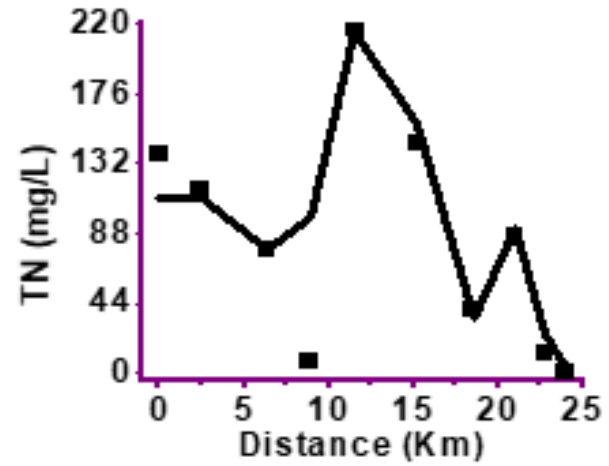
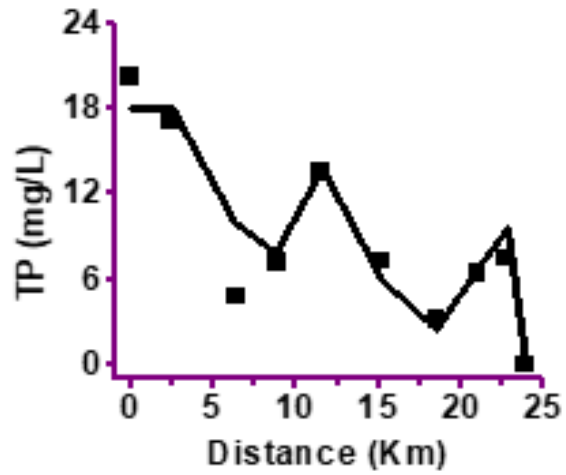
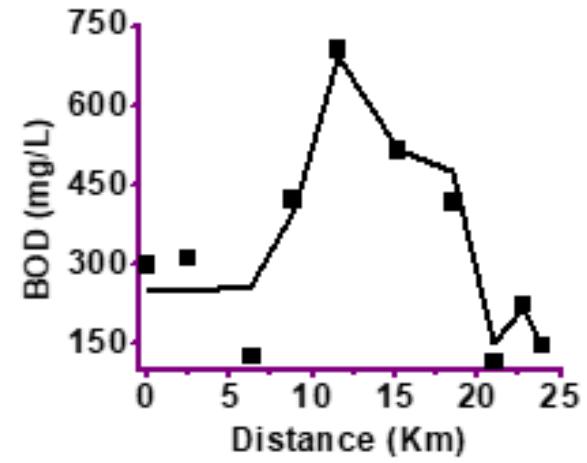
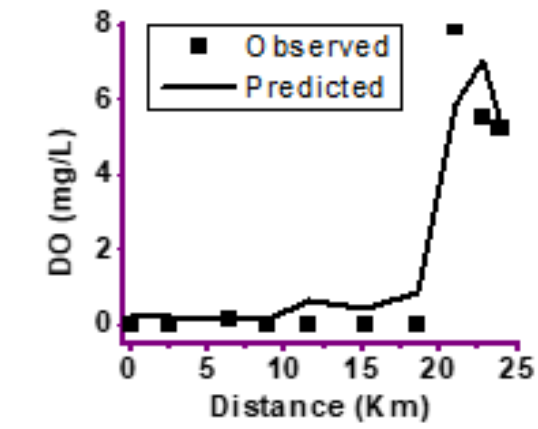


Figure 4-7 Model calibration

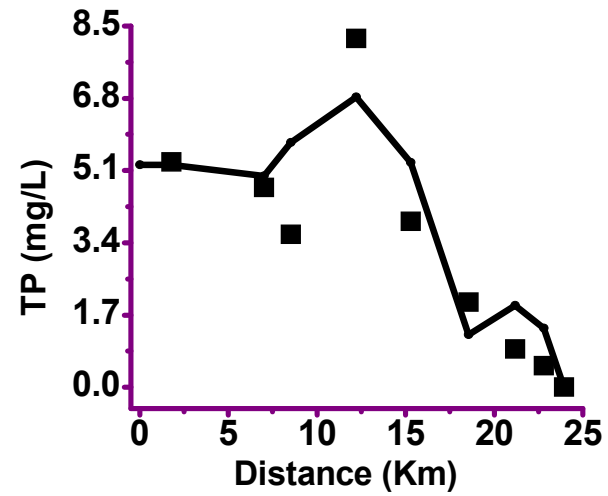
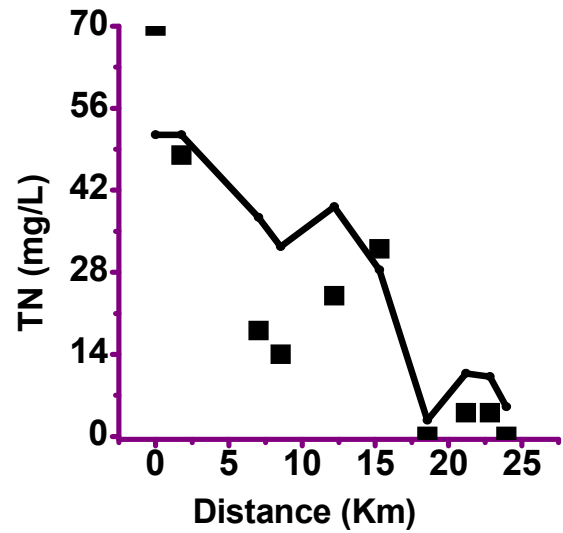
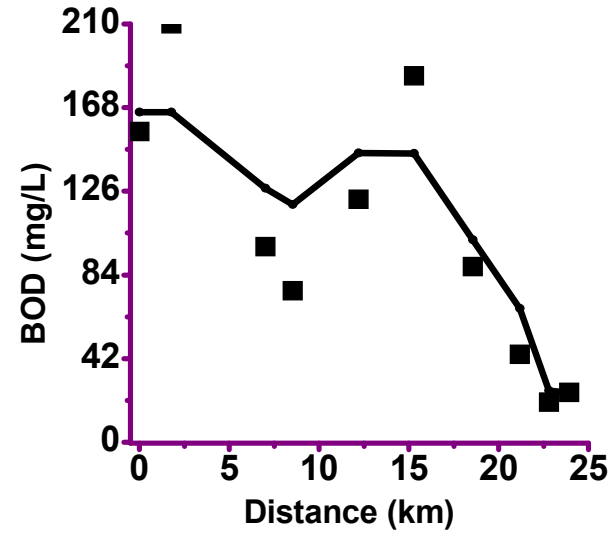
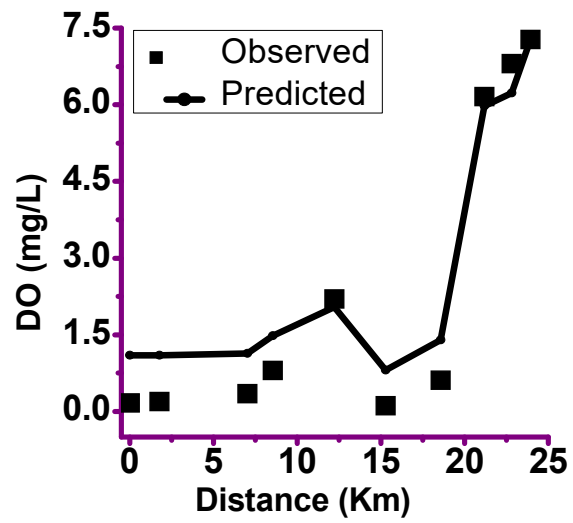


Figure 4-8 Model verification

4.3.2. Sensitivity analysis

Sensitivity analysis is crucial to quantify the error associated with the variation of the parameters of greater importance and to evaluate the model robustness. Moreover, sensitivity analysis allows to understand the parameters to which the model is more sensitive and more effort should be applied in order to precisely determine those parameters (Oliveira *et al.* 2012). The sensitivity analysis has been done to identify the parameters having maximum influence on the significant parameters: DO, BOD, TP, and TN output. The analysis was performed by varying the values of the rate parameters, Manning's n factor and flowrates while keeping all other parameters constant. It was found that the model was highly sensitive to point source flow, Manning's n, InorgP Settling velocity and Sed denitrification transfer coeff. Similar approach was observed in other reports (Kannel *et al.* 2007b) (Sharma *et al.* 2015a).

The results of the sensitivity analysis are presented in **Table 4-9** where it is seen that the DO model is highly sensitive to Manning's n and Sed denitrification transfer coeff. BOD is significantly sensitive to point source flow, InorgP Settling velocity, OrgP Settling velocity, Denitrification and Sed denitrification transfer coeff. The TN concentration is primarily affected by point source flow, InorgP Settling velocity and Periphyton Basal respiration rate. Whereas TP concentration is impacted by point source flow, InorgP Settling velocity, Manning's n, Sed denitrification transfer coeff and HTS Respiration rate. Other reports with similar approach include (Sharma *et al.* 2015a).

Table 4-9 Sensitivity analysis results

Parameter	Degree of parameters variation	DO	BOD	TN	TP
		degree of disturbance due to parameters variation			
Point source flow	20%	0.68%	20.18%	19.46%	17.00%
InorgP Settling velocity	20%	0.64%	1.78%	2.68%	1.57%
OrgP Settling velocity	20%	0.67%	1.45%	0.70%	0.65%
Denitrification	20%	0.28%	1.19%	0.90%	0.94%
Sed denitrification transfer coeff	20%	1.44%	1.01%	0.70%	1.07%
Periphyton Basal respiration rate	20%	0.18%	0.98%	1.30%	0.87%
Mannings n	20%	6.65%	0.91%	0.84%	2.09%
HTS Respiration rate	20%	0.43%	0.79%	0.99%	1.06%
CBODslow Hydrolysis rate	20%	0.26%	0.52%	0.47%	0.59%
Oxygen inhib parameter CBOD oxidation	20%	0.24%	0.50%	0.25%	0.22%
OrgN Settling velocity	20%	0.61%	0.46%	0.47%	0.39%
Periphyton Death rate	20%	0.43%	0.46%	0.35%	0.50%
HTS Death rate	20%	0.65%	0.32%	0.40%	0.68%
Reaeration user model parameter A	20%	0.59%	0.29%	0.07%	0.34%
OrgP Hydrolysis	20%	0.60%	0.28%	0.35%	0.51%
OrgN Hydrolysis	20%	0.30%	0.16%	0.36%	0.23%
Nitrification	20%	0.43%	0.04%	0.15%	0.09%
CBODfast Oxidation rate	20%	0.11%	0.04%	0.08%	0.03%
Periphyton Max Growth rate	20%	0.70%	0.02%	0.11%	0.17%

4.4. Multi-objective Waste-load Allocation Model

The ongoing effort to address the pollution problem of LAR requires a combination of regulatory and technical tools. In this regard, this study aims to show usefulness of SO based WLA models for the WQ management of LAR. Especially, such models assist environmental agencies in the WQ decision making process. Application of similar models to realistic systems was observed in other papers (Yandamuri *et al.* 2006; Cho & Lee 2014). The proposed two models, in this study, are cost-performance (CP) and cost-equity-performance (CEP) models.

4.5. Cost-performance model

This model takes cost of waste treatment and performance measures as objective function and constraint respectively. In this aspect, implementation of the Q2Kw - GA model retains those runs with the least possible treatment cost and least violations of specified standards. Violation of standards is expressed as number of violations, magnitude of total violations and magnitude of maximum violations at the WQ checkpoints.

The Pareto-optimal front describing the cost-performance trade-off for DO, BOD, TN and TP are presented in **Figure 4-9 (a) - (d)**. In these trade-off curves, the two extreme points stand for the Maximum Performance (MP) and the Least Cost (LC) pareto-optimal solutions. The midpoint represents compromise (CM) solution. This approach was observed in other reports (Yandamuri *et al.* 2006).

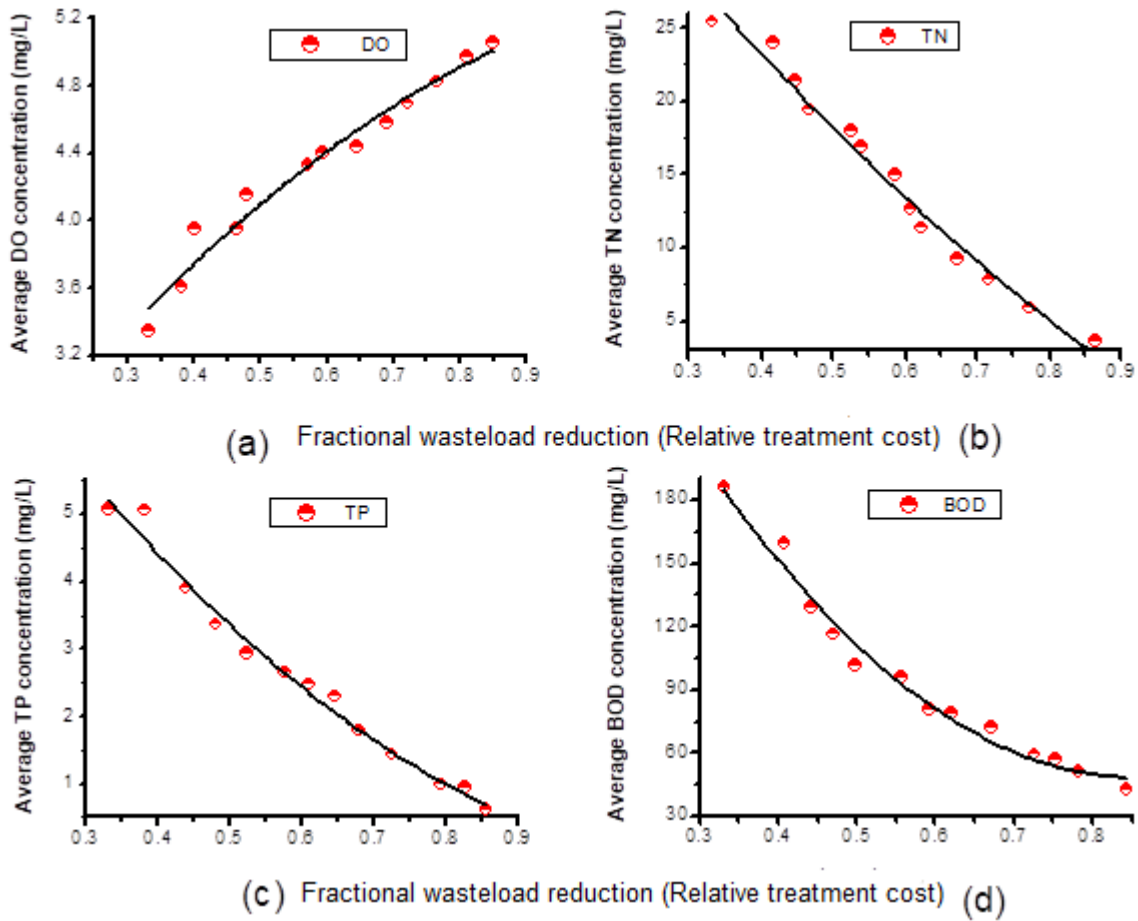


Figure 4-9 Cost-performance trade-off graphs

The decision maker at the relevant authority can make decisions based on DO, BOD, TN, TP performances and corresponding treatment cost. For example, a desired DO performance increase from 3.6 mg/L to 5 mg/L on average requires 50% WL reduction (treatment cost equivalent). Furthermore, the decision process can look into treatment cost effects as a result of better DO, BOD, TN and TP performances. As such, the trade-off curves can help environmental decision making under conditions of budget constraints and desired performance measures.

