UNIVERSITY OF NAIROBI

SCHOOL OF COMPUTING AND INFORMATICS

A GIS Tool for Water Quality Monitoring: A Case on Nairobi River.

BY

NDITHI, HENRY KIVUVA

P56/60449/2010

SUPERVISOR

MR ANDREW MWAURA

July, 2012

Submitted in partial fulfillment of the requirement of a Master of Science Degree in Information Systems of the University of Nairobi.

DECLARATION

I, **Henry Kivuva Ndithi** do declare that this research project is my original work and where there is work or contribution of other individuals, it has been duly acknowledged. To the best of my knowledge, this research work has not been carried out before or previously presented to any other educational institution for similar purposes. No part of this research report should be reproduced in part or whole without written consent of the author or that of the University of Nairobi.

Signature.. Date 25/07/2012 ....

Student: Henry K. Ndithi

P56/60449/2010

This research project has been submitted for examination with my approval as the University of Nairobi supervisor.

Signature.. Date 8/Aug/2012 ....

Supervisor: Andrew Mwaura.

Lecturer, School of Computing and Informatics, University of Nairobi.

ABSTRACT

Although many studies have been conducted to assess different aspects of pollution along the Nairobi Rivers, there is a need to collate this information in order to identify areas of concern and gauge the extent to which the rivers have been degraded. A Geographical Information System (GIS) tool was developed and used to construct thematic maps for water quality of the Nairobi Rivers Environmental data were integrated and an overall picture about the spatial variation in the water quality of the rivers was defined The water quality maps were derived from the results of previous studies for physico-chemical parameters like Total dissolved solids (TDS), Hydrogen Ions Concentrations (PH), Total Hardness (TH) and Nitrate (NO_3) concentration. The GIS maps showed not only contaminant distribution but also illustrated the need to improve the water quality management methods.

Using the developed GIS water quality monitoring tool, water quality specialists can concentrate on analyzing data and presenting their results without bothering about the details of the software application being used. The GIS tool was developed using Quantum GIS (QGIS). This is a free open source desktop GIS. The data being anlysed was stored in MS Access database, which was linked to the QGIS. A map of Nairobi River basin was sourced from the ministry of Environment and natural resources. This map was scanned, Georeferenced and digitized. The coordinates of the monitoring points along the river basin and the corresponding water quality data were stored in the database. The database was queried to generate data meeting specified criteria, e.g. TDS of the monitoring points in a specified period. The generated data was then interpolated using the Inverse Distance Weighting (IDW) interpolation method, to generate water quality maps for each of the parameters; TDS,PH,TH, and NO_3. A combined quality map was then generated by integrating the four thematic maps.

From this study, it was concluded that a multi-purpose GIS software tool can be developed and used to automate many functions of water quality analysis. Further work can be done to improve the tool by incorporating a menu to select all the quality analysis commands. Also a geospatial database, like POSTGIS, could be used instead of MS Access. In addition, a Global Positioning System (GPS) device could be used to accurately identify the coordinates of the monitoring points along the river basin.

Keywords: GIS, Water quality monitoring, Geospatial database, Phisico-chemical parameters, spatial interpolation.

DEDICATION

*To my mother, **Maritina Mbelengwa**, in MEMORIAM*

ACKNOWLEDGEMENT

My sincere gratitude goes to my supervisor **Mr. Andrew Mwaura** for his guidance, viable suggestions and priceless advice throughout my project work. These tremendously contributed to my success. May GOD bless him.

My Special gratitude goes to my dear wife ***Roseline***, My loving Children ***Sylvia***, ***Nelson*** and ***Martin*** for giving me ease of time and encouragement in the entire period of my study.

Finally, I owe everything to the GOD almighty for the precious gift of life and good health, without which nothing is possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFRALTI	African Advanced Level Telecommunication Institute
AHP	Analytic Hierarchy Process
DBMS	Database Management System
DFD	Data Flow Diagram
DSS	Decision Support System
EPA	Environmental Protection Agency
ER	Entity Diagram
GIS	Geographical Information System.
GPS	Global Positioning System
IDW	Inverse Distance Weighting
JPEG	Joint Photographic Experts Group
KICC	Kenyatta International Conference Centre
MCL	Maximum Contaminant Level
NEMA	National Environment Management Authority
NO_3^-	Nitrates
PH	Hydrogen ion Potential
QGIS	Quantum Geographical Information System
RS	Remote Sensing
SDI	Spatial Data Infrastructure
SQL	Structured Query Language
TDS	Total Dissolved Solids
TH	Total Hardness
UNEP	United Nations Environmental Programme
UON KST	University of Nairobi Kenya science Training
WHO	World Health Organisation

CHAPTER 1: INTRODUCTION

1.0 Background Information

Increasing spatial data sharing and interoperability throughout the world have resulted to availability of large amount of data to be used in development of spatial data infrastructures (SDI) in a rapid and unprecedented pace.

In spite of production of such a huge amount of spatial and thematic data, proper use and management of the data are important issues to achieve sustainable development. It has been clear that conventional approaches of statistical analyses of geospatial data in analogue form and in small amount cannot be efficiently implemented for digital data and in large volume, which are being produced nowadays. Therefore, conventional analysis approaches cannot be implemented to explore the hidden relationship between and among spatial data and their future trends, which are quite important functionalities in optimum geospatial data management. Such functionalities can be efficiently employed using geospatial knowledge discovery.

One of the important applications of GIS is environmental data management. GIS can be used to provide scientists and managers with a range of scenarios for spatial distribution of the data and predict future trends of the data to avoid possible environmental crisis. Geospatial data analysis can be used to assess hidden relationships of the crisis and environmental pollutions, sources, causes and amount of pollutions to take necessary measures for environmental protection.

In 1972 Clean Water Act established a National Pollution Discharge Elimination System (NPDES), which requires an easily revoke permit for any industry, municipality or other entity dumping wastes in surface water (Cunningham and Saigo, 1999).

Most of those efforts have been aimed at point sources, especially to build or upgrade thousands of municipality sewage treatment plants.

Attempts were made by Indian geographers to delineate 'Physical complexes' on the principle of relationship among the natural elements. One study consists of 16 macro and 58 microphysical complexes based on regional grouping of administrative districts (Singh, 1995). Such systems are considered as starting point in integrating the dimensions of environment with spatial monitoring and forecasting. This system had many problems in building a unified database due to increasing volume of source information.