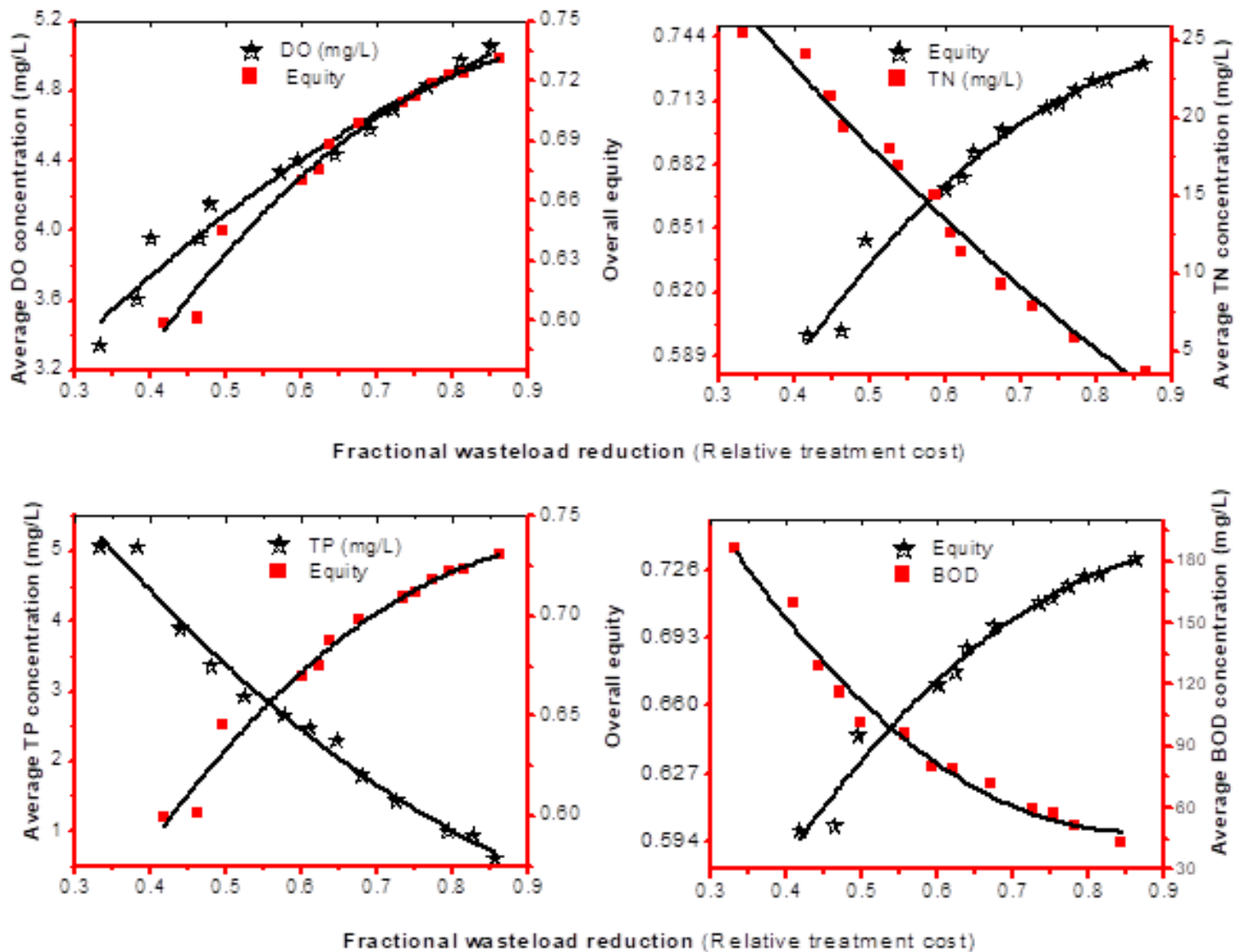


Figure 4-10 Cost-equity-performance trade-off

3.2 Cost-equity-performance (CEP) model

This model considers cost and equity as objective functions while performance measures is taken as constraint. **Figure 4-10** shows selected representations of CEP model. In these graphs, various combinations of treatment cost, performance measures and equity measures are represented for DO, BOD, TN and TP. These trade-off curves allow the decision maker to arrive at optimal solutions without seriously affecting interests of dischargers. In this way, fairness among dischargers can be ensured, in addition to improvement in WQ and reduction in treatment cost.

A desired 5 mg/L average oxygen concentration requires about 87% *WL* reduction with 0.73 equity measure. However, in the case of industries, cost burden for waste treatment is usually worrisome especially those under budget constraint condition. This will lead the decision maker to search for a compromised solution. For instance, a decision to lower this *WL* by 40% results in a 3.6 mg DO/L and that violates the standard. In such a way, the approach can be used to select a compromise solution using the trade-off curves. Moreover, the environmental agencies can make their analyses and decisions based on the remaining parameters such as TP, BOD and TN.

Furthermore, effect of *WL* reduction on the overall environmental performance can be considered. In this regard, overall WQ improvement due to concentration reduction of all parameters is considered. This analysis is illustrated in **Figure 4-11**. For example, a 0.5 *WL* reduction results in an overall performance of 5.4. The same *WL* reduction results in 4.1 mg/L of DO. Environmental decision based on this approach can be improved by examining performance of the separate parameters for the desired overall performance.

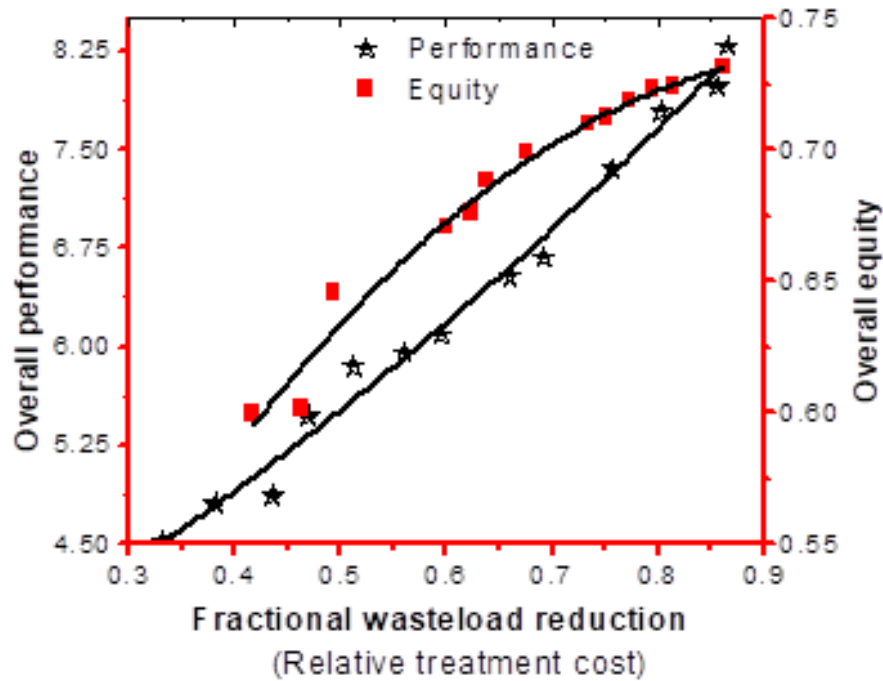


Figure 4-11 Equity and performance versus *WL* reduction

4.5.1. Model's Applicability to the Local Condition

Similar to suggestion for developing countries (Hranova 2006a), the practical implementation of the WLA modeling in Ethiopia for the purpose of River management can be considered for impaired large rivers (i.e., rivers that are not based on effluent discharge). However, due to extensive data requirement (that is collected on regular monitoring programs) of these approaches, its consideration as a management tool is not possible in the near future. Data base organization and capacity building aimed at future application of models as management tools is important for improving WQ of impaired waters in Ethiopia .

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

The data from LAR was subjected to CA, DA, PCA and FA to recognize spatial pollution pattern, identify discriminating parameters, reduce data dimensionality and investigate the latent factors. The general grouping of 27 sampling sites for LAR based on Ward's method came up with three clusters: Cluster I grouped together relatively least polluted water from the upstream and a major tributary to LAR having its source from Gefersa Dam. Cluster II contains nearly 63% of the sampling sites with a relatively medium pollution level, and Cluster III groups three streams coming from the heavily inhabited areas with high level of commercial activity in the AK, L and K sub-cities. The DA resulted in 85.2% correct assignation with two parameters: T-P and DO. Although PCA/FA data dimensionality reduction was not significant, from the original parameters of sixteen to ten with 81.9% of data variability explained, the methods revealed the situation in LAR pollution patterns. Strong PCA/FA loadings for ten parameters such as TSS, T-NH₃, T-N, T-P, PO₄, BOD, NO₃-N, and NO₂ can be mainly due to sewage and detergent discharge from domestic, beverage and food processing industries and agricultural runoff from the peri-urban section of LAR watershed. The odor from the river was due to T-NH₃, which was released by the disintegration of organic matter. Moreover, FA identified domestic and industrial waste discharges as the common causes for pollution problems of LAR. The results obtained from the multivariate analysis indicate that CA, DA, PCA and FA can be used to systematically assess the WQ variation of LAR economically with limited resource and effort. To this end,

three sampling sites, instead of twenty-seven, one from each cluster, and ten significant parameters, instead of fifteen, can be employed to characterize LAR WQ.

WQI was determined based on CCME-WQI method using local discharge standards and data collected during 2015 (November/October) and 2017 (January/February). The result obtained summarizes LAR WQ as generally *poor*. This rank, according to CCME_WQI interpretation, categorizes LAR as impaired. Considering lengthy calculation requirement of WQI determination, a validated ANN model was applied successfully for WQI determination of LAR. ANN optimal model resulted in acceptable prediction with RMSE, MSE and R^2 values of 1.97, 0.07 and 0.92 respectively. Therefore, it is concluded that CCME-WQI can be used to summarize LAR WQ status in a way that is easy to understand for policy makers and the public. In addition, to minimize time and effort burden of repeated WQI determination, a modelling approach based on ANN (parametric model) can be employed successfully for acceptable results.

In this study, Q2Kw was calibrated, validated and a sensitivity analysis was done for LAR based on simulated parameters such as: DO, BOD, TN and TP. During sensitivity analysis, 20% increase in values of parameters such as point source flow and Manning's n did not result in significant WQ improvement of LAR. This indicates that a higher degree of effort is required for better WQ result. The R^2 values of DO, BOD, TN and TP obtained after calibration: 0.91, 0.90, 0.81 and 0.89; and those after validation: 0.98, 0.82, 0.77 and 0.75, respectively indicate that the performance of Q2Kw model is satisfactory and the model can be applied successfully to generate future scenarios for the study area.

The results obtained from the application of optimal MO-WLA models on LAR demonstrate the importance SO models for a sustainable WQ decision making. The factors considered in these proposed models (i.e., cost-performance and cost-equity-performance) such

as treatment cost, equity and performance measures consider economic, social and environmental concerns. Thus, this feature is useful in addressing pollution problems fairly in various land use conditions and under budget constraint. Particularly, the cost-performance model outputs show the need to consider budget constraints in specifying desired performance measure. This model fails to consider equity as a factor. Hence, it may not be appropriate for conditions with uneven discharge from polluters. On the other hand, cost-equity-performance model can be used for such cases. The trade-off curves from the cost-equity-performance model, for a specified WQ performance, treatment cost increases with increased equity among dischargers. Here, the trade-off curves present the environmental agency with the opportunity of finding compromised solutions. Therefore, the approach is useful for sustainable WQ decision making.

5.2. Significance of the Work for Rehabilitation of the LAR

The most important issues related to the Little Akaki River are deteriorating water quality, poor water quality monitoring and control, indiscriminate domestic and industrial waste disposal, inactive public participation and low awareness. Rehabilitation works aimed at improving the water quality condition of the River and its surrounding require approaches that can cope up with the complexities and shifting pollution conditions. To this end, this work can be used: to increase public participation through information about the River condition using WQI, and to achieve economical and regular monitoring activity using CA and PCA. Moreover, this work can be helpful in decision making during water quality assessment and control schemes using QUAL2Kw-Genetic Algorithm based simulation-optimization model.

5.3. Recommendations for Future Work

In this study, the recommendations for future work are:

- WQ assessment of sewer networks for pollution sources identification and solution proposal.
- Assessment of other regulatory approaches other than the existing discharged based approach

6. REFERENCES

- Abbasi T. and Abbasi S. A. (2012). *Water quality indices*. Elsevier, Amsterdam.
- Abdul-Wahab S. A., Bakheit C. S. and Al-Alawi S. M. (2005). Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environ. Model. Softw.* **20**(10), 1263-71.
- Abiy A. Z., Demissie S. S., MacAlister C., Dessu S. B. and Melesse A. M. (2016). Groundwater Recharge and Contribution to the Tana Sub-basin, Upper Blue Nile Basin, Ethiopia. In: *Landscape Dynamics, Soils and Hydrological Processes in Varied Climates*, Springer, pp. 463-81.
- Adgolign T. B., Rao G. S. and Abbulu Y. (2016). WEAP modeling of surface water resources allocation in Didessa Sub-Basin, West Ethiopia. *Sustainable Water Resources Management* **2**(1), 55-70.
- Afifi A. A. and Clark V. (1997). *Computer aided multivariate analysis*. Chapman & Halucrc, Boca Raton.
- Ahammed S. S., Tasfina S., Rabbani K. A. and Khaleque M. A. (2016). An investigation into the water quality of Buriganga-A river running through Dhaka. *International Journal of Scientific & Technology Research* **5**(3), 36-41.
- Alberto W. D., del Pilar D. a. M. a., Valeria A. M. a., Fabiana P. S., Cecilia H. A. and de los Ángeles B. M. a. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality- A case study: Suquía River Basin (Córdoba–Argentina). *Water Res* **35**(12), 2881-94.
- Alemayehu T., Sulieman H. and Ayalew M. (2009). Use of treated wastewater for managed aquifer recharge in highly populated urban centers: a case study in Addis Ababa, Ethiopia. *Environ Geol* **58**(1), 55-9.
- Allam A., Tawfik A., Yoshimura C. and Fleifle A. (2016). Multi-objective models of waste load allocation toward a sustainable reuse of drainage water in irrigation. *Environmental Science & Pollution Research* **23**(12), 11823-34.
- Ambrose Jr R. B., Wool T. A. and Barnwell Jr T. O. (2009). Development of Water Quality Modeling in the United States. *Environmental Engineering Research* **14**(4), 200-10.
- Andrade L. N. d., Mauri G. R. and Mendonça A. S. F. (2013). General Multiobjective Model and Simulated Annealing Algorithm for Waste-Load Allocation. *J. Water Resour. Plann. Manage.* **139**, 339-44.
- APHA, AWWA and WEF (1998). *Standard methods for the examination of water and wastewater*. American Public Health Association (APHA), American Water Works Association (AWWA), Water Environment Federation (WEF), Washington, DC.
- Aral M. (2010). *Environmental modeling and health risk analysis (Acts/Risk)*. Springer Science & Business Media, Dordrecht.
- Armah F. A., Obiri S., Yawson D. O., Onumah E. E., Yengoh G. T., Afrifa E. K. and Odoi J. O. (2010). Anthropogenic sources and environmentally relevant concentrations of heavy

- metals in surface water of a mining district in Ghana: a multivariate statistical approach. *J Environ Sci Health A* **45**(13), 1804-13.
- Arsiso B. K., Tsidu G. M., Stoffberg G. H. and Tadesse T. (2018). Influence of urbanization-driven land use/cover change on climate: The case of Addis Ababa, Ethiopia. *Physics and Chemistry of the Earth, Parts A/B/C* **105**, 212-23.
- Aruga R., Gastaldi D., Negro G. and Ostacoli G. (1995). Pollution of a river basin and its evolution with time studied by multivariate statistical analysis. *Analytica Chimica Acta* **310**(1), 15-25.
- Aschale M., Sileshi, Y., Kelly-Quinn, M., & Hailu, D. (2015). Potentially toxic trace element contamination of the Little Akaki River of Addis Ababa, Ethiopia. *J Nat Sci Res* **5**(1), 1-13.
- Asit K. B., Cecilia T., Benedito B. and Diego J. R. (2006). *The Singapore Water Story: Sustainable Development in an Urban City-state*. Springer Science & Business Media.
- Ayele G. T., Teshale E. Z., Yu B., Rutherford I. D. and Jeong J. (2017). Streamflow and sediment yield prediction for watershed prioritization in the Upper Blue Nile River Basin, Ethiopia. *Water* **9**(10), 782.
- Ayewew T. (2001). Numerical groundwater flow modeling of the Central Main Ethiopian Rift lakes basin. *SINET: Ethiopian Journal of Science* **24**(2), 167-84.
- Ayewew T., Demlie M. and Wohnlich S. (2008). Application of numerical modeling for groundwater flow system analysis in the Akaki catchment, central Ethiopia. *Math Geosci* **40**(8), 887-906.
- Ayeni A. and Soneye A. (2013). Interpretation of surface water quality using principal components analysis and cluster analysis. *J Geogr Reg Plan* **6**(4), 132-41.
- Bahadur R., Amstutz D. E. and Samuels W. B. (2013). Water contamination modeling—a review of the state of the science. *Journal of Water Resource and Protection* **5**(02), 142.
- Barańkiewicz D., Chudzińska M., Szpakowska B., Świerk D., Gołdyn R. and Dondajewska R. (2014). Storm water contamination and its effect on the quality of urban surface waters. *Environ Monit Assess* **186**(10), 6789-803.
- Barzegar R. and Moghaddam A. A. (2016). Combining the advantages of neural networks using the concept of committee machine in the groundwater salinity prediction. *Model. Earth Syst. Environ.* **2**, 26.
- Beles E. (2014). *Environmental and social impact assessment of the wastewater treatment plant and sewer lines expansion and rehabilitation in the kaliti catchment*. Addis Ababa Water and Sewerage Authority (AAWSA), Addis Ababa, Ethiopia.
- Benedini M. and George T. (2013). *Water quality modelling for rivers and streams*. Dordrecht: Springer.
- Berhanu B., Seleshi Y. and Melesse M. A. (2014). Surface Water and Groundwater Resources of Ethiopia: Potentials and Challenges of Water Resources Development. In: *Nile River Basin: Ecohydrological Challenges, Climate Change and Hydropolitics* Assefa M, Wossenu A and Shimelis G (eds), Springer.

- Berhe F., Melesse A., Hailu D. and Sileshi Y. (2013). MODSIM-based water allocation modeling of Awash River Basin, Ethiopia. *CATENA* **109**, 118-28.
- Betrie G., Van Griensven A., Mohamed Y., Popescu I., Mynett A. and Hummel S. (2011). Linking SWAT and SOBEK using open modeling interface (OPENMI) for sediment transport simulation in the Blue Nile River basin. *Transactions of the ASABE* **54**(5), 1749-57.
- Bhuiyan M. A., Rakib M., Dampare S., Ganyaglo S. and Suzuki S. (2011). Surface water quality assessment in the central part of Bangladesh using multivariate analysis. *KSCE J Civil Engineer* **15**(6), 995-1003.
- Bilgin A. and Konanç M. U. (2016). Evaluation of surface water quality and heavy metal pollution of Coruh River Basin (Turkey) by multivariate statistical methods. *Environ Earth Sci* **75**(12), 1028–46.
- Bottino F., Ferraz I. C., Mendiondo E. M. and Calijuri M. d. C. (2010). Calibration of QUAL2K model in Brazilian micro watershed: effects of the land use on water quality. *Acta Limnologica Brasiliensia* **22**(4), 474-85.
- Bouguerne A., Boudoukha A., Benkhaled A. and Mebarkia A.-H. (2017). Assessment of surface water quality of Ain Zada dam (Algeria) using multivariate statistical techniques. *Intl. J. River Basin Management* **15**(2), 133-43.
- Brion G. M. and Lingireddy S. (2000). Identification of pollution sources via neural networks. In: *Artificial neural networks in hydrology* Govindaraju RS and Rao AR (eds), Springer, The Netherlands.
- Brown L. C. and Barnwell T. O. (1987). *The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: Documentation and user manual*. Athens, Georgia: US Environmental Protection Agency, Office of Research and Development. Environmental Research Laboratory.
- Brown R. M., McClelland N. I., Deininger R. A. and O'Connor M. F. (1972). A water quality index-crashing the psychological barrier. In: *Indicators of environmental quality* Thomas WA (ed.), Springer, Boston.
- Brūmelis G., Lapiņa L., Nikodemus O. and Tabors G. (2000). Use of an artificial model of monitoring data to aid interpretation of principal component analysis. *Environ Model Softw* **15**, 755-63.
- Burn D. H. (1987). Alternative formulations for water quality management models. In: *Engineering hydrology* Feldman AD (ed.), ASCE, New York, N.Y.
- Burn D. H. and Lence B. J. (1992). Comparison of optimization formulations for waste-load allocations. *J Environ Eng.* **118**(4), 597-612.
- Burn D. H. and Yulianti J. S. (2001). Waste-load allocation using genetic algorithms. *Journal of Water Resources Planning and Management* **127**(2), 121-9.
- Caduto MJ, (1990). *Pond and Brook: a guide to nature in freshwater environments*. UPNE.
- Camargo R. d. A., Calijuri M. L., Santiago A. d. F. and Couto E. d. A. d. (2010). Water quality prediction using the QUAL2Kw model in a small karstic watershed in Brazil. *Acta Limnologica Brasiliensia* **22**(4), 486-98.

- Carmichael J. J. and Strzepek K. M. (2000). A multiple-organic-pollutant simulation/optimization model of industrial and municipal wastewater loading to a riverine environment. *Water Resources Research* **36**(5), 1325-32.
- Carpenter S. R., Caraco N. F., Correll D. L., Howarth R. W., Sharpley A. N. and Smith V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. App.* **8**(3), 559-68.
- CCME (2001). *Canadian water quality guidelines for the protection of aquatic life: CCME Water Quality Index 1.0, User's Manual*. The Canadian Council of Ministers of the Environment (CCME), Winnipeg, Canada.
- Chaemiso S. E., Abebe A. and Pingale S. M. (2016). Assessment of the impact of climate change on surface hydrological processes using SWAT: a case study of Omo-Gibe river basin, Ethiopia. *Modeling Earth Systems and Environment* **2**(4), 205.
- Chapman D. (1992a). *Water quality assessments*. Chapman & Hall, London.
- Chapman D. (1992b). *Water quality assessments: a guide to use of biota, sediments and water in environmental monitoring* Chapman & Hall, Great Britain.
- Chapra S., Pelletier G. and Tao H. (2007). A modeling Framework for Simulating River and Stream Water Quality. In: USEPA (ed.), GA, USA.
- Charbonneau P. and Knapp B. (1995). *A user's guide to PIKAIA 1.0*. National Center for Atmospheric Research.
- Chebud Y. A. and Melesse A. M. (2009). Numerical modeling of the groundwater flow system of the Gumera sub-basin in Lake Tana basin, Ethiopia. *Hydrological Processes: An International Journal* **23**(26), 3694-704.
- Chen Q., Mei K., Dahlgren R. A., Wang T., Gong J. and Zhang M. (2016). Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression. *Sci Total Environ* **572**, 450-66.
- Cho J., Ahn K., Chung W. and Gwon E. (2003). Waste load allocation for water quality management of a heavily polluted river using linear programming. *Water Science and Technology* **48**(10), 185-90.
- Cho J. H. and Lee J. H. (2014). Multi-objective waste load allocation model for optimizing waste load abatement and inequality among waste dischargers. *Water Air & Soil Pollution* **225**(3), 1892.
- Chun M. H., Sulaiman, W. N. A., & Samah, M. A. A. (2012). A Case Study on Public Participation for the Conservation of a Tropical Urban River. *Polish Journal of Environmental Studies* **21**(4).
- Close M., Abraham P., Humphries B., Lilburne L., Cuthill T. and Wilson S. (2016). Predicting groundwater redox status on a regional scale using linear discriminant analysis. *J Contam Hydrol* **191**, 19-32.
- Cohon J. L. (1978a). *Multiobjective programming and planning*. Academic Press, New York.
- Cohon J. L. (1978b). *Multiobjective programming and planning*. Academic, New York.

- Cox B. (2003). A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers. *Science of the Total Environment* **314**, 335-77.
- CSA (2013). Population projection of Ethiopia for all regions at wereda level from 2014 – 2017. In, Central Statistical Agency (CSA), Addis Ababa, Ethiopia.
- Cude C. G. (2001). Oregon water quality index: a tool for evaluating water quality management effectiveness. *J. Am. Water. Resour. Assoc.* **37**(1), 125-37.
- Dargahi B. and Setegn S. G. (2011). Combined 3D hydrodynamic and watershed modelling of Lake Tana, Ethiopia. *Journal of Hydrology* **398**(1-2), 44-64.
- Das S., Majumder M., Roy D. and Mazumdar A. (2010). Determination of urbanization impact on rain water quality with the help of water quality index and urbanization index. In: *Impact of Climate Change on Natural Resource Management* Jana BK and Majumde M (eds), Springer, London.
- Davis L. (1991). *Handbook of genetic algorithms*. Van Nostrand Reinhold, New York.
- Dawson C. and Wilby R. (2001). Hydrological modelling using artificial neural networks. *Prog Phys Geogr* **25**(1), 80-108.
- de Andrade L. N., Mauri G. R. and F. Mendonça A. S. (2013). General multiobjective model and simulated annealing algorithm for waste-load allocation. *Journal of Water Resources Planning & Management* **139**, 339-44.
- Deksissa T., Meirlaen J., Ashton P. J. and Vanrolleghem P. A. (2004). Simplifying dynamic river water quality modelling: A case study of inorganic nitrogen dynamics in the Crocodile River (South Africa). *Water, Air, and Soil Pollution* **155**(1-4), 303-20.
- Demuth H. and Beale M. (1993). *Neural network toolbox for use with Matlab: user's guide verion 3.0*. The MathWorks, Inc, Massachusetts, USA.
- Deo M. C. and Thirumalaiah K. (2000). Real time forecasting using neural networks. In: *Artificial Neural Networks in Hydrology* Govindaraju RS and Rao AR (eds), Springer, Dordrecht.
- Desta H. and Lemma B. (2017). SWAT based hydrological assessment and characterization of Lake Ziway sub-watersheds, Ethiopia. *Journal of Hydrology: Regional Studies* **13**, 122-37.
- Di Toro D. M., Fitzpatrick J. J. and Thomann R. V. (1983). Documentation for water quality analysis simulation program (WASP) and model verification program (MVP).
- Dile Y. T., Berndtsson R. and Setegn S. G. (2013). Hydrological response to climate change for gilgel abay river, in the lake tana basin-upper blue Nile basin of Ethiopia. *PloS one* **8**(10), e79296.
- DOE (2007). *Malaysia environmental quality report 2006*. Sasyaz Holdings Sdn Bh, Petaling Jaya, Malaysia.
- Dogan E., Sengorur B. and Koklu R. (2009). Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. *J Environ Manag* **90**(2), 1229-35.

- Douglas-Mankin K., Srinivasan R. and Arnold J. (2010). Soil and Water Assessment Tool (SWAT) model: Current developments and applications. *Transactions of the ASABE* **53**(5), 1423-31.
- Du P., Li Y. P. and Huang G. H. (2013). Inexact chance-constrained waste-load allocation model for water quality management of Xiangxihe River. *Journal of Environmental Engineering* **139**, 1178-97.
- Dunteman G. H. (1989). *Principal components analysis*. Sage, Newberry Park.
- Ehteshami M., Farahani N. D. and Tavassoli S. (2016). Simulation of nitrate contamination in groundwater using artificial neural networks. *Model. Earth Syst. Environ.* **2**, 28.
- EPA and UNIDO (2003). Guideline ambient environment standards for Ethiopia. In, Environmental Protection Authority of Ethiopia (EPA), United Nations Industrial Development Organization (UNIDO), Addis Ababa, Ethiopia.
- EU (2000). *EU Water Framework Directive*, European Union (EU).
- FAO (1976). *Water quality for agriculture*. Food and Agriculture Organization of the United Nations (FAO), Rome.
- Fasil D., Aschalew L., Yared T. and Kibru T. (2013). The Water Quality Degradation Of Upper Awash River, Ethiopia. *Ethiopian Journal of Environmental Studies and Management* **6**(1).
- FDREMT (2011). *Transport policy of Addis Ababa*. Federal Democratic Republic of Ethiopia Ministry of Transport (FDREMT), Addis Ababa, Ethiopia.
- Feven S. (2007). *Spatial and temporal water quality trend analysis using sediment cores and water samples from Aba Samuel Lake*. master's thesis, earth sciences department, Addis Ababa University, Addis Ababa.
- Florescu D., Ionete R., Sandru C., Iordache A. and Culea M. (2011). The influence of pollution monitoring parameters in characterizing the surface water quality from Romania southern area. *Rom. Journ. Phys* **56**, 1001-10.
- Foias C., Manley O., Rosa R. and Temam R. (2001). *Navier-Stokes equations and turbulence*. Cambridge University Press, United Kingdom.
- Gao L. and Li D. (2014). A review of hydrological/water-quality models. *Frontiers of Agricultural Science and Engineering* **1**(4), 267.
- Gazzaz N. M., Yusoff M. K., Aris A. Z., Juahir H. and Ramli M. F. (2012). Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar Pollut Bull* **64**(11), 2409-20.
- Gebreyohannes T., De Smedt F., Walraevens K., Gebresilassie S., Hussien A., Hagos M., Amare K., Deckers J. and Gebrehiwot K. (2017). Regional groundwater flow modeling of the Geba basin, northern Ethiopia. *Hydrogeology Journal* **25**(3), 639-55.
- Georgilakis P. S. (2009). *Spotlight on modern transformer design*. Springer, London.
- Gerbens-Leenes P., Hoekstra A. Y. and Van der Meer T. (2009). The water footprint of energy from biomass: A quantitative assessment and consequences of an increasing share of bio-energy in energy supply. *Ecological economics* **68**(4), 1052-60.