In 1996, the U.S. Environmental Protection Agency (EPA) announced that toxic, chemical sewage or other pollutants contaminated about 16000 segments of surface water in U.S. and its territories. However, the method to compare this situation with past pollution levels and with the other countries, to make an efficient decision was not specified (Cunningham and Saigo, 1999).

Coordination of the above mentioned samples can be done in a GIS environment. Especially, when the volume of data is increased, geospatial data analysis is an appropriate candidate to extract more information from such a data warehouse.

1.0.1 The Importance of Water Resource Monitoring.

Water is one of the most important requirements in daily life. Implementation of proper and practical policies to evaluate the water resources through integrated exploitation, management and planning is vital. This is to ensure the availability of enough qualified water for the whole planet in addition to maintain the hydrologic, biologic and chemical ecosystem in a proper way.

Considering the vital role of water resources in national, regional and global ecosystems, assessment of existing situation and evaluation of the potentials of available water resources are vital steps for proper water resource management.

The existence of a number of environmental effects such as increasing the earth temperature has changed the spatiotemporal distribution of rain. This will lead to desertification, flooding, etc. (Cunningham and Saigo, 1999). The relationship between development and environment, population increase and the importance of maintaining food security are among the important challenges in water demand and supply management. Therefore, water supply and quality protection have been considered in Agenda 21 in Rio Conference in 1992 (www.un.org).

The importance of water quality management has forced a number of countries to investigate ways for pollution control. In this direction, GISs are among the most useful approaches. Recently, many efforts have been undertaken to use GISs for water quality assessment and monitoring in different scales such as streams, rivers, lakes, seas and oceans. For example "The Alabama Watershed Demonstration" project links land use patterns and water quality through GIS (Flynn, 1999). Also some successful efforts about satellite and GIS tools to assess lake quality have been reported in University of Minnesota (Brezonik et al., 2002).

In some previous researches predefined indicators and relationships were considered and GISs were used to manage these situations. However, in some cases no certain relationships between parameters and their extraction have been directly reported. In such situations, some statistical analyses can be used which lead us to geospatial data mining.

1.1 Problem Statement and Purpose of Project

Urban rivers, streams and wetlands are prone to pollution (Natumanya et al, 2009) and deterioration of water quality reduces a river's amenity and aesthetic value.

Although numerous studies have been conducted assessing different aspects of pollution along the Nairobi Rivers (e.g. Issaias 2000, Mwathi et al. 1997, Nyikuri 1994, Ohayo-Mitoko 1996) there is a pressing need to collate this information in order to identify areas of concern and to gauge the extent to which the rivers have been degraded.

1.2 Proposed Solution

The proposed study developed a software tool to summarize large amounts of spatially and temporal distributed data in an intelligent form. The system will allow water quality experts to select monitoring stations, along the Nairobi river basin, interactively and analyze the chemical data. The results will be reported symbolically on a map .The tool will provide scientists, managers and policy makers with a range of scenarios of data distribution to help them predict future trends; to avoid possible environmental crisis and disasters

1.3 Objectives

1.3.1 General objective

The overall objective was to develop a GIS based software tool for use in monitoring water quality along the Nairobi river basin.

1.3.2 Specific objectives

1) To develop a GIS database to store digital map of the Nairobi River basin and associated water quality data.
2) To customize the GIS application for water quality monitoring, for easy analysis and retrieval of information.
3) To test the performance of the GIS water quality monitoring tool using data acquired from NEMA, UNEP, published scientific research papers and/or other reliable sources.

1.4 Research questions

1) What parameters determine the water quality in River basins?
2) What is the structure of the database used to store water quality data along a river basin?
3) Which GIS is best suited to analyse and generate water quality maps?
4) What is the best format for displaying the geospatial pollution levels along a River basin?

1.5 Justification

Pollution of water bodies by the ever increasing human activities (agriculture, discharge of industrial and domestic wastewaters, solid waste disposal, runoffs, etc) is a serious problem (Chapman, 1996). In 1996, the U.S. Environmental protection Agency (EPA) announced that toxic, chemical sewerage and other pollutants contaminated about 16000 segments of surface water in U.S. and its territories. However, the method to compare this situation with past pollution levels and with other countries, to make efficient decisions was not specified (Cunningham and Saigo, 1999). Coordination of these mentioned samples can be done in a GIS environment.

1.6 Scope of study

This was a case study of the Nairobi River.

Secondary Data from different sampling sites along the river basin was analyzed, based on previous studies. Pollution data on the Nairobi water basin was sourced from published scientific papers, in the period 1999 up to 2010. Maps of the river basin were sourced from National Environmental Management Authority (NEMA), the ministry of environment and national resources, and from survey of Kenya.

1.7 Assumptions and limitations

1.7 .1 Assumptions

The study assumed that:

1) Data on pollution levels for Nairobi Rivers was available and accessible to the researcher.
2) Maps of Nairobi river basin were available from either National Environmental Management Authority (NEMA) or the ministry of environment and national resources or from survey of Kenya.
3) The secondary data on pollution along the Nairobi river basin were not poorly collected, inaccurate, or flawed.

1.7.2 Limitations

The following limitations were anticipated:

1) Since we are using secondary data, there is no way of going back for further information.
2) The researcher will keep the analysis within the boundaries of the originally collected data

1.8 Definition of important terms

a) **Geospatial information**: is data referenced to a place. A set of geographical coordinates which can often be gathered, manipulated, and displayed in real time.
b) **Geographical information system (GIS)** : is a computer system capable of capturing , storing , analyzing and displaying geographically referenced information.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

In the past century, the natural environment has provided various types of resources in support of rapid industrialization and urbanization. As the world's population grows over time, human beings have progressively made greater demands on environmental resources through an unprecedented increase in technology capacity, energy consumption, international trade, and social complexity. Information technologies are becoming more and more important for environmental management, due to increasing need for large-scale computational capability in order to handle the sophistication of environmental decision-making. To explore ultimate limitation of the nature, an analysis of the environment in terms of physical, chemical, and biological processes and their interactions is becoming critical. Significant efforts are required to analyze relevant information, simulate related processes, evaluate resulting impacts, and generate sound decision alternatives. System-based approaches developed in the past two decades have enabled us to investigate complex interactions fundamental to the co-evolution of engineered and natural systems. Recent advances in information technology lean towards making effective search for sustainable development strategies via integrative efforts between multi-dimensional, multi-scale data analysis and environmental system modeling. This could facilitate decision-makers to intimately link the domain knowledge with envisioned social, economic, ecological, and environmental objectives, leading to a new interdisciplinary field, environmental informatics. This new field brings together a variety of information-technology-based measures, in connection with versatile environmental monitoring networks and in association with multidisciplinary mathematical modeling skills to provide risk-informed, consensus-oriented, and cost-effective solution (Chang et al., 2001, 2002).