- Ghenaïet A. and Tayeb B. (2009). Optimum design of turboprop engines using genetic algorithm. *Journal of Propulsion and Power* **25**(6), 1345-50.
- Gholizadeh M. H., Melesse A. M. and Reddi L. (2016). Discriminant analysis application in spatiotemporal evaluation of water quality in South Florida. *J. Hydroinform* **18**(6), 1019-32.
- Gikas G. D. (2014). Water quality of drainage canals and assessment of nutrient loads using QUAL2Kw. *Environmental Processes* **1**(4), 369-85.
- Gizachew K. (2008). *Gis and remote sensing based solid waste landfill site selection: A case of Addis Ababa city, Ethiopia*. Thesis, Addis Ababa University, Addis Ababa.
- Govindaraju R. S. and Rao A. R. (2013). *Artificial neural networks in hydrology*. Springer, Amsterdam.
- Gu R. and Dong M. (1998). Water quality modeling in the watershed-based approach for waste load allocations. *Water Science and Technology* **38**(10), 165-72.
- Gumindoga W., Rientjes T., Haile A. and Dube T. (2014). Predicting streamflow for land cover changes in the Upper Gilgel Abay River Basin, Ethiopia: A TOPMODEL based approach. *Physics and Chemistry of the Earth, Parts A/B/C* **76**, 3-15.
- Hailemariam A. (2017). The second biggest african country undergoing rapid change: Ethiopia. In: *Africa's population: in search of a demographic dividend* Groth H and May J (eds), Springer, Cham.
- Hair J. F., Anderson R. E., Tatham R. L. and Black W. C. (1995). *Multivariate data analysis*. Macmillan, New York.
- Hajjigholizadeh M. and Melesse A. M. (2017). Assortment and spatiotemporal analysis of surface water quality using cluster and discriminant analyses. *CATENA* **151**, 247-58.
- Heddam S. (2016). New modelling strategy based on radial basis function neural network (RBFNN) for predicting dissolved oxygen concentration using the components of the Gregorian calendar as inputs: case study of Clackamas River, Oregon, USA. *Model. Earth Syst. Environ.* **2**, 167.
- Helena B., Pardo R., Vega M., Barrado E., Fernandez J. M. and Fernandez L. (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res* **34**(3), 807-16.
- Hinkelmann R. (2006). *Efficient numerical methods and information-processing techniques for modeling hydro-and environmental systems*. Springer Science & Business Media, Germany.
- Hobson A., Neilson B., von Stackelberg N., Shupryt M., Ostermiller J., Pelletier G. and Chapra S. (2014). Development of a minimalistic data collection strategy for qual2kw. *Journal of Water Resources Planning and Management* **141**(8), 04014096.
- Holguin-Gonzalez J. E. (2013). *Integrated ecological modelling for decision support in river management*. PhD thesis, Ghent University, Ghent, Belgium.
- Horton R. K. (1965). An index number system for rating water quality. *Journal of Water Pollution Control Federation* **37**(3), 300-6.

- Hossain M., Sujaul I. and Nasly M. (2014). Application of QUAL2Kw for water quality modeling in the Tunggak River, Kuantan, Pahang, Malaysia. *Research Journal of Recent Sciences* **3**(6), 6-14.
- Hranova R. (2006a). DIFFUSE POLLUTION – PRINCIPLES, DEFINITIONS AND REGULATORY ASPECTS. In: *Diffuse Pollution of Water Resources: Principles and Case Studies in the Southern African Region* Hranova R (ed.), Taylor & Francis, London, UK.
- Hranova R. (2006b). *Diffuse Pollution of Water Resources: Principles and Case Studies in the Southern African Region*. CRC Press.
- Inoue T. and Ebise S. (1991). Runoff characteristics of COD, BOD, C, N and P loadings from rivers to enclosed coastal seas. *Mar Pollut Bull* **23**, 11-4.
- Javan K., Lialestani M. R. F. H. and Nejadhossein M. (2015). A comparison of ANN and HSPF models for runoff simulation in Gharehsoo River watershed, Iran. *Model. Earth Syst. Environ.* **1**(4), 41.
- Jayawardena A. W. (2014). *Environmental and Hydrological Systems Modelling*. CRC Press.
- Ji Z.-G. (2008). *Hydrodynamics and water quality: Modeling rivers, lakes, and estuaries*. John Wiley & Sons, New Jersey.
- Jolliffe I. (1986). *Principal component analysis*. Springer, New York.
- Juahir H., Zain S. M., Toriman M. E., Mokhtar M. and Man H. C. (2004). Application of artificial neural network models for predicting water quality index. *Malaysian Journal of Civil Engineering* **16**(2), 42-55.
- Juahir H., Zain S. M., Yusoff M. K., Hanidza T. I. T., Armi A. S. M., Toriman M. E. and Mokhtar M. (2011). Spatial water quality assessment of Langat River Basin (Malaysia) using environmetric techniques. *Environ Monit Assess* **173**(1), 625-41.
- Jung K. Y., Lee K.-L., Im T. H., Lee I. J., Kim S., Han K.-Y. and Ahn J. M. (2016). Evaluation of water quality for the Nakdong River watershed using multivariate analysis. *Environ Technol Innov* **5**, 67-82.
- Kamruzzaman J., Begg R. K. and Sarker R. A. (2006). *Artificial neural networks in finance and manufacturing*. Idea Group Inc., United States of America.
- Kannel P. R., Kanel S. R., Lee S., Lee Y.-S. and Gan T. Y. (2011). A review of public domain water quality models for simulating dissolved oxygen in rivers and streams. *Environmental Modeling & Assessment* **16**(2), 183-204.
- Kannel P. R., Lee S., Kanel S. R., Lee Y.-S. and Ahn K.-H. (2007a). Application of QUAL2Kw for water quality modeling and dissolved oxygen control in the river Bagmati. *Environmental Monitoring & Assessment* **125**(1), 201-17.
- Kannel P. R., Lee S. and Lee Y.-S. (2008). Assessment of spatial–temporal patterns of surface and ground water qualities and factors influencing management strategy of groundwater system in an urban river corridor of Nepal. *Journal of Environmental Management* **86**(4), 595-604.

- Kannel P. R., Lee S., Lee Y.-S., Kanel S. and Pelletier G. (2007b). Application of automated QUAL2Kw for water quality modeling and management in the Bagmati River, Nepal. *Ecological modelling* **202**(3), 503-17.
- Karunanithi N., Grenney W. J., Whitley D. and Bovee K. (1994). Neural networks for river flow prediction. *J Comput Civ Eng* **8**(2), 201-20.
- Khan M. Y. A., Gani K. M. and Chakrapani G. J. (2015). Assessment of surface water quality and its spatial variation: A case study of Ramganga River, Ganga Basin, India. *Arab J Geosci* **9**, 28.
- Khan M. Y. A., Gani K. M. and Chakrapani G. J. (2017). Spatial and temporal variations of physicochemical and heavy metal pollution in Ramganga River—a tributary of River Ganges, India. *J. Environ Earth Sci* **76**(5), 231.
- Khuan L. Y., Hamzah N. and Jailani R. (2002). Prediction of water quality index (WQI) based on artificial neural network (ANN). In: *Student conference on research and development*, Shah Alam, Malaysia
- Kim J. and Mueller C. W. (1978). *Introduction to factor analysis*. Sage Publications, Beverly Hills.
- Kim M., Kim Y., Kim H., Piao W. and Kim C. (2015). Enhanced monitoring of water quality variation in Nakdong River downstream using multivariate statistical techniques. *Desalin Water Treat* **57**(27), 12508-17.
- Kivaisi A. K. (2001). The potential for constructed wetlands for wastewater treatment and reuse in developing countries: a review. *Ecological engineering* **16**(4), 545-60.
- Kucuksezgin F. (1996). Multivariate analysis of water quality parameters in Izmir Bay, Eastern Aegean. *Toxicol Environ Chem* **55**, 135-44.
- Kühnert C., Bernard T., Arango I. M. and Nitsche R. (2014). Water quality supervision of distribution networks based on machine learning algorithms and operator feedback. *Procedia Engineering* **89**, 189-96.
- Kumar A., Bisht B., Joshi V., Singh A. and Talwar A. (2010). Physical, chemical and bacteriological study of water from rivers of Uttarakhand. *J Hum Ecol* **32**(3), 169-73.
- Kundu S., Coumar M. V., Rajendiran S. and Rao A. S. (2015). Phosphates from detergents and eutrophication of surface water ecosystem in India. *Curr. Sci* **108**(7), 1320–5.
- Larose D. (2014). *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, New York.
- Lawson J. (2005). *River Basin Management: Progress Towards Implementation of the European Water Framework Directive*. CRC Press.
- Lek S., Delacoste M., Baran P., Dimopoulos I., Lauga J. and Aulagnier S. (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecol Model* **90**(1), 39-52.
- LI Y., XU L. and LI S. (2009). Water quality analysis of the Songhua River Basin using multivariate techniques. *J Water Resour Prot* **1**(2), 110-21.

- Ling T.-Y., Soo C.-L., Liew J.-J., Nyanti L., Sim S.-F. and Grinang J. (2017). Application of multivariate statistical analysis in evaluation of surface river water quality of a tropical river. *J Chem* **13**.
- Liu C.-W., Lin K.-H. and Kuo Y.-M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Sci. Tot. Environ* **313**(1), 77-89.
- Lung W.-S. (1993). *Water Quality Modeling: Application to Estuaries*. CRC Press.
- Magoulès F. and Zhao H.-X. (2016). *Data mining and machine learning in building energy analysis*. John Wiley & Sons, Inc., Hoboken, USA.
- Manly B. F. J. (2004). *Multivariate Statistical Methods - A Primer*. Chapman & Hall/CRC, Boca Raton.
- Mannina G. (2010). Uncertainty assessment of a water-quality model for ephemeral rivers using GLUE analysis. *Journal of Environmental Engineering* **137**(3), 177-86.
- McKenna J. (2003). An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. *Environ Model Softw* **18**(3), 205-20.
- McKinney RE, (2004). *Environmental pollution control microbiology: a fifty-year perspective*. CRC Press.
- Mekonnen K. N., Ambushe A. A., Chandravanshi B. S., Abshiro M. R. and McCrindle R. I. (2012). Potentially toxic elements in some fresh water bodies in Ethiopia. *Toxicol. Environ. Chem.* **94**(10), 1980-94.
- Melaku S., Wondimu T., Dams R. and Moens L. (2007). Pollution status of Tinishu Akaki River and its tributaries (Ethiopia) evaluated using physico-chemical parameters, major ions, and nutrients. *Bull. Chem. Soc. Ethiop.* **21**(1), 13-22.
- Mengistu G. and Alemayehu S. (2013). Prevalence and burden of primary headache disorders among a local community in Addis Ababa, Ethiopia. *The journal of headache and pain* **14**(1), 30.
- Metcalf & Eddy (2005). *Wastewater Engineering Treatment and Reuse*. Tata McGraw-Hill Edition, New Delhi.
- Metcalf T. S. and Charbonneau P. (2003). Stellar structure modeling using a parallel genetic algorithm for objective global optimization. *Journal of Computational Physics* **185**(1), 176-93.
- Mohammadpour R., Shaharuddin S., Zakaria N. A., Ghani A. A., Vakili M. and Chan N. W. (2016). Prediction of water quality index in free surface constructed wetlands. *Environ Earth Sci* **75**(2), 139.
- Monica N. and Choi K. (2016). Temporal and spatial analysis of water quality in Saemangeum watershed using multivariate statistical techniques. *Paddy Water Environ* **14**(1), 3-17.
- Mostafavi S. A. and Afshar A. (2011). Waste load allocation using non-dominated archiving multi-colony ant algorithm. *Procedia Computer Science* **3**, 64-9.
- MoWIE (2015). *Urban wastewater management strategy*, Ministry of Water Irrigation and Electricity of Ethiopia (MoWIE), Addis Ababa, Ethiopia.

- MoWIE (2017). *Urban water supply & sanitation project phase II: Environmental and social management framework final*, Ministry of Water, Irrigation and Electricity (MoWIE), Addis Ababa, Ethiopia.
- Mujumdar P. P. and Subbarao V. V. (2004). Fuzzy waste load allocation model: simulation-optimization approach. *Journal of computing in civil engineering* **18**(2), 120-31.
- Murty Y., Bhallamudi S. M. and Srinivasan K. (2006). Non-uniform flow effect on optimal waste load allocation in rivers. *Water resources management* **20**(4), 509-30.
- Mvungi A., Hranova R. and Love D. (2003). Impact of home industries on water quality in a tributary of the Marimba River, Harare: implications for urban water management. *Physic. Chem. Earth* **28**(20), 1131-7.
- NEPA (1988). Environmental quality standards for surface water: GB 3838-88. In, *The National Standards of the People's Republic of China*.
- Nirmalakhandan N. (2002). *Modeling tools for environmental engineers and scientists*. CRC Press, Florida, United States of America.
- Nourani V., Baghanam A. H., Adamowski J. and Gebremichael M. (2013). Using self-organizing maps and wavelet transforms for space–time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. *J Hydrol* **476**, 228-43.
- O'Shea B. and Jankowski J. (2006). Detecting subtle hydrochemical anomalies with multivariate statistics: an example from 'homogeneous' groundwaters in the Great Artesian Basin, Australia. *Hydrol Process* **20**(20), 4317-33.
- OECD (2018). *OECD Studies on Water Managing the Water-Energy-Land-Food Nexus in Korea Policies and Governance Options: Policies and Governance Options*. OECD Publishing.
- Oliveira B., Bola J., Quinteiro P., Nadais H. and Arroja L. (2012). Application of Qual2Kw model as a tool for water quality management: Cértima River as a case study. *Environmental monitoring and assessment* **184**(10), 6197-210.
- Oljira E. (2006). *Numerical groundwater flow simulation of Akaki river catchment*. master's thesis, earth science department, Addis Ababa University, Addis Ababa, Ethiopia.
- Ongley E. D. and Booty W. G. (1999). Pollution remediation planning in developing countries: Conventional modelling versus knowledge-based prediction. *Water international* **24**(1), 31-8.
- Ouyang Y., Nkedi-Kizza P., Wu Q., Shinde D. and Huang C. (2006). Assessment of seasonal variations in surface water quality. *Water Res* **40**(20), 3800-10.
- Park N. (2010). *Advances in Geosciences: Hydrological Science*. World Scientific.
- Park S. S. and Lee Y. S. (2002). A water quality modeling study of the Nakdong River, Korea. *Ecological modelling* **152**(1), 65-75.
- Parsaie A. and Haghiabi A. H. (2015). Predicting the longitudinal dispersion coefficient by radial basis function neural network. *Model. Earth Syst. Environ.* **1**(4), 34.
- Pelletier G. and Chapra S. (2005). QUAL2Kw theory and documentation (version 5.1), a modeling framework for simulating river and stream water quality. *Washington: Department of ecology*.

- Pelletier G. and Chapra S. (2008a). *QUAL2Kw theory and documentation (version 5.1): a Modeling Framework for Simulating River and Stream Water Quality*. Olympia, Washington: Department of ecology.
- Pelletier G. and Chapra S. (2008b). *Qual2kw user manual (Version 5.1): A modeling framework for simulating river and stream water quality*. Olympia, Washington: Department of Ecology.
- Pelletier G. J., Chapra S. C. and Tao H. (2006). QUAL2Kw—A framework for modeling water quality in streams and rivers using a genetic algorithm for calibration. *Environmental Modelling & Software* **21**(3), 419-25.
- Pelletier G. J., Chapra, S.C. (2005). QUAL2Kw theory and documentation (version 5.1), A Modeling Framework for Simulating River and Stream Water Quality. In, Washington: Department of ecology.
- Pereira L. S., Ian C. and Iacovides I. (2009). *Coping with water scarcity*. Springer.
- Powley H. R., Dürr H. H., Lima A. T., Krom M. D. and Van Cappellen P. (2016). Direct discharges of domestic wastewater are a major source of phosphorus and nitrogen to the mediterranean sea. *Environ. Sci. Technol* **50**(16), 8722-30.
- Price R. K. and Vojinovi Z. (2011). *Urban hydroinformatics: data, models, and decision support for integrated urban water management*. IWA publishing.
- Qadir A., Malik R. N. and Husain S. Z. (2008). Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. *Environ Monit Assess* **140**(1), 43-59.
- Rao P. V., (2002). *Textbook of environmental engineering*. PHI Learning Pvt. Ltd.
- Reichert P., Borchardt D., Henze M., Rauch W., Shanahan P., Somlyódy L. and Vanrolleghem P. (2001). *River water quality model No. 1*. IWA Publishing, London, UK.
- Rencher A. C. (2002). *Methods of multivariate analysis*. Wiley, New York.
- Rowe D. R. and Abdel-Magid I. M., (1995). *Handbook of wastewater reclamation and reuse*. CRC press.
- Sabri N., Abdullah M., Mat S., Elfithri R. and Khalik W. (2016). Evaluation of hydrochemistry variation in water quality of Cempaka Lake, Malaysia using multivariate statistical analysis. *J. Mater. Environ. Sci.* **7**(12), 4403-10.
- Said A. M., Ahmadun F. I.-R., Paim L. H. and Masud J. (2003). Environmental concerns, knowledge and practices gap among Malaysian teachers. *Int J Sustain High Educ* **4**, 305-13.
- Sakizadeh M. (2016). Artificial intelligence for the prediction of water quality index in groundwater systems. *Model. Earth Syst. Environ.* **2**(1), 8.
- Santhi C., Srinivasan R., Arnold J. G. and Williams J. (2006). A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. *Environmental Modelling & Software* **21**(8), 1141-57.
- Saremi A., Sedghi M., Manshouri M. and Kave F. (2010). Development of multiobjective optimal model for Haraz River. *World Applied Sciences* **11**(8), 924-9.
- Schaffer J. D. (1985). *Multiple objective optimization with vector evaluated genetic algorithms*. In *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pp. 93–100.

- Schütze M., Butler D. and Beck B. M. (2011). *Modelling, simulation and control of urban wastewater systems*. Springer Science & Business Media, London.
- Setegn S. G., Dargahi B., Srinivasan R. and Melesse A. M. (2010). Modeling of Sediment Yield From Anjeni-Gauged Watershed, Ethiopia Using SWAT Model 1. *JAWRA Journal of the American Water Resources Association* **46**(3), 514-26.
- Shamshirband S., Gocić M., Petković D., Javidnia H., Ab Hamid S. H., Mansor Z. and Qasem S. N. (2015). Clustering project management for drought regions determination: a case study in Serbia. *Agricultural and forest meteorology* **200**, 57-65.
- Sharma D. and Kansal A. (2013). Assessment of river quality models: a review. *Reviews in Environmental Science and Bio/Technology* **12**(3), 285-311.
- Sharma D., Kansal A. and Pelletier G. (2015a). Water quality modeling for urban reach of Yamuna river, India (1999–2009), using QUAL2Kw. *Applied Water Science* **7**(3), 1535-59.
- Sharma N., Zakaullah M., Tiwari H. and Kumar D. (2015b). Runoff and sediment yield modeling using ANN and support vector machines: a case study from Nepal watershed. *Model. Earth Syst. Environ.* **1**, 23.
- Shrestha S. and Kazama F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environ. Model. Software* **22**(4), 464-75.
- Silveira A. d. (2002). Problems of modern urban drainage in developing countries. *Water Science and Technology* **45**(7), 31-40.
- Simonovic S. P. (2009a). *Managing Water Resources: Methods and Tools for a Systems Approach*. In, earthscan, London.
- Simonovic S. P. (2009b). *Managing Water Resources: Methods and Tools for a Systems Approach*. UNESCO Publishing, France.
- Singh K. P., Basant A., Malik A. and Jain G. (2009a). Artificial neural network modeling of the river water quality: a case study. *Ecol. Model.* **220**(6), 888-95.
- Singh K. P., Malik A., Mohan D. and Sinha S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Res* **38**(18), 3980-92.
- Singh K. P., Malik A. and Sinha S. (2005). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques—a case study. *Analytica Chimica Acta* **538**, 355-74.
- Singh S. K., Singh C. K., Kumar K. S., Gupta R. and Mukherjee S. (2009b). Spatial-temporal monitoring of groundwater using multivariate statistical techniques in Bareilly District of Uttar Pradesh, India. *J Hydrol Hydromech* **57**, 45-54.
- SINGH V. P. and FREVERT D. K. (2006). *WATERSHED MODELS*. Taylor & Francis, Boca Raton.
- Somlyódy L. and Varis O. (1992). *Water quality modeling of rivers and lakes*, International Institute of Applied Systems Analysis, Laxenburg Austria.

- Srinivas N. and Deb K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary computation* **2**(3), 221-48.
- Sterratt D., Graham B., Gillies A. and Willshaw D. (2011). *Principles of computational modelling in neuroscience*. Cambridge University Press New York.
- Streeter H. and Phelps E. (1925). A study of the pollution and natural purification of the Ohio River. *Public Health Bulletin* **146**.
- Susilowati Y., Mengko T., Rais J. and Leksono B. E. (2004). Water quality modeling for environmental information system. In: *Circuits and Systems, 2004. Proceedings. The 2004 IEEE Asia-Pacific Conference on*, IEEE, pp. 929-32.
- Sutadian A. D., Muttill N., Yilmaz A. G. and Perera B. (2016). Development of river water quality indices: a review. *Environ Monit Assess* **188**(1), 58.
- Swanson R. A. and Holton E. F. (2005). *Research in organizations: Foundations and methods in inquiry*. Berrett-Koehler Publishers, San Francisco.
- Tan K. C., Khor E. F. and Lee T. H. (2005). *Multiobjective Evolutionary Algorithms and Applications*. Springer, United States of America.
- Taoufik G., Khouni I. and Ghrabi A. (2017). Assessment of physico-chemical and microbiological surface water quality using multivariate statistical techniques: a case study of the Wadi El-Bey River, Tunisia. *Arab J Geosci* **10**, 181.
- Tesfaye N. (2007). *Mathematical Modeling of Point Source Pollutants Transport in the Little Akaki River*. M.Sc., Chemical Engineering, Addis Ababa University, Addis Ababa.
- Tibebe M., Melesse A. M. and Zemadim B. (2016). Runoff estimation and water demand analysis for Holetta River, Awash Subbasin, Ethiopia using SWAT and CropWat models. In: *Landscape Dynamics, Soils and Hydrological Processes in Varied Climates*, Springer, pp. 113-40.
- Tong S. T. and Chen W. (2002). Modeling the relationship between land use and surface water quality. *Journal of Environmental Management* **66**(4), 377-93.
- Tung Y.-K. (1992a). Multiple-objective stochastic waste-load allocation. *Water resources management* **6**(2), 117-33.
- Tung Y.-K. and Hathhorn W. E. (1989). Multiple-Objective Waste Load Allocation. *Water resources management* **3**, 129-40.
- Tung Y. K. (1992b). Multiple-objective stochastic waste-load allocation. *Water Resour. Manage.*, **6**(2), 117-33.
- Tunstall S. M., Penning-Rowsell E., Tapsell S. and Eden S. (2000). River restoration: public attitudes and expectations. *Journal of the Chartered Institution of Water and Environmental Management* **14**(5), 363-70.
- Turner D. F., Pelletier G. J. and Kasper B. (2009). Dissolved oxygen and pH modeling of a periphyton dominated, nutrient enriched river. *Journal of Environmental Engineering* **135**(8), 645-52.
- Tyson J., Guarino C., Best H. and Tanaka K. (1993). Management and institutional aspects. *Water Science and Technology* **27**(12), 159-72.

- UN-DESA (2017). World Population Prospects: The 2017 Revision. In, United Nations Department of Economic and Social Affairs (UN DESA), New York: United Nations.
- UN-Habitat (2001). *Cities in a globalizing world: global report on human settlements*. Taylor & Francis, New York.
- UN-Habitat (2002). *Sustainable urbanisation: achieving agenda 21*, Nairobi, Kenya.
- UN-Habitat (2007). *Situation analysis of informal settlements in Addis Ababa: cities without slums sub-regional programme for Eastern and Southern Africa*, United Nations Human Settlements Programme, Nairobi, Kenya.
- UN-Habitat (2010). *The Ethiopia Case of Condominium Housing: The Integrated Housing Development Programme*, United Nations Human Settlements Programme, Nairobi, Kenya.
- USEPA (1996). *Environmental indicators of water quality in the United States*. United States Environmental Protection Agency (USEPA), Washington, D.C., USA.
- USEPA (2011). *A National Evaluation of the Clean Water Act Section 319 Program*, U.S. Environmental Protection Agency (USEPA).
- USEPA (2012). *Guidelines for water reuse*. U.S. Environmental Protection Agency, Washington, D.C.
- Van Rooijen D. and Tadesse G. (2009). Urban sanitation and wastewater treatment in Addis Ababa in the Awash Basin, Ethiopia. In: *34th WEDC international conference on water, sanitation and hygiene: sustainable development and multi-sectoral approach*. edn, Addis Ababa, Ethiopia.
- Van Rooijen D. J., Biggs T. W., Smout I. and Drechsel P. (2010). Urban growth, wastewater production and use in irrigated agriculture: a comparative study of Accra, Addis Ababa and Hyderabad. *Irrig Drain Syst* **24**, 53-64.
- Vanderkruk K., Owen K., Grace M. and Thompson R. (2010). *Review of Existing Nutrient, Suspended Solid and Metal Models*, Monash University, Clayton, Victoria, Australia.
- Vaughan I., Diamond M., Gurnell A., Hall K., Jenkins A., Milner N., Naylor L., Sear D., Woodward G. and Ormerod S. (2009). Integrating ecology with hydromorphology: a priority for river science and management. *Aquatic Conservation: Marine and Freshwater Ecosystems* **19**(1), 113-25.
- Vega M., Pardo R., Barrado E. and Debán L. (1998). Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Res* **32**(12), 3581-92.
- Von Korff Y., Daniell K. A., Moellenkamp S., Bots P. and Bijlsma R. M. (2012). Implementing participatory water management: recent advances in theory, practice, and evaluation. *Ecology and Society* **17**(1).
- von Sperling M. and de Lemos Chernicharo C. A. (2002). Urban wastewater treatment technologies and the implementation of discharge standards in developing countries. *Urban water* **4**(1), 105-14.

- Voutsas D., Manoli E., Samara C., Sofoniou M. and Stratis I. (2001). A study of surface water quality in Macedonia, Greece: Speciation of nitrogen and phosphorus. *Water, Air, & Soil Pollution* **129**, 13-32.
- Wagh V. M., Panaskar D. B., Muley A. A., Mukate S. V., Lolage Y. P. and Aamalawar M. L. (2016). Prediction of groundwater suitability for irrigation using artificial neural network model: a case study of Nanded tehsil, Maharashtra, India. *Model. Earth Syst. Environ.* **2**(4), 196.
- Wang J., Da L., Song K. and Li B.-L. (2008). Temporal variations of surface water quality in urban, suburban and rural areas during rapid urbanization in Shanghai, China. *Environmental Pollution* **152**(2), 387-93.
- Wang Y.-B., Liu C.-W., Liao P.-Y. and Lee J.-J. (2014). Spatial pattern assessment of river water quality: implications of reducing the number of monitoring stations and chemical parameters. *Environmental monitoring and assessment* **186**(3), 1781-92.
- Ward J. H. (1963). Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* **58**, 236-44.
- Wardlaw R. and Sharif M. (1999). Evaluation of genetic algorithms for optimal reservoir system operation. *Journal of Water Resources Planning and Management* **125**(1), 25-33.
- WBG (2015). *Addis Ababa, Ethiopia Enhancing Urban Resilience* The World Bank Group (WBG), Washington DC, USA.
- Webster J. (1995). Automated techniques for the monitoring of water quality. In: *Chemometrics in environmental chemistry - applications* Einax J (ed.), Springer, Verlag.
- Welde K. and Gebremariam B. (2017). Effect of land use land cover dynamics on hydrological response of watershed: Case study of Tekeze Dam watershed, northern Ethiopia. *International Soil and Water Conservation Research* **5**(1), 1-16.
- Weldegebriel Y., Chandravanshi B. S. and Wondimu T. (2012). Concentration levels of metals in vegetables grown in soils irrigated with river water in Addis Ababa, Ethiopia. *Ecotoxicol Environ Saf* **77**, 57-63.
- Whitehead P., Wilby R., Battarbee R., Kernan M. and Wade A. J. (2009). A review of the potential impacts of climate change on surface water quality. *Hydrological Sciences Journal* **54**(1), 101-23.
- Worku Y., Ejigu S., Erge W. and Jemaneh L. (2017). Chemical, physical, and microbiological characteristics of various sources of water in and around Addis Ababa. *The Ethiopian Journal of Health Development (EJHD)* **13**(3).
- Wu E. M.-Y. and Kuo S.-L. (2012). Applying a multivariate statistical analysis model to evaluate the water quality of a watershed. *Water Environ Res* **84**(12), 2075-85.
- Yandamuri S. R., Srinivasan K. and Bhallamudi S. M. (2006). Multiobjective optimal waste load allocation models for rivers using nondominated sorting genetic algorithm-II. *Journal of Water Resources Planning and Management* **132**, 133-43.
- Yesuf H. M., Assen M., Alamirew T. and Melesse A. M. (2015). Modeling of sediment yield in Maybar gauged watershed using SWAT, northeast Ethiopia. *CATENA* **127**, 191-205.

- Zhang G., Patuwo B. E. and Hu M. Y. (1998). Forecasting with artificial neural networks: the state of the art. *Int J Forecast* **14**, 35-62.
- Zhang R., Qian X., Yuan X., Ye R., Xia B. and Wang Y. (2012). Simulation of Water Environmental Capacity and Pollution Load Reduction Using QUAL2K for Water Environmental Management. *Int. J. Environ. Res. Public Health* **9**, 4504-21.
- Zheng L.-y., Yu H.-b. and Wang Q.-s. (2016). Application of multivariate statistical techniques in assessment of surface water quality in Second Songhua River basin, China. *J. Cent. South Univ* **23**, 1040-51.

APPENDIX I: METEOROLOGICAL DATA

(Original data obtained from National Meteorology Agency, Data collected from: Addis Ababa observation station, 2,386m amsl, (38.7475, 9.01891)).