Traditional mathematical simulation models are useful tools for the forecasting of environmental processes. For example, Li and Chen (1994) proposed a model for simulating organics removal and oxygen consumption by biofilms in an open-channel. Masliev and Somlyody (1994) advanced a probabilistic method for uncertainty analysis and parameter estimation for dissolved oxygen models. Kazmi and Hansen (1997) developed a numerical model for water quality simulation and applied it to a case study in the Yamuna River, India. While state-of-the-science models characterizing the fate and transport of contaminants in different environmental compartments or medias are indeed necessary to rigorously understand the short-term and long-term dynamics of pollutant behavior, they may be of limited value in finding out the casual effect and cost-benefit relationship for immediate policy planning and regulatory studies. System dynamics model exhibits promising potential to assess feedback mechanisms for identifying system response from a broad sense.

Besides, optimization techniques have been widely used in the field of environmental management and pollution control. The results provide basis for making decisions related to allocation of waste loadings, deployment of monitoring network, and implementation of pollution abatement activities (Chen and Chang, 1998). Chang et al. (1997a) and Chang and Wei (1999) developed a multi-objective mixed-integer programming, which incorporated with geographical information system (GIS) for routing and scheduling collection vehicles in solid waste management systems. Alidi (1998) utilized a goal-programming model to aid in the integrated waste management, using the analytic hierarchy process (AHP) technique for determining the weights and priorities for a given set of goals. Such kind of extension work does enrich the application spectrum in the context of systems analysis.

Integrated modeling systems with the aid of simulation, regression, and optimization analyses have utilized existing disciplines common in operations research and management science applications to design various environmental management systems for different study regimes (Yen et al., 2003; Yen and Chang, 2003). They may significantly help address the forcing of human-induced impacts, identify the responses in the environmental systems, and assess consequence due to such disturbance in our society. Overall, it enables scientists, engineers, and managers to project consequences of management alternatives, provide insightful planning, and formulate environmental policy such that effective decision-making schemes can be identified. The key challenge is how to make complementary use of models with different features, scales, and complexity as well as data collected from multiple types of sensors to pursue a full understanding of the air-sea-land-biosphere interaction mechanisms under the impacts of infrastructure operation, resources consumption, and global change. These modeling frameworks themselves have to be highly modularized and are adapted to multiple computation platforms in dealing with various types of issues in environmental systems. In an attempt to find a balance between competing social, economic, ecological, and environmental factors in the context of sustainable development, seamless integration of soft information and quantitative results obtained from integrative modeling studies may exhibit the beauty of environmental systems analysis.

2.1 Challenges of Environmental Systems Modeling

This goal, however, involves several challenges. The first challenge is the characterization of uncertainties that exist in many intertwined system parameters that could make environmental systems extremely complicated. Applicability of modeling techniques to environmental management is affected by many factors. First of all, environmental systems are complicated, where some factors and interrelationships are hard to be expressed as mathematical formulas. For example, nonlinearity that exists in a system can hardly be effectively reflected. Secondly, information about some system parameters is often unavailable, such that rough estimations have to be made. Also, a large portion of available information may not be quantifiable. This type of information could simply be the implicit knowledge from decision makers Instead. Thus, the input into a modeling system may only be a small part of the entire information in a study system. Consequently, the modeling output is inadequate to support decision-making. The remaining part of the work should be a solid investigation on ambiguous and unquantifiable information using innovative information technologies. Thirdly, a significant part of quantifiable information may not exist as deterministic data. This brings about the difficulty in uncertainty expression, as well as solving the models that contain uncertain parameters and/or relationships (Sasikumar et al., 1999). Many researchers tried to deal with the uncertainties through the inclusion of fuzzy, stochastic and other inexact modeling approaches in the context of optimization analysis (Sasikumar et al., 1999). In particular Huang et al. (1996) proposed inexact optimization models for watershed environmental planning, and applied them to two real-world case studies. Another challenge is the quantitative description of how the risk is involved in decision-making related to the uncertainty in modeling process. While risk involved in environmental assessment might become another source of uncertainty, uncertainties in environmental systems may cause a certain level of risk that might affect the final decision in environmental management. Consequently, risks and uncertainties associated with a variety of system behaviors, objectives, and their interrelationships have received significant attention from both environmental management and information science (Beck, 1987). An associated challenge is to develop the capability of minimizing the uncertainties and risks using advanced information technologies. To overcome the challenges and to enhance model feasibility and applicability, there is a need for incorporating current information processing techniques with up-to-date monitoring and measurement technologies to aid in environmental systems modeling.

2.2. Impacts of Environmental Informatics

In the past decades, a number of computer-based modeling techniques were developed for studying environmental management systems and providing related decision supports (Chang and Wang, 1996; Chang et al., 1997a,b). Especially, many comprehensive decision support systems (DSS) were designed and applied to real-world problems. Such computer-based systems have interactive, graphical, and dynamic characteristics and can be directly used for addressing specific management issues and assisting individuals in their problem-solving processes (Soncini Sessa et al., 1999). The strategic effort in this regard is to provide scientific answers to overarching questions in cases when early warning, special operation, and emergency response need to be particularly taken into account.