A. Mean Precipitation Data (mm)

Year	2015												2016												2017							
Month	01	02	03	04	05	06	07	08	09	10	11	12	01	02	03	04	05	06	07	08	09	10	11	12	01	02	03	04	06	07		
Date	1	0.0	0.0	0.0	0.0	0.0	0.6	12.6	0.0	2.9	0.0	0.0	0.0	0.0	0.0	0.0	3.8	12.3	0.0	7.1	1.2	2.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4
	2	0.0	0.0	0.0	0.0	0.0	0.0	3.2	4.3	0.7	0.0	0.0	0.0	0.0	0.0	0.0	1.2	15.6	0.0	11.6	4.0	5.4	0.0	0.0	1.9	0.0	0.0	0.0	0.0	0.0	34.4	
	3	0.0	0.0	0.0	0.0	0.0	0.0	1.2	2.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	3.4	0.0	0.0	2.6	1.1	3.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8
	4	0.0	0.0	0.0	0.0	2.1	0.0	0.5	2.0	8.0	0.0	0.0	0.0	0.0	0.0	11.0	0.0	0.0	2.0	9.0	8.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.0	
	5	0.0	0.0	0.0	0.0	2.1	0.0	3.0	0.0	8.7	0.0	3.3	0.0	0.0	0.0	0.0	0.0	11.1	41.2	0.0	14.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	7.2
	6	0.0	0.0	0.0	0.0	1.8	0.0	5.8	4.2	0.0	0.0	2.7	0.0	0.0	0.0	10.3	12.8	21.2	0.0	8.2	11.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	38.1
	7	0.0	0.0	0.0	0.0	0.0	1.0	5.1	0.4	47.8	0.0	0.0	0.0	0.0	0.0	2.2	15.0	5.6	3.2	4.6	47.7	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	
	8	0.0	0.0	0.0	0.0	14.2	0.0	1.8	17.1	12.8	0.0	0.0	0.0	0.0	0.0	3.6	0.0	0.0	2.2	4.4	16.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.2	
	9	0.0	0.0	0.0	0.0	5.2	0.0	14.5	14.4	4.5	0.0	1.7	0.0	3.4	0.0	12.5	0.0	3.2	1.3	11.6	3.3	6.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
	10	0.0	0.0	0.0	2.9	7.0	0.0	1.4	9.6	1.7	0.0	0.0	0.0	19.7	0.0	2.3	6.5	4.0	0.9	9.3	17.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	23.2	
	11	0.0	0.0	0.0	0.0	0.0	0.9	10.5	0.0	0.8	0.0	0.0	0.0	2.6	0.0	0.0	0.0	20.2	0.0	9.8	15.6	14.6	6.3	0.0	0.0	0.0	0.0	0.0	0.0	3.5	6.2	
	12	0.0	0.0	0.0	0.0	0.0	7.5	3.5	11.3	2.0	0.0	0.0	0.0	3.8	0.0	0.0	9.0	0.0	0.0	2.2	6.0	0.7	1.7	0.0	0.0	0.0	0.0	0.0	0.0	1.0	3.8	
	13	0.0	0.0	0.0	0.5	0.0	0.0	16.8	8.8	11.9	0.0	0.0	0.0	0.2	0.0	0.0	15.7	11.8	0.0	3.5	0.0	8.8	0.0	0.0	0.0	0.0	19.3	0.0	0.0	0.0	5.5	
	14	0.0	0.0	0.0	11.6	0.0	0.0	2.3	16.3	10.2	0.0	0.0	0.0	19.7	0.0	0.0	8.3	0.0	3.7	2.8	17.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	
	15	0.0	0.0	0.0	0.0	0.0	0.1	13.4	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	12.0	0.0	9.2	0.0	0.1	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	1.3	
	16	0.0	0.0	0.0	0.0	0.0	16.9	1.7	32.3	0.4	0.0	0.0	0.0	0.4	0.0	0.0	1.3	0.0	10.2	10.2	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	
	17	0.0	0.0	0.0	0.0	0.0	3.0	13.2	22.6	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.6	15.0	5.7	3.3	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	36.7	
	18	0.0	0.0	0.0	0.0	0.0	0.0	5.4	24.0	5.9	0.0	0.0	0.0	0.0	0.0	0.0	3.4	13.2	7.1	7.7	23.4	1.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.4	3.5	
	19	0.0	0.0	0.0	0.0	0.0	3.8	0.0	26.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	7.3	3.5	8.6	15.2	0.0	0.0	0.0	0.0	0.0	4.8	0.0	0.0	23.0	
	20	0.0	0.0	0.0	0.0	0.0	1.2	18.1	33.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	12.3	12.0	0.5	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	4.8	
	21	0.0	0.0	12.8	0.0	0.0	2.8	4.7	7.5	16.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.8	11.8	3.7	1.0	4.5	0.0	0.0	0.0	0.0	0.5	0.0	1.0	3.2	
	22	0.0	0.0	3.5	0.0	2.5	21.6	3.3	6.0	4.2	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	6.2	2.5	5.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	
	23	0.0	0.0	9.8	0.0	0.0	21.0	20.4	7.4	7.0	0.0	0.0	0.0	0.0	0.0	0.3	8.3	0.0	8.2	0.5	5.8	12.3	0.0	0.0	0.0	0.0	0.0	2.3	0.0	2.7	2.2	
	24	0.0	0.0	0.0	0.0	2.1	22.5	8.5	22.8	0.0	0.0	0.0	0.0	0.0	12.1	0.0	2.3	0.0	0.8	4.0	12.6	0.1	0.0	0.5	0.0	0.0	0.7	11.7	0.0	12.1	5.0	
	25	0.0	0.0	1.1	0.0	12.6	22.5	3.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	2.9	15.2	7.0	9.4	3.4	0.0	0.0	0.0	0.0	0.6	0.0	0.0	3.2	9.1	
	26	0.0	0.0	0.0	0.0	27.6	23.0	4.5	0.2	0.0	0.0	0.0	0.1	0.0	0.0	2.2	2.7	0.0	0.2	4.8	37.3	1.5	0.0	0.0	0.0	0.0	0.1	0.0	0.0	2.6	3.5	
	27	0.0	0.0	0.0	0.0	0.5	20.5	1.1	0.3	0.0	0.0	0.0	0.0	2.5	0.0	1.3	2.3	0.0	10.0	17.9	0.3	0.0	0.0	1.2	0.0	0.0	6.5	0.0	7.2	5.6		
	28	0.0	0.0	0.0	0.0	1.5	19.4	6.8	1.1	1.4	0.0	0.0	0.0	6.8	0.0	0.0	0.0	0.0	7.1	1.8	7.6	0.0	0.0	1.9	0.0	0.0	0.6	8.0	6.2	3.8		
	29	0.0	0.0	0.0	0.0	30.1	23.0	9.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.0	0.0	4.0	3.0	0.0	13.0	0.0	0.0	0.0	0.0	10.2	17.3	13.0	35.0		
	30	0.0	0.0	0.0	0.0	0.0	2.6	17.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.6	0.0	14.2	2.2	0.2	40.1	0.0	0.0	0.0	0.0	0.0	8.0	5.5	1.8		

B. Mean Relative Humidity Data (%)

Year	2016												2017												2018								
Month	01	02	04	04	05	06	07	08	09	10	11	12	01	02	03	05	06	07	08	09	10	11	12	01	02	03	04	06	07	08	09		
Date	37.8	31.0	39.0	38.0	37.8	60.8	72.3	70.5	69.5	66.5	58.3	47.8	50.0	62.3	38.5	77.5	49.3	55.5	70.8	79.5	80.8	35.0	53.8	32.3	38.3	41.0	41.0	53.0	59.0	79.5	74.3		
2	36.3	34.3	33.8	40.8	44.0	69.3	61.3	77.0	57.5	51.0	62.5	49.8	37.0	56.8	39.5	68.0	46.3	59.0	70.8	80.5	62.3	40.0	68.0	28.3	45.0	37.8	43.0	56.0	68.3	67.3	74.5		
3.0	36.0	30.3	27.8	43.8	51.5	55.5	72.3	76.0	54.8	59.0	51.5	48.0	56.8	39.5	47.5	63.5	52.3	51.0	75.5	82.0	45.5	35.0	42.3	28.8	48.8	38.5	42.5	50.0	83.5	70.3	79.5		
4.0	39.0	27.5	24.5	42.3	60.8	37.0	70.0	71.8	58.8	46.5	56.8	45.8	54.3	48.3	54.8	54.5	52.3	53.5	71.8	60.8	48.5	36.8	36.8	28.8	53.3	37.5	42.0	50.3	80.3	68.8	81.0		
5.0	48.5	33.3	34.5	41.3	69.8	38.8	71.3	76.5	79.8	53.0	67.8	48.8	51.3	44.5	53.0	55.3	61.8	53.0	##	45.5	49.5	40.5	35.3	29.0	45.0	35.5	42.5	58.0	77.0	74.0	78.0		
6.0	49.8	29.3	35.0	62.8	69.0	38.8	59.7	73.0	68.0	57.0	71.0	47.8	50.3	48.8	52.8	68.0	61.8	58.3	75.0	49.5	48.0	52.8	38.3	23.0	45.0	38.3	37.0	48.3	78.3	60.5	82.3		
7.0	54.0	39.0	0.0	39.5	72.0	59.0	62.7	82.3	67.5	60.5	69.8	60.3	48.5	45.0	51.8	71.3	68.8	66.5	82.5	56.5	45.5	41.3	39.8	28.3	45.3	32.5	43.8	50.3	76.5	81.0	79.5		
8.0	49.0	32.5	35.3	25.0	82.0	56.3	71.7	74.0	74.0	48.0	60.5	47.0	46.3	52.8	61.3	59.0	73.0	52.5	83.8	73.0	49.5	47.0	43.0	21.3	48.3	42.3	42.3	54.0	85.5	86.3	70.5		
9.0	52.3	34.0	34.3	40.3	77.5	53.5	79.0	68.0	74.8	52.3	55.3	62.3	48.3	55.5	59.0	63.5	57.8	56.5	80.3	62.8	50.8	52.0	40.0	26.5	32.5	34.8	44.0	52.8	74.0	84.5	70.8		
10.0	48.0	30.5	34.8	45.3	79.0	62.8	59.3	76.8	72.8	55.5	48.5	74.8	56.3	48.5	61.3	83.3	62.8	45.8	84.3	65.0	44.8	46.3	39.3	27.8	29.8	33.3	34.8	54.0	70.5	84.8	83.3		
11.0	55.3	36.0	38.5	33.8	64.0	0.0	42.0	61.8	66.5	55.0	53.0	59.8	59.3	46.5	43.0	85.0	54.0	46.5	74.8	65.3	68.5	36.5	42.0	27.3	33.5	37.0	38.3	64.0	79.0	81.5	78.5		
12.0	62.3	32.5	44.0	39.8	56.8	66.8	63.7	61.5	66.5	52.0	52.8	55.0	69.0	46.0	42.8	64.0	47.3	61.3	74.3	64.8	70.5	44.5	43.3	28.0	36.5	32.5	37.5	60.5	77.5	76.8	56.3		
13.0	56.8	40.3	40.3	56.8	54.8	58.8	69.3	67.3	66.8	48.0	60.0	56.3	80.5	42.8	37.3	60.0	48.0	60.3	65.3	71.3	51.5	40.3	45.8	30.5	51.3	37.3	40.3	54.3	78.3	78.0	53.0		
14.0	49.3	40.5	41.5	51.8	45.5	60.5	67.0	57.3	51.0	46.5	42.0	64.5	70.3	51.5	34.0	49.3	79.3	58.5	70.8	67.5	55.8	31.3	38.8	26.5	69.3	34.8	38.0	49.8	79.0	75.5	81.0		
15.0	44.5	44.5	35.0	66.0	44.0	58.5	59.7	79.5	62.0	45.0	42.5	60.3	67.5	41.0	35.5	42.0	63.0	54.3	76.5	53.3	46.3	38.3	37.5	31.3	53.3	31.0	40.3	49.3	62.8	73.8	72.5		
16.0	44.5	49.3	28.8	36.5	41.5	64.5	75.0	64.5	58.0	46.5	89.0	60.8	63.5	43.5	35.5	38.3	56.8	60.0	61.3	63.5	51.0	44.8	45.0	28.5	52.0	32.8	38.0	47.0	66.8	71.5	71.3		
17.0	47.8	50.8	33.8	44.0	47.8	63.3	61.0	75.0	76.0	49.3	47.5	55.0	74.8	47.8	41.0	60.3	72.8	54.8	75.0	70.0	41.3	42.3	48.5	27.5	51.0	36.5	33.5	50.0	75.3	72.3	81.5		
18.0	52.3	42.8	34.3	42.8	40.0	60.5	70.7	77.0	68.3	52.0	46.5	50.5	63.8	45.8	43.3	65.0	86.0	57.5	80.0	73.3	43.3	44.8	48.0	28.8	56.3	41.0	37.0	53.0	69.3	68.5	81.5		
19.0	57.8	39.0	40.0	43.3	36.0	63.3	63.3	76.8	60.0	46.8	52.8	53.3	52.8	45.8	43.0	69.3	72.5	53.5	86.0	67.8	50.5	42.0	46.5	27.8	62.3	45.8	39.8	62.0	69.3	74.3	75.5		
20.0	59.8	42.5	53.0	44.0	30.3	62.5	45.7	72.0	51.5	47.8	48.8	51.3	54.3	38.0	59.5	52.3	65.3	68.8	81.8	69.3	45.8	20.5	45.5	34.8	55.8	48.5	43.8	52.3	79.3	68.0	83.8		
21.0	25.5	49.5	0.0	36.3	34.5	67.3	49.0	65.5	61.5	47.5	46.0	54.5	59.3	38.5	38.8	54.0	76.8	64.5	70.0	64.0	46.3	44.5	51.8	35.5	41.0	45.3	39.0	48.3	76.8	63.8	76.3		
22.0	50.8	48.5	36.3	41.5	58.3	68.8	70.0	59.5	77.8	54.0	47.3	52.3	53.3	35.5	41.5	49.0	62.0	50.8	64.5	55.0	57.0	44.5	57.8	33.8	39.0	48.8	41.5	49.3	71.5	72.0	73.3		
23.0	48.5	40.0	64.3	31.5	60.8	68.0	74.3	66.0	62.5	50.3	50.5	44.0	46.0	36.3	52.8	58.8	57.3	58.8	73.0	62.8	53.0	46.8	36.5	33.3	40.0	54.5	33.0	51.3	79.0	77.8	77.8		
24.0	45.0	47.5	53.3	32.0	0.0	74.5	76.7	69.8	59.3	46.8	43.8	49.3	39.5	56.3	37.0	62.8	80.0	53.8	71.5	71.3	48.3	55.0	34.8	30.8	51.5	56.3	38.8	66.8	67.5	66.0	69.0		
25.0	41.5	48.5	55.3	42.3	69.8	68.0	68.3	83.0	61.0	45.3	57.8	60.0	44.0	62.8	54.3	67.0	81.0	53.3	78.0	71.3	44.5	57.3	46.8	36.3	57.3	52.3	45.5	70.5	77.5	78.3	70.5		
26.0	45.5	50.8	46.0	46.8	55.3	73.8	61.7	74.0	58.0	52.8	42.3	65.5	49.8	63.0	46.5	64.8	78.8	55.0	67.0	66.3	42.3	66.3	37.5	38.3	45.8	56.0	42.8	65.5	86.0	70.5	59.8		
27.0	37.0	56.3	47.5	43.5	65.5	80.3	65.3	62.5	56.0	48.0	57.5	60.0	52.5	50.5	47.5	61.3	65.8	63.3	69.8	67.3	45.3	64.8	33.3	35.3	61.8	53.5	45.8	77.5	78.3	54.8	53.8		
28.0	28.3	39.8	52.0	37.3	72.3	86.0	63.3	75.8	56.8	56.0	41.0	57.0	60.0	49.8	40.8	47.0	72.8	57.5	71.5	57.8	50.0	73.5	38.8	35.8	55.5	65.5	50.0	75.0	73.3	77.5	57.0		
29.0	29.3	0.0	48.8	39.5	74.5	83.5	65.0	74.3	63.3	56.0	56.3	49.5	63.5	36.3	48.5	46.0	75.0	58.0	73.8	69.8	50.0	54.3	37.0	44.8	0.0	63.8	74.0	55.3	77.5	89.0	75.3		
30.0	27.3	0.0	44.5	35.3	57.5	71.5	76.7	67.0	50.0	58.0		48.0	68.5	47.6	45.8	43.5	74.5	50.3	73.3	85.3	43.8	47.3	42.5	48.5	0.0	56.8	70.3	73.8	78.3	77.3	90.0		

C. Sun Hours Data

Year	2015												2016												2017								
Month	01	02	03	04	05	06	07	08	09	10	11	12	01	02	03	05	06	08	09	10	11	12	01	02	03	04	06	07	08	09			
Date	1	9.8	11.0	10.3	10.6	10.5	3.3	2.6	2.7	8.0	10.5	5.5	6.9	8.0	9.2	8.8	4.2	7.6	2.5	0.7	3.4	10.9	5.8	10.2	10.0	10.2	9.3	6.5	0.4	2.0	1.9		
	2	10.4	10.5		8.9	7.9	2.0	3.0	0.2	9.5	10.7	6.5	10.3	10.2	10.3	9.2	3.0	6.5	1.6	0.0	1.6	9.8	2.3	10.5	10.0	10.3	8.3	9.0	4.5	3.4	3.5		
	3	9.8	10.5	9.7	9.9	6.3	4.6	1.2	3.3	9.5	9.7	10.1	7.3	10.3	8.6	7.6	6.6	6.0	0.1	2.5	9.4	10.3	10.3	10.3	10.3	10.0	10.6	8.0	0.0	4.2	1.4		
	4	10.3	12.0	10.4	8.1	4.2	9.7	3.2	6.5	10.5	6.2	3.0	10.0	5.6	10.4	7.4	8.2	7.8	0.4	9.4	9.7	10.2	10.5	9.3	6.6	10.0	8.5	5.2	1.3	4.5	2.5		
	5	9.5	10.5	10.3	8.9	0.5	9.0	4.0	1.3	4.0	1.2	8.0	8.8	10.0	8.7	8.3	8.9	3.0	0.0	10.1	4.7	10.2	8.6	9.2	10.4	10.3	9.8	6.1	3.2	1.2	2.6		
	6	10.3	9.5	9.7		2.3	7.6	7.1	2.5	1.8	8.8	5.6	10.3	8.3	10.3	6.5	4.2	5.6	1.0	9.5	8.6	10.2	10.4	10.4	8.2	9.7	10.6	6.5	3.4	3.6	2.5		
	7	9.8	10.3		6.7	1.8	2.7	5.2	1.0	2.1	8.3	2.8	6.2	10.0	9.5	6.4	2.5	4.0	2.0	7.5	8.7	9.7	10.0	10.4	10.2	10.3	6.5	7.0	2.2	0.7	2.4		
	8	10.4	10.0	10.5		0.1	5.6	5.0	1.6	0.1	8.2	5.5	7.0	9.0	7.8	5.2	5.4	4.5	0.6	0.0	10.4	0.6	9.5	10.5	9.6	9.2	9.6	4.1	0.8	0.0	5.5		
	9	9.7	9.8	10.3		1.2	3.6	3.5	6.2	0.1	9.4	10.0	10.0	7.2	9.8	6.3	4.4	5.6	0.1	3.2	10.1	1.3	10.4	10.4	10.1	10.2	10.5	7.0	0.1	2.0	5.1		
	10	10.1	10.9	9.9	7.1	1.1	6.2	5.8	2.2	0.0	7.1	6.7	1.0	8.5	9.5	4.9	0.4	5.2	0.1	5.2	6.4	10.2	10.4	10.4	9.8	10.0	10.6	0.0	5.2	3.0	3.3		
	11	8.5	10.5	10.4	10.8	3.0		4.5	1.0	0.1	6.7	6.7	2.2	8.8	9.3	10.3	0.2	3.9	2.7	5.1	3.3	8.4	10.3	8.2	5.2	10.6	10.8	4.2	1.5	1.2	4.8		
	12	6.6	9.6	7.3	10.1	8.1	2.5	7.0	8.1	0.0	9.8	2.2	3.0	3.0	9.8	10.3	5.6	10.2		3.3	2.3	1.3	9.2	10.5	5.5	9.8	10.6	2.7	4.2	2.0	9.7		
	13	6.0	10.5	10.0	3.5	4.6	6.3	1.2	2.1	5.6	9.6	8.3	3.5	2.2	9.7	10.3	6.6	7.5	5.2	3.6	5.7	8.4	10.2	10.4	3.1	8.3	8.7	5.1	1.5	0.6	10.4		
	14	10.0	6.2	10.3	4.2	10.0	4.6	2.6	4.2	3.4	9.2	9.0	3.0	4.7	9.8	10.4	9.8	1.2	5.4	3.5	6.3	10.3	8.6	9.9	1.6	10.7	10.3	5.4	2.3	2.3	3.9		
	15	9.7	6.8	10.4	4.3	7.3	3.5	3.0	1.8	5.6	7.8	10.3	6.8	4.2	10.2	10.6	10.3	5.0	3.0	9.8	6.2	10.4	9.4	10.3	9.8	9.0	9.8	10.0	3.5	4.8	4.0		
	16	10.2	5.5	10.5	7.5	10.0	5.3	0.4	5.7	10.3	10.2	9.3	7.0	4.4	10.0	10.7	8.3	6.0	6.5	7.2	8.8	10.2	10.0	10.0	10.4	10.0	7.8	6.6	3.7	5.7	1.5		
	17	10.0	6.4	10.2	3.6	10.0	2.6	6.9	3.0	0.1	9.2	5.6	7.3	3.5	7.8	9.7	2.4	2.6	2.3	2.2	10.2	8.3	8.7	10.3	9.6	10.1	10.1	5.5	2.8	5.8	1.8		
	18	10.0	9.8	10.3	10.3	7.2	7.0	6.0	3.4	9.3	9.4	9.3	10.0	6.5	8.9	8.5	3.5	0.3	1.5	1.7	9.2	10.0	10.4	10.2	8.8	9.2	8.2	4.0	4.1	2.5	1.8		
	19	6.3	9.2	8.5	10.2	7.2	4.6	4.7	5.0	8.8	10.2	10.3	8.5	7.0	10.2	9.0	5.6	1.1	0.0	2.3	10.3	9.9	10.3	10.4	5.3	7.8	10.2	3.5	1.2	5.5	2.4		
	20	6.4	10.4	2.6	7.6	9.8	7.2	5.1	4.5	8.5	8.8	10.0	7.7	8.7	9.0	8.0	8.0	0.3	0.1	3.1	8.5	10.2	8.4	10.2	10.3	7.0	7.0	8.2	1.2	5.0	2.8		
	21	5.5	8.5		10.4	9.3	5.0	5.9	5.1	4.7	7.1	10.4	12.0	7.6	10.0	8.3	9.7	3.6	1.9	4.5	5.5	7.8	8.7	10.0	10.4	6.7	8.5	8.4	2.2	5.5	3.5		
	22	9.0	10.5	10.0	8.8	5.4	1.8	2.0	10.0	3.0	8.0	10.5	10.2	10.0	9.2	4.8	5.6	3.8	5.7	10.0	2.0	10.1	8.6	9.5	10.3	7.8	6.0	6.5	1.9	3.3	1.4		
	23	9.9	10.4	2.0	10.7	8.0	0.4	3.2	6.1	7.0	7.9	8.7	8.2	10.2	9.3	5.5	4.2	6.7	4.5	5.2	6.2	7.5	10.2	10.3	10.0	4.2	10.7	6.3	1.6	2.2	4.2		
	24	10.1	8.6	3.2	7.8		0.6	2.7		7.0	10.0	10.0	9.5	8.2	8.0	3.4	3.2	2.0	5.5	2.5	7.8	5.6	10.4	9.7	8.0	5.2	9.8	5.0	4.3	1.4	5.2		
	25	10.5	7.4	2.8	9.5		3.4	5.8	0.5	6.0	7.5	9.3	6.0	10.2	4.5	6.0	3.3	3.0	1.6	1.8	8.2	7.0	8.5	10.0	3.5	5.0	7.6	1.2	2.0	2.9	1.3		
	26	10.0	9.6	3.2	3.0	3.0	2.5	6.2	4.5	5.7	8.2	10.4	7.5	10.3	6.1	5.5	2.6	3.0	6.5	3.7	10.2	3.5	10.4	10.2	6.3	2.7	8.7	3.9	0.7	0.6	5.9		
	27	10.4	10.2	10.2		4.0	0.1	4.0	4.2	5.3	9.9	9.5	7.0	8.6	8.0	4.7	6.1	2.2	5.3	1.8	10.5	2.0	10.3	10.3	2.0	7.0	7.4	0.8	3.5	1.8	8.2		
	28	10.3	10.0	10.0	2.6	1.7	0.0	5.7	4.5	6.2	8.0	7.8	7.3	6.7	10.0	7.0	6.8	1.2	4.0	5.2	10.3	1.4	9.0	10.0	6.0	5.3	3.1	2.1	2.5	3.0	6.5		
	29	10.4		8.2	10.5	3.6	2.0	3.0	1.1	5.0	4.2	9.6	9.9	3.4	10.0	6.2	9.3	0.9	1.8	5.1	8.2	10.2	10.0	7.8		3.0	0.0	6.5	5.0	0.0	0.6		
	30	10.6		7.3	10.5	6.5	2.6	1.0	8.7	8.1	7.0	10.2	8.5	6.4	9.1	7.9	7.8	3.8	2.2	1.2	10.5	8.8	10.1	10.0		8.8	2.0	5.2	2.5	1.5	0.0		

D. Wind Speed Data (m/s)

Year	2016												2017						
Month	04	01	02	03	04	05	06	07	08	09	10	11	12	01	02	05	08	09	
Date	1	0.9	0.9	0.9	1.0	1.1	0.4	0.6	0.4	0.8	0.3	0.4	0.6	0.5	0.6	0.7	0.8	0.4	0.5
	2	0.7	0.9	1.1	0.8	0.6	0.6	0.6	0.5	1.0	0.4	0.7	0.6	0.5	0.6	0.3	0.5	0.5	0.5
	3	1.1	1.6	1.0	0.8	0.5	0.9	0.4	0.5	0.7	0.3	0.7	0.6	1.0	0.7	0.7	0.7	0.4	0.3
	4	0.7	1.3	1.2	0.7	0.5	1.0	0.4	0.5	0.5	0.7	0.8	0.6	1.1	0.5	0.6	1.3	0.4	0.8
	5	0.4	1.2	1.0	0.6	0.6	0.8	0.3	0.5	0.2	0.7	0.5	0.6	0.7	0.7	0.6	0.6	0.3	0.4
	6	0.0	0.8	0.6	0.7	0.7	0.5	0.5	0.6	0.4	0.6	0.6	0.7	0.6	0.6	0.8	0.6	0.4	0.4
	7		0.9	0.7	0.4	0.8	0.3	0.3	0.3	0.3	0.5	0.7	0.8	0.5	0.4	0.7	0.8	0.3	0.3
	8		0.5	0.7	0.7	0.5	0.9	0.3	0.5	0.4	0.3	0.7	0.7	0.5	0.5	0.8	0.9	0.1	0.5
	9	0.8	0.7	0.8	0.5	0.6	0.5	0.4	0.5	0.5	0.5	0.6	0.7	0.6	0.6	0.8	1.0	0.3	0.3
	10	0.6	0.7	0.9	0.8	0.5	0.3	0.4	0.4	0.4	0.5	0.6	1.9	0.6	0.7	0.6	0.7	0.2	0.3
	11	0.7	0.8	1.0	0.4	0.8	0.2	0.7	0.5	0.5	0.3	0.4	0.2	0.6	0.7	0.6	0.9	0.4	0.3
	12	1.4	0.6	0.9	1.3	0.6	0.4	0.6	0.3	0.4	0.3	0.3	0.5	0.8	0.5	0.4	0.6	0.3	0.6
	13	1.0	0.4	0.9	1.4	0.5	0.3	0.4	0.4	0.3	0.3	0.5	0.8	0.7	0.4	0.6	0.3	0.3	0.7
	14	0.7	0.6	1.1	1.2	0.6	0.8	0.4	0.3	0.3	0.5	0.5	0.8	0.6	0.4	0.6	0.9	0.3	0.4
	15	0.8	0.4	1.0	1.1	0.7	0.9	0.5	0.6	0.3	0.7	0.5	0.6	0.6	0.5	1.2	0.5	0.3	0.4
	16	1.0	0.5	1.0	1.3	0.9	0.6	0.5	0.3	0.4	0.5	0.7	0.7	0.7	0.7	0.5	0.2	0.5	0.4
	17	0.5	0.5	0.7	1.0	0.6	0.5	0.6	0.4	0.6	0.6	0.6	0.7	0.8	0.8	0.9	0.3	0.5	0.3
	18	1.0	0.6	0.8	1.0	0.0	0.4	0.2	0.4	0.3	0.3	0.5	0.5	0.7	0.7	1.2	0.9	0.2	0.3
	19	0.8	0.8	1.0	1.2	0.5	0.4	0.5	0.5	0.3	0.3	0.6	0.4	1.5	0.6	0.5	0.1	0.1	0.3
	20	0.7	1.2	1.0	1.3	0.0	0.4	0.5	0.4	0.5	0.4	0.6	1.3	0.7	0.5	0.7	0.3	0.4	0.2
	21	0.9	0.9	0.6	1.0	1.0	0.5	0.4	0.1	0.3	0.4	0.5	0.8	0.7	0.5	0.6	0.7	0.4	0.4
	22	0.8	0.6	0.8	0.8	1.0	0.4	0.7	0.3	0.4	0.8	0.4	0.7	0.8	0.7	0.5	0.6	0.3	0.3
	23	0.8	0.6	1.7	1.0	0.6	0.4	0.8	0.5	0.3	0.4	0.5	0.5	0.9	0.8	0.7	0.8	0.2	0.3
	24	0.8	0.7	1.0	1.0	0.5	0.4	0.4	0.7	0.4	0.5	0.6	0.4	0.8	0.7	0.7	0.5	0.3	0.5
	25	1.0	0.6	0.7	1.0	0.6	0.4	0.4	0.6	0.4	0.3	0.7	0.4	0.9	0.6	0.8	0.6	0.2	0.3
	26	0.6	0.6	0.8	1.0	0.8	0.5	0.3	0.3	0.4	0.6	0.7	1.0	0.6	0.6	0.7	0.4	0.3	0.7
	27	0.8	0.7	1.3	0.0	0.6	0.4	0.6	0.4	0.4	0.5	0.7	0.5	0.6	0.4	0.6	0.4	0.2	0.6
	28	1.2	0.8	1.2	0.0	0.4	0.6	0.6	0.3	0.3	0.7	0.7	0.5	0.5	0.6	0.8	0.2	0.3	0.5
	29	1.1	0.9	1.1	1.0	0.5	0.6	0.4	0.8	0.4	0.3	0.6	0.3	0.5	0.5		0.3	0.2	0.3
	30	1.1	1.1	0.9	1.0	1.0	1.0	0.5	0.9	0.5	0.3	0.7	0.8	0.8	0.6		0.3	0.3	0.1

APPENDIX II: MAXIMUM and MINIMUM VALUES of RAW WQ DATA

Sr. No.	Description	Maximum		Minimum	
		Dry Season	Wet Season	Dry Season	Wet Season
1	TSS, mg/L	1,587	1,062	5	20
2	T-NH ₃ , µg/L	184,800	83,000	40	400
3	NO ₃ , µg/L	97,800	12,100	0	0
4	NO ₂ , µg/L	1,700	740	0	0
5	T-N, µg/L	365,000	97,000	0	0
6	T-P, µg/L	30,800	18,400	0	0
7	T-alk.,mg/L	1,221	975	110	7.5
8	TOC, mg/L	135	21.4	0.5	2.1
9	COD, mg/L	1,672.56	1,580	20	25
10	BOD, mg/L	470	720	5	20
11	DO, mg/L	8.6	9.53	0	0
12	T°,°C	17.0	23.4	8	9.8
13	EC, µS/cm	810	5,130	300	300
14	pH	8.44	8.95	4.5	6.7