Nowadays, information technologies are seen to play a major role in sustainability-based decision-making processes. A typical computer-based technology that has been widely used in assisting environmental systems analysis is GIS. GIS is effective in handling complicated spatial information that is essential for many environ mental studies, as well as providing platforms for integrating various models, systems, and interfaces (Huang et al., 1999). In the past decade, many GIS-aided environmental modeling and decision-support systems have been developed (Huang et al., 1999). The geo-coding exercises carried out by many agencies over the last few years, coupled with advances in geo-information processing tools have made available large volumes of datasets that can be used during environmental decision making process. Remote sensing (RS) is another important computer-based technology for supporting environmental systems modeling to perform systems analysis (Goksel, 1998). Space-borne, air-borne, and ground-borne remote sensing technologies vastly supplement the ground-based sampling scheme in the context of environmental monitoring and measurement. Most of RS projects produce large volumes of spatial information, while GIS is an effective tool for storing, manipulating and analyzing them. Consequently, a number of integrated environmental modeling and RS-GIS studies have been reported. Recent advances in the technical integration of GIS and RS in connection with global position system (GPS) and database management systems (DBMS) successfully streamline the information flows among stakeholders (Atkinson and Tate, 1999).

Associated with recently renewed interest in utilization of information technologies in environmental studies, several fundamental and applied aspects related to applications of new technologies draw much attention from the scientific community. Firstly, there has been a unanimous recognition of the importance of OpenGIS® defined by open interfaces and protocols among heterogeneous computer platforms. Specifications support interoperable solutions that "geo-enable" the wireless and location-based applications throughout the Internet environment. This advance could allow remote users to share enormous amount of spatial data across various computer platforms. Secondly, recent advances in high performance computing have shown great potential to improve the prediction accuracy in the practice of environmental systems modeling. High performance computing is needed when we have to assess large-scale environmental changes. Thirdly, large-scale database search in conjunction with artificial intelligence techniques, such as artificial neural networks, fuzzy reasoning, knowledge-based expert systems and other data mining tools, also have the significant potential to be of much use in environmental decision-making schemes. Developing a valuable decision support system is an inherently collaborative endeavor involving a diverse group of people with different backgrounds, values and experiences. Information technologies, such as knowledge acquisition, data mining, uncertainty analysis, and expert system technologies, could be helpful to increase data integrity and reliability of decision-making. Figure 2.1 describes the integrative structure between environmental systems modeling, environmental monitoring and measurement, and environmental informatics. Advances of sensor synergy skills and telecommunication technology make large-scale ground-based sampling scheme feasible. The main challenges associated with such an effort rest upon model synthesis when dealing with various features, scales, and complexities. They include handling connection among various simulation, optimization, and assessment models as well as the related information technologies and platforms, linkage between inputs and outputs of various technologies, quantification of socio-economic factors, and solution procedure for the resulting large-scale integrated models. These difficulties have affected practical applicability of the integrated approach (Vijayan et al.,1999). Confronting these complexities, it is essential to gain insight into integrative efforts in order to identify effective approaches to overcoming or mitigating the challenges.

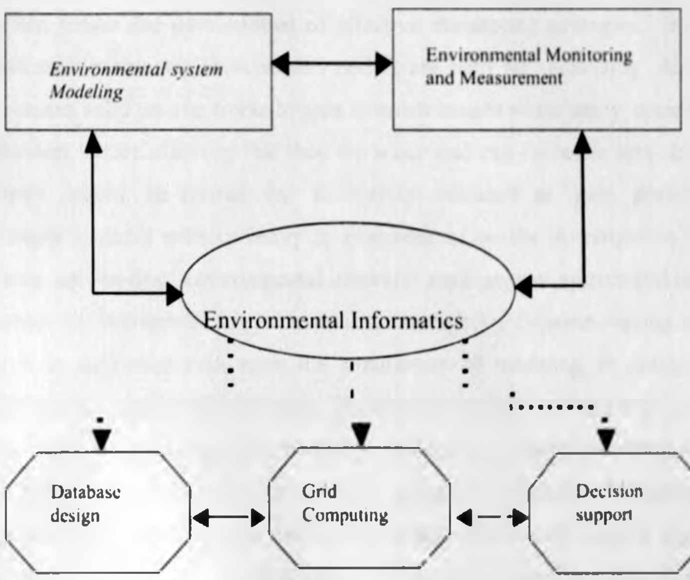


Fig. 2.1 Components of environmental informatics and their interactions

2.3 Modern Environmental Decision Making.

Many challenges exist in the applications of modeling techniques to environmental management. Most of environmental models can only deal with limited spatial and temporal units in a system due to difficulties in computational requirement and data availability. However, what decision makers desire to know might be either detailed plans based on much finer units or just a broad justification. This could lead to incompatibility between the researcher's outputs and the user's demands, and raise the question about usefulness of modeling solutions. Moreover, the collection of environmental statistics is fraught with difficulties, due to wide range of environmental phenomena, data sources, and agencies involved, as well as the complexities of their temporal and spatial characteristics (Briggs, 1995). Consequently, many environmental data are subject to serious discretion in regards to uncertainties, inconsistencies, and errors. To obtain improved reliability and certainty, solid works on validation of input-data prior to being used for further analysis are desired, where information technology could take crucial roles. The insufficiency of data about pollution sources, mitigation measures, natural conditions, and environmental quality records, as well as the lack of information relevant to cultural, social, economic and political factors often hinder the development of effective monitoring strategies. Inevitable on-site investigation could ease such problems. Thus, when models are used for providing decision support, researchers have to conduct solid on-site works to gain as much insight of the study system as the local managers and stakeholders before claiming that they are wiser and can do better jobs. Involvement of information technologies would be desired for facilitating research to gain profound knowledge and understanding of study systems without heavy involvement of on-site investigation. This motivates many agencies to build up on-line environmental database management system and decision support system to aid in various environmental impact assessment and policy decision-making under a fast and friendly environment. It is suggested that, upon the completion of modeling or design based on the available data, the entire job is merely halfway done. The remaining half is to examine how information that are unavailable but may present as implicit knowledge of decision makers or stakeholders could be collected by various remote sensing instruments and acquired through innovative information technologies such as artificial intelligence and data mining. This would require high performance computation in most cases. Figure 2.2 demonstrates a typical system configuration using 4S information technology concept (GIS/GPS/RS/DBMS) that enables us to facilitate various spatial analyses, data sharing and distribution, and modeling assessment. This computer node could become one of the OpenGIS and/or grid computing smart nodes to share essential data and knowledge distributed over the internet to achieve a high-end sophisticated knowledge management goal or support vast computational capacity for performing share-vision modeling analysis (Chen et al., 2003). When using this kind of DSS to help solve environmental problems, the final difficulties might include how system can be sustained and how all system components can be comprehensively formulated and maintained in the information arena.

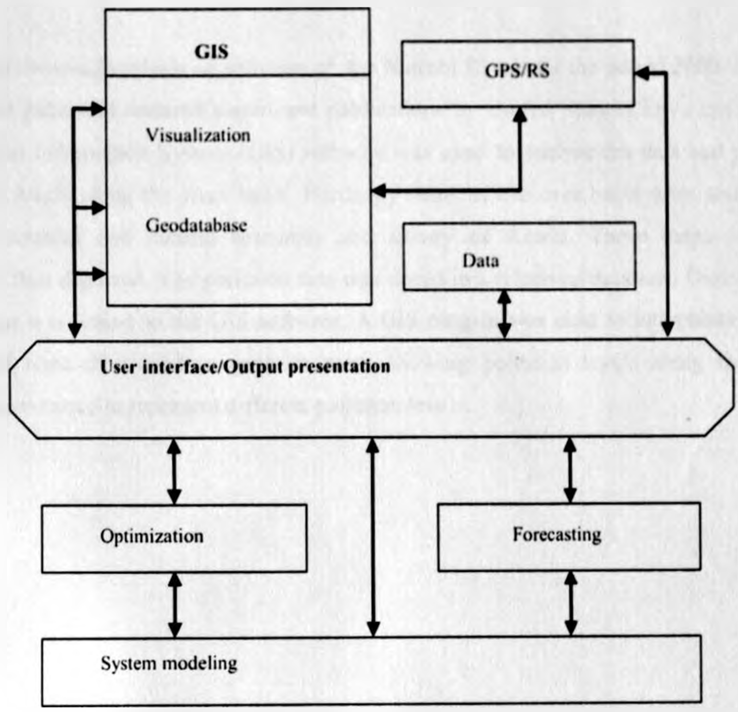


Fig 2.2 Outline of a computing system for environmental decision support

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Overview

The research study involved analysis of pollution of the Nairobi Rivers for the period 2000-2010. The data was collected from published research papers and publications by United Nations Environmental program (UNEP). Geospatial Information System (GIS) software was used to analyse the data and generate maps showing pollution levels along the river basin. Hardcopy maps of the river basin were sourced from the ministry of environment and natural resources and survey of Kenya. These maps were scanned, georeferenced and then digitized. The pollution data was stored in a relational database. During the analysis stage, this database was linked to the GIS software. A GIS plug-in was used to interpolate the data. The generated resulted were then used to generate maps showing pollution levels along the river basin. Different colours were used to represent different pollution levels.

3.2 System Design.

3.2.1 System Context diagram

Fig 3.1 shows a context diagram of the GIS tool

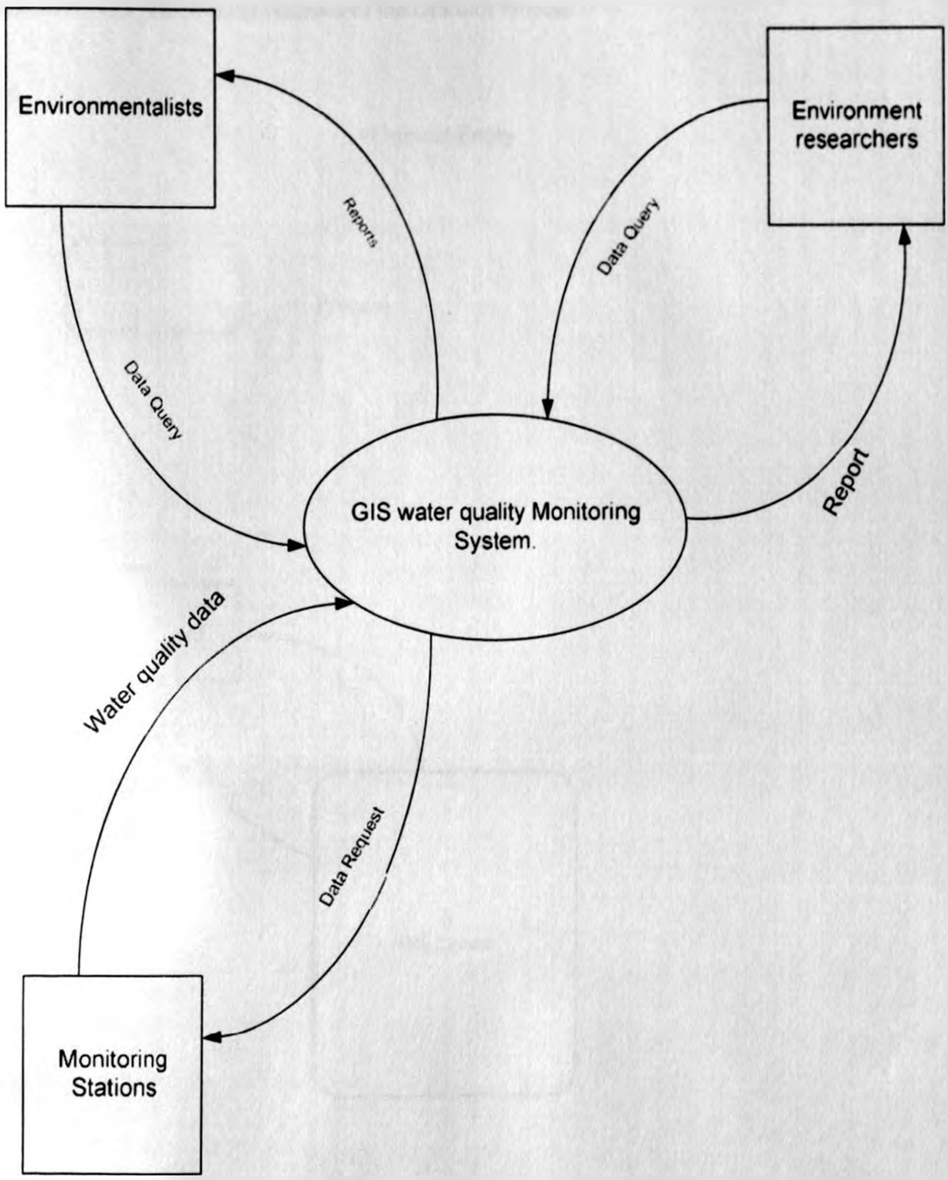


Fig 3.1 GIS tool context diagram.

This top-level diagram or level 0 diagram models the whole system as a single process box whose sides represent the boundary of the system. It identifies all external entities and related input and output flows.

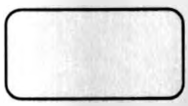
The GIS water quality monitoring system will interact with monitoring stations, environmentalists, and researchers. All these will consume data and also generate data for the system.

3.2.2 System Data Flow diagrams (DFDs).

Fig.3.2 shows a level 0 DFD diagram of the GIS tool Process.

Key

=External Entity



= Process

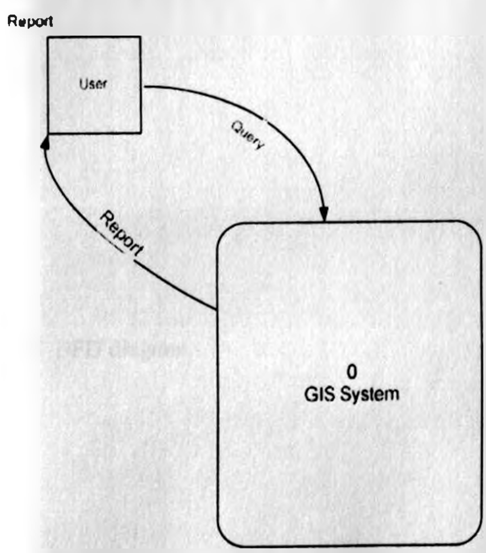


Fig. 3.2 Level 0 DFD diagram

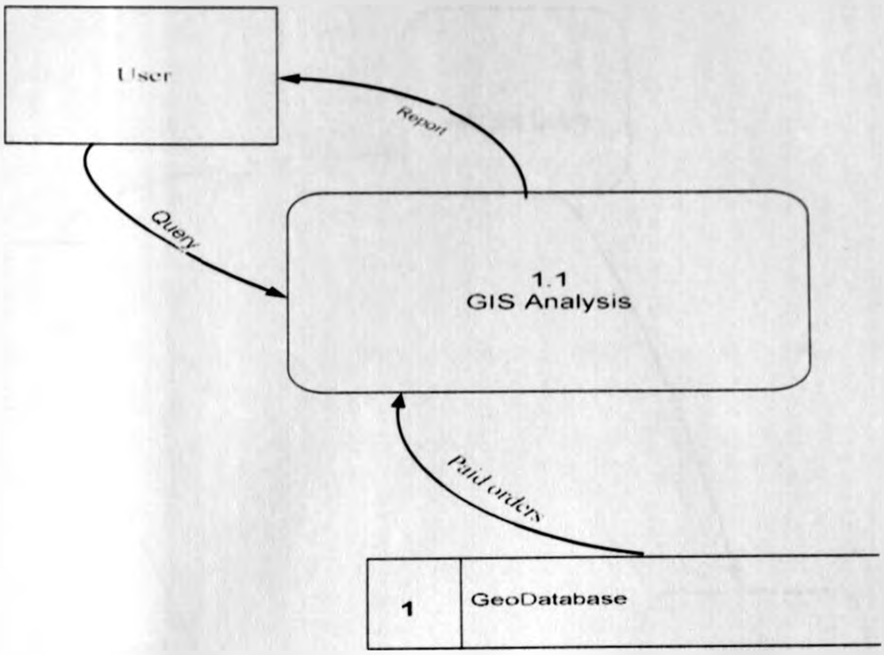


Fig. 3.3Level 1 DFD diagram

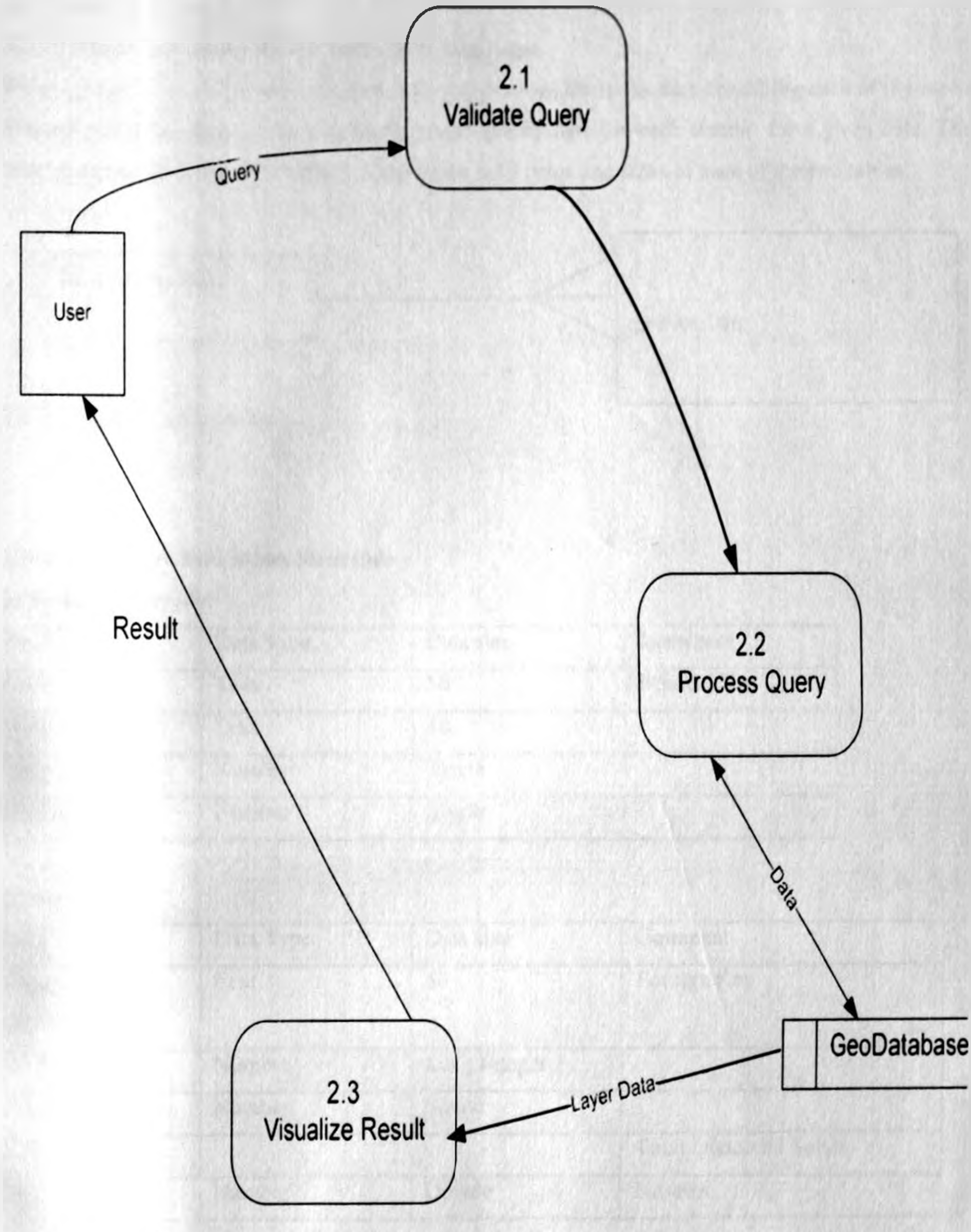


Fig. 3.4Level 2 DFD diagram

3.2.3 Data Base Design.

3.2.3.1 Data Base Entity Relationship (ER) Diagrams.

Two database tables were created ; StationDescription- which holds data describing each of the monitoring Stations and StationData -which holds the water quality data for each station for a given date. Their ER diagram is shown in fig. 3.5. Table 3.1 shows the field types and sizes of each of the two tables.

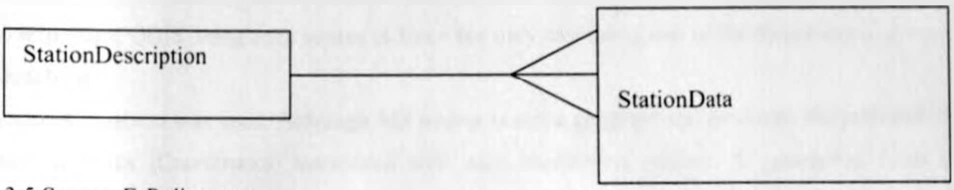


Fig 3.5 System E-R diagram

Table 3.1: Data Base tables Structure

a) StationDescription

Field name	Data Type	Data size	Comment
Code	Text	50	Primary Key
StationName	Text	50	
Latitude	Number	Single	
Longitude	Number	Single	

b) StationData

Field name	Data Type	Data size	Comment
StationIdentity (STID)	Text	50	Foreign Key
Year	Number	Long Integer	
PH	Number	Single	
TDS			Total Dissolved Solids
NO3	Number	Double	Nitrates

3.3 System Implementation.

3.3.1 Hardware and Software Choices.

a) Geospatial Information System (GIS)

The development platform was Geospatial information System (GIS) using Quantum GIS (QGIS). This is a free open source desktop GIS application that provides data viewing, editing, and analysis capabilities. Also important, QGIS being open source is free – the only cost being that of the bandwidth to download it.

b) Database

MS access database was used. Although MS access is not a geographical database, the pollution data had geospatial fields (Coordinates) associated with each monitoring station. A connection from the GIS software was established to the database during data analysis.

3.3.2 Data Processing

3.3.2.1 Database Creation.

A database was created to store the water quality parameters. MS Access database was used. This is a desktop relational database management System (DBMS). This was chosen because it is a readily available and most of the targeted users are familiar with it.

The wizard facility in MS-Access was used to readily create the database tables and their relationships. One main form was created for entering the parameters of the monitoring stations along the Nairobi river basin viz station name, station code, latitude and longitude. A sub-form was created to enter the water quality data for each selected station. The database was queried using SQL statements to generate views (Virtual tables) meeting specified criteria.

3.3.2.2 Map Preparation

A hardcopy map of Nairobi and its environs was sourced from the ministry of environment and Natural resources. This map covered the River basin and the surrounding areas; where the river pollutants emanate. The map was scanned using a digitizing tablet to create a JPEG image.

3.3.2.3 Georeferencing

The scanned image of the Nairobi river basin was not in a format aligned to real world coordinate system. It was necessary to align this image with existing geographically referenced data. This process is called georeferencing.

a) Reference Points

The first stage in the georeferencing process was to choose a number of reference points on our map. Six control points were identified; these were UON KST campus, Baba Dogo primary school, Railway training Institute, AFRALTI and KICC. The geographical coordinates (Latitude, Longitude) of each of these points were determined using Google map. The choice of the points was such that they are well distributed throughout the scanned image. Table 3.2 shows these reference points and their coordinates.

Table 3.2: Reference Points

Control Point Name	Latitude	Longitude
UON KST Campus	-1.30069076	36.7642199884
Baba Dogo Primary School	-1.2433376	36.8862523999
Rail Way training Institute	-1.3166504	36.8445383000
Afralti	-1.2634089	36.7569765000
KICC	-1.288134	36.821999

b) **Georeferencing Process**.

The following steps describe the major steps to georeference an image.

i. The georeferencing plug-in is invoked from the plug-in manager.
ii. The map image being georeferenced is uploaded and the right co-ordinate system defined (e.g WGS 1984 datum)
iii. The scanned image is linked to the ground control points.
iv. The georeferencing plug in is executed. A properly georeferenced map will be generated. The resulting image is saved.

Selecting more ground reference points increases the accuracy of the resulting map.

Table 3.3 shows the control points used to test the accuracy of the georeferenced map and the corresponding results, by comparing their map coordinates and their corresponding Google map coordinates.

Table 3.3: Control points

Control Point Name	Georeferenced coordinates (Long, Lat)	Google Map Coordinates (Long, Lat)
Kamukunji Police Station	(36.8334, -1.28530)	(36.833376, -1.28447)
Norfolk Hotel, NRB	(36.81553, -1.27210)	(36.816667,-1.28333)
Strathmore High School	(36.77573, -1.27210)	(36.777238, -1.270186)

3.2.2.4 Digitisation

Digitisation is the process of making features we can see on the map image editable. It also makes these features to be assigned additional spatial and non-spatial attributes. This means making versions of objects which have an attribute table associated with them; hence are vectors.

There are three types of digital features: point, line and polygon. Their creation is normally guided by a background map layer.

The point layer was used to represent the monitoring points, towns etc. The line feature represented the rivers and their tributaries. The polygon feature was used to create dams, forest covers, marshes etc.

In order to create and store the features to digitize, we first created a new layer. Then a choice was made on what vector type to display on the newly created layer.

The benefits of digitizing features are that being vector features means:

- Symbolizing features in any way deemed.
- easily edited to reflect changes
- Ease in making measurements e.g. quantity (number of features), length (e.g. of roads), area (e.g. Forests coverage).

3.3.2.5 Interpolation.

This is the process of using a sample set of point locations to create a raster surface (or a Map) based on the data attribute values. To do this we used the inverse distance weighting (IDW) interpolation method. Surface interpolation is used when you have data for certain locations and then you want to estimate the values in-between each location. The output is a raster layer that contains a value everywhere within the extent of the data.

With IDW method, a weight is attributed to the point to be measured. The amount of this weight is dependent on the distance of the point to another unknown point. These weights are controlled on the bases of power of ten. With increase of power of ten, the effect on the points that are farther diminishes. Lesser power distributes the weights more uniformly between neighbouring points. In this method the distance between the points count, so the points of equal distance have equal weights (Burrough and MacDonnell, 1998)

The weight factor is calculated with the use of the following formula;

$$\lambda_i = \frac{D_i^{-\alpha}}{\sum_{i=1}^{n} D_i^{-\alpha}}$$

Where λ_i = the weight of the point, D_i = distance between point i and the unknown point .

α = the power ten of weight.

3.4 Criteria for Acceptability and Rejection in Water Quality Analysis.

The criteria for suitability and non-suitability of the water in the Nairobi Rivers were determined. This was performed based on the water quality standards stipulated by the World health organization (WHO). Ranks were assigned for each parameter depending on the respected values, as shown in table 3.4.

Table 3.4: WHO Criteria for acceptability and Rejection in water quality analysis.

S/N	Parameter	Rank	CRITERIA	REMARKS
1	TDS	1	<500	DESIRABLE
		2	500- 1000	ACCEPTABLE
		3	>1000	NOT ACCEPTABLE
2	TH	1	<500	DESIRABLE
		2	500- 1000	ACCEPTABLE
		3	>1000	NOT ACCEPTABLE
3	NO_3^-	1	<45	DESIRABLE
		2	45- 100	ACCEPTABLE
		3	>100	NOT ACCEPTABLE
4.	PH		6.5- 8.5	RECOMMENDED LEVELS.

3.5 Generating the Water Quality Maps

Four quality maps for the parameters of PH, NO_3-, TDS and TH were created using the QGIS software. A final integrated quality map was created by overlaying these four thematic maps; which were produced using the inverse distance weighted (IDW) Interpolations.

Three areas where delineated within the study area based on the quality of water as portable, portable in absence of better alternate sources and non portable zones.

CHAPTER 4: TEST RESULTS AND DISCUSSIONS.

4.0 Introduction

To test the developed GIS water monitoring Tool, data from twelve stations along the Nairobi River were used. The stations together with their Geo Coordinates are shown in table 4.1

Table 4.1: Monitoring Stations

Sample Station Description

Code	Name	Lat	Long
001R	Ondiri	-1.24883	36.65729
002R	Dagoretti_Bridge	-1.28033	36.70969
003R	Naivasha_Bridge	-1.26892	36.75198
004R	J_Gichuru_Bridge	-1.26934	36.7709
005R	Museum_Hill	-1.275431	36.81243
008R	Kamukunji_Bridge	-1.28739	36.84199
009T	Kariobangi_North_Bridge	-1.2496	36.87821
010R	Kariobangi_South_Bridge	-1.2621	36.88237
011R	Carnivore_Bridge	-1.30962	36.8223
012T	Ruai_Bridge	-1.307375	36.88899
013R	Dandora_SW	-1.256874	36.88643
015R	Nairobi_Falls	-1.23835	36.91091

4.1 Test Data

Table 4.2 shows the test data used to test the developed tool. This was sample data extracted from Appendix 1

Table 4.2: Pollution Data

Pollution Data								
Year	Code	Name	Lat	Long	PH	TH(mg/l)	TDS(mg/l)	NO3(m/l)
2000	001R	Ondiri	-1.24883	36.65729	5.3	77	479	94
2000	002R	Dagoretti_Bridge	-1.28033	36.70969	6.4	51	448	184
2000	003R	Naivasha_Bridge	-1.26892	36.75198	7.2	73	843	2253
2000	004R	J_Gichuru_Bridge	-1.26934	36.7709	7.1	99	995	3804
2000	005R	Museum_Hill	-1.275431	36.81243	7.4	98	966	4077
2000	008R	Kamukunji_Bridge	-1.28739	36.84199	6.9	79	1213	838
2000	009T	Kariobangi_North_Bridge	-1.2496	36.87821	6.8	125	1344	201
2000	010R	Kariobangi_South_Bridge	-1.2621	36.88237	7.7	51	1133	184
2000	011R	Carnivore_Bridge	-1.30962	36.8223	7.5	95	1310	94
2000	012T	Ruai_Bridge	-1.307375	36.88899	8.7	143	1845	161
2000	013R	Dandora_SW	-1.256874	36.88643	8	98	1391	158
2000	015R	Nairobi_Falls	-1.23835	36.91091	7.7	115	1495	301

4.2 Findings

Water quality maps are useful in assessing the usability of water for different purposes figures 4.1, 4.2, 4.3, 4.4 and 4.5 shows spatial distributions of PH, nitrates, Total hardness, and total dissolved solid distribution concentrations and the combined water quality map along the Nairobi Rivers, respectively. The GIS tool created water quality maps for each parameter following the classification show in Table 3.4.

The result obtain from the maps can be summarized as follows

4.2.1 PH.

The results of figure 4.1 shows that the rivers were acidic upstream and moderately basic downstream. The increase in PH can be attributed to organic pollution and waste discharge draining into the river system.

The maps clearly