APPENDIX III: PSEUDO-CODE FOR Q2Kw-PIKAIA-GA based WLA MODELING

Step 1	<p>Declaration of variables</p> <ul style="list-style-type: none">- <i>Public declarations</i><ul style="list-style-type: none">- Ctrl(14): array but have any type of values- Default (14): default values for Ctrl(14) array- n: number of adjustable parameters in the fitness function (42 in this case)- x: parameter values to the fitness function, $ff . x = (x_1, x_2, \dots, x_n)$- f: fitness value from the fitness function at $x(f=ff(n,x))$- xtemp: temporary array for $x()$ to build and pass to ff- xbest: final best value from last generation - <i>Checking variables</i><ul style="list-style-type: none">- NMAX: maximum number of adjustable parameters ($n \leq NMAX$) in the rates sheet (for optimization purpose, it refers to fractional removal from pollution sources)- PMAX: maximum number of population- GMAX: maximum number of genes per chromosomes - <i>Phenotypes</i><ul style="list-style-type: none">- Phparent- Phoffspring
--------	--

- *Pikai declaration*
 - npop: population size
 - ngen: number of generations
 - lp: random phenotype population
 - np: individual in a population
 - Fdif: fitness differential individuals
 - Important declarations
 - Oldph: place where population are stored
 - newph: temporary location
 - ifit: access to the information in terms of rank
 - oldph(1:n,ifit(np)) represents the fittest
 - imut= mutation mode
 - irep= reproduction plans
 - ielit=elitism
 - pcross= crossover rate
 - pmut = initial mutation rate
 - Fdif: fitness differential
 - lfit: ranking array (rank key index)

- *Remarks*
 - the fittest phenotype is stored in oldph (1:n, ifit(np))
 - Jfit: actual ranks of the individuals stored in oldph(1:n,i)
 - Xnum: rate constant ID number for each x parameter ff

Step 2	Call sect1 for Pre-initialization check of the
--------	--

	variables	If irestart = 0 Then	'gp
Step 3	Give error message, if there is a problem	'generate random initial population	'gp
Step 4	Check the values of n, npop and ng are in the acceptable range, if not give an error message.	Sheets("Rates").Select Range("x8").Select .br/>.br/>	'MY 'MY
Step 5	Give the header row and pointer for output population during each generation in the sheet "Auto-optimization". Such as Generation, Rank, Quantiles, Adjustable parameter headings, User defined parameters and Fitness.	For k = 1 To N oldph(k, ipop) = urand() xtemp(k) = oldph(k, ipop) ActiveCell.Value = xtemp(k) ActiveCell.Offset(1, 0).Select .br/>.br/>	'gp 'gp 'gp 'MY 'MY
Step 6	If, lrestart = 0, generate random initial population Oldph (k,ipop): xtemp(k) as an input for the fitness function - The Pikaia-ff(n,xtemp) will be executed and the result will be saved in fitns(ipop) [PIKAIA-ff will: do the fitness function by calling Pikaia_RunQ2K, Pikaia_ff=ActiveCell.Value]	Next k 'fitns(ipop) = ff(n, oldph(1, ipop)) fitns(ipop) = pikaia_ff(N, xtemp()) Range("m8").Select Next ipop .br/>.br/>	'gp 'gp 'gp 'MY 'gp
Step 7	Else, Check if there is a file called oldph-file is in the directory. - Display an error message, if the file does not exist	Sheets("Rates").Select Range("x8").Select .br/>	'MY 'MY
Step 8			

	<p>Else, If irestart is different from zero, open oldph-file.</p> <ul style="list-style-type: none"> - Read the value to pmut. - If pmut < 0 or pmut > 1, invalid file (check all the values are between 0 and 1) (by reading the values from oldph-file, it checks the current configuration in accordance with the oldph-file values of n, ctrl. - Write the oldph value from the file 	<pre> For k = 1 To N 'gp xtemp(k) = oldph(k, i) 'gp ActiveCell.Value = xtemp(k) 'MY ActiveCell.Offset(1, 0).Select 'MY Next k 'gp . . . Sheets("Rates").Select 'MY Range("y9").Select 'MY ActiveCell.End(xlDown).Select 'MY ActiveCell.Offset(1, 0).Select 'MY ActiveCell.Value = Sheets("Qual2K").Range("aa40") 'MY . . . </pre>
<p><i>End of population initialization</i></p>		
<p>Step 9</p>	<p>Rank initial population by fitness order using: Rnkpop-function, after evaluation.</p>	<pre> Range("z9").Select 'MY ActiveCell.End(xlDown).Select 'MY </pre>
<p>Step 10</p>	<p>Write the output of the ranked population for each generation in the "Auto-optimization" sheet including Fitness value.</p>	<pre> ActiveCell.Offset(1, 0).Select 'MY ActiveCell.Value = Sheets("Qual2K").Range("h45") 'MY . . . </pre>

Main Generation loop starts here.		
For igen = 1 to ngen		
Step 11	Crossover/mutation For ipop = 1 to npop/2, - Select mom & dad - Do encode - Do crossover - Do mutate - Do decode	Range("aa9").Select 'MY ActiveCell.End(xlDown).Select 'MY ActiveCell.Offset(1, 0).Select 'MY ActiveCell.Value = Sheets("Qual2K").Range("g60") 'MY . . .
Step 12	Write the values to the "Auto-optimization" sheet - Save best parameters	Range("x8").Select 'MY Next I gp
Step 13	Save the output to the oldph-file	
<i>End when the specified number of generations is complete.</i>		

APPENDIX IV: DATA-TABLES FOR WLA

1	EQUITY
----------	---------------

		B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11		
	X_i/X_m	0.05977178	1.892773048	0.70331462	0.970295236	0.727223	0.814888607	1.22134	1.024089839	1.98841	0.516029705	1.08187		
CBODs	W_i/W_m	0	0	0.69349234	2.21597976	4.531996	0	0	2.47007604	0	0.939253228	0.1492	48.00915455	
	$abs(X_i/X_m - W_i)$	0.05977178	1.892773048	0.00982228	1.245684524	3.804773	0.814888607	1.22134	1.445986201	1.98841	0.423223523	0.93267	13.83933395	0.07225781
ON	W_i/W_m	0.04	2.38	0.45	1.08	2.45	0.34	0.03	0.52	2.20	0.91	0.60	23,740.43	
	$abs(X_i/X_m - W_i)$	0.01835816	0.485288806	0.24850108	0.105354413	1.724923	0.475259501	1.19313	0.503750763	0.2159	0.38973209	0.4822	5.842398043	0.17116259
OP	W_i/W_m	0.08	2.78	0.53	1.24	0.95	0.55	1.10	0.89	1.14	1.50	0.22	1,126.50	
	$abs(X_i/X_m - W_i)$	0.02364248	0.88463559	0.17286739	0.274274148	0.220489	0.263005142	0.11847	0.130862136	0.8458	0.98779856	0.85984	4.7816804	0.2091315
NH4	W_i/W_m	0.02	1.60	1.14	0.82	1.11	0.25	0.46	1.17	2.17	1.89	0.37	12,403.07	
	$abs(X_i/X_m - W_i)$	0.04292225	0.295825524	0.44158304	0.151649505	0.38441	0.568187407	0.76066	0.144895526	0.18265	1.377097695	0.71139	5.061272604	0.19757877
NO2+NO4	W_i/W_m	-	0.02	0.11	3.33	4.93	0.09	0.30	0.38	1.72	0.13	-	5,774.45	
	$abs(X_i/X_m - W_i)$	0.05977178	1.870321709	0.59452133	2.356622848	4.206267	0.724447509	0.92482	0.647592835	0.27037	0.389172244	1.08187	13.12577911	0.07618595
COD	W_i/W_m	0.01	1.91	0.59	1.78	1.88	0.22	0.61	1.19	1.64	1.02	0.15	139.3190073	
	$abs(X_i/X_m - W_i)$	0.04881572	0.015406988	0.11751905	0.808060265	1.150889	0.590837229	0.61085	0.17033299	0.34829	0.50475641	0.93314	5.298890747	0.18871874
CBODf	W_i/W_m	0.04	2.21	0.53	1.55	0.60	0.38	1.24	0.58	2.66	1.05	0.15	#REF!	
	$abs(X_i/X_m - W_i)$	0.01582958	0.318382654	0.17422037	0.583392781	0.12668	0.433283664	0.02094	0.446540785	0.67489	0.5337275	0.93478	4.262674975	0.23459448

OVERALL EQUITY	1.14962985
OVERALL EQUITY RECIPROCAL	0.869845

Where X_i : GA generated fractional WL removal factor for every pollution source

X_m : Average fractional WL removal factor for pollution sources in a given population

W_i : WL reduction for a pollution source obtained from a product of X_i by WL of a given WQ parameter

W_m : Pollution sources average WL reduction in a given population

2 POINT SOURCE

POINT SOURCES																																
#	Environmental Standards Control				Wasteload		CBODf wasteload after fraction		Wasteload		CBODs wasteload after fraction		Wasteload		ON wasteload after fraction removal (wi)		Wasteload		OP wasteload after fraction		Wastelo.		NH3 wasteload after fraction		Wastelo.		NO2+NO3 wasteload after fraction		Wastelo.		COD wasteload after fraction	
	Monitoring Site	Dom	DOs	CBODf	CBODf	CBODs	ON	ON	OP	OP	NH3	NH3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3	NO2+NO3
0.03	23.93	9	0.27	218.8	6.6	0	0.00	53,250.0	1,597.5	4,760.0	142.8	#####	426.0	-	-	80	2.40															
0.95	23.61	6	5.70	347.8	330.4	0	0.00	96,560.0	91,732.0	5,005.0	4,754.8	#####	40,375.0	171.0	162.5	440	418.00															
0.353	22.76	3	1.06	223.9	79.1	139.42	49.22	49,700.0	17,544.1	2,572.5	908.1	#####	28,946.0	2,230.0	787.2	363.52	128.32															
0.487	21.00	1	0.49	476.7	232.1	322.92	157.26	85,200.0	41,492.4	4,375.0	2,130.6	#####	20,697.5	#####	24,072.4	799.92	389.56															
0.365	18.53	0	0.00	245.8	89.7	881.16	321.62	259,150.0	94,589.8	4,445.0	1,622.4	#####	28,105.0	#####	35,697.0	1127.16	411.41															
0.409	15.21	0	0.00	139.4	57.0	0	0.00	32,031.7	13,100.9	2,310.0	944.8	#####	6,237.3	1,600.0	654.4	120	49.08															
0.613	11.54	0	0.00	302.8	185.6	0	0.00	1,775.0	1,088.1	3,080.0	1,888.0	#####	11,647.0	3,500.0	2,145.5	218.16	133.73															
0.514	8.86	0	0.00	167.9	86.3	341.04	175.29	39,050.0	20,071.7	2,975.0	1,529.2	#####	29,555.0	5,300.0	2,724.2	509.04	261.65															
0.998	6.37	0	0.00	398.7	397.9	0	0.00	85,200.0	85,029.6	1,960.0	1,956.1	#####	54,890.0	#####	12,431.1	360	359.28															
0.259	2.48	0	0.00	605.6	156.8	257.36	66.66	134,900.0	34,939.1	9,940.0	2,574.5	#####	47,863.2	3,544.0	917.9	863.36	223.61															
0.543	0.00	0	0.00	40.5	22.0	19.5	10.59	42,600.0	23,131.8	700.0	380.1	#####	9,366.8	-	-	60	32.58															
Average	0.50	14.03	1.73	0.68	287.99	149.41	178.31	70.97	79946.97	38574.27	3829.32	1711.94	#####	25282.61	#####	7235.65	449.20	219.06														

3 MONITORING SITES ON THE RIVER LAR									
MONITORING SITES ON THE RIVER LAR									
QUAL2Kw SIMULATION OUTPUTS					Measured values at Monitoring Sites				
Monitoring Site	DOs	CBODf (wi)	TN (wi)	TP (wi)	Dom	CBODf (wi)	TNm (wi)	TPm (wi)	
23.93	5.18	146.0	5,410.0	-	5.18	145.97	210.00	-	
23.61	5.10	92.2	2,781.2	-	5.53	223.35	12,300.00	7,500.00	
22.76	6.37	219.8	25,151.1	9,533.1	7.94	113.92	86,000.00	6,500.00	
21.00	5.23	153.2	91,901.4	6,724.1	3.80	212.76	85,000.00	4,105.00	
18.53	0.76	476.3	34,509.1	2,500.9	-	486.10	123,333.33	6,150.00	
15.21	0.33	517.7	157,365.8	6,000.9	-	604.08	194,200.00	12,700.00	
11.54	0.49	689.9	216,239.1	13,714.9	-	488.66	160,000.00	30,800.00	
8.86	0.11	406.6	98,017.1	7,890.8	0.09	275.06	42,600.00	6,000.00	
6.37	0.06	262.5	78,574.4	10,055.9	-	395.73	115,000.00	8,400.00	
2.48	0.10	247.8	110,466.2	18,008.7	-	305.79	126,500.00	18,650.00	
0.00	0.10	247.8	110,466.2	18,008.7	-	305.79	126,500.00	18,650.00	
Ave. conc.	2.166027025	314.5394573	84625.59576	8403.461897	2.048181818	323.3821387	97422.12121	10859.54545	
No of viol.	7	11	11	9	8	11	10	10	

*Simulation outputs of Q2Kw based WQ data were determined after each run of WL reduction. Afterwards, number of violations and total magnitude of violations were determined.

4 | MAGNITUDE OF STANDARD VIOLATION

Standa	DO		BOD		TN		TP	
	5.0	5.0	5.0	5.0	1,000.0	1,000.0	1,000.0	1,000.0
	DO Viol. Magnitude (measured)	DO Viol. Magnitude (simulated)	BOD Viol. Magnitude (measured)	BOD Viol. Magnitude (simulated)	TN Viol. Magnitude (measured)	TN Viol. Magnitude (simulated)	TP Viol. Magnitude (measured)	TP Viol. Magnitude (simulated)
	-	-	141.0	141.0	-	4,410.0	-	-
	-	-	218.3	87.2	11,300.0	1,781.2	6,500.0	-
	-	-	108.9	214.8	85,000.0	24,151.1	5,500.0	8,533.1
	1.2	-	207.8	148.2	84,000.0	90,901.4	3,105.0	5,724.1
	5.0	4.2	481.1	471.3	122,333.3	33,509.1	5,150.0	1,500.9
	5.0	4.7	599.1	512.7	193,200.0	156,365.8	11,700.0	5,000.9
	5.0	4.5	483.7	684.9	159,000.0	215,239.1	29,800.0	12,714.9
	4.9	4.9	270.1	401.6	41,600.0	97,017.1	5,000.0	6,890.8
	5.0	4.9	390.7	257.5	114,000.0	77,574.4	7,400.0	9,055.9
	5.0	4.9	300.8	242.8	125,500.0	109,466.2	17,650.0	17,008.7
	5.0	4.9	300.8	242.8	125,500.0	109,466.2	17,650.0	17,008.7

	DO		BOD		TN		TP	
	Dom	DOs	BODm	BODs	TNm	TNs	TPm	TPs
Total magnitude of violations	36.115	33.056	3502.204	3404.934	1061433	919881.55	109455	83438.081
Magnitude of max. viol.	5	4.937	599.0811	684.93585	193200	215239.13	29800	17008.674
Number of Violations	4	4	11	11	10	11	10	9

INDEX	DO		BOD		TN		TP	
	Dom		BODm		TNm		TPm	
Total magnitude of violations, ETS	0.085		0.0278		0.1334		0.238	
Magnitude of max. viol., EV	0.013		-0.143		-0.114		0.429	
Number of Violations	0		0.00		-0.1		0.1	

	DO	BOD	TN	TP
Overall Separate Performances	0.0973	-0.12	-0.08	0.77

OVERALL PERFORMANCE MEASURE	0.67
------------------------------------	-------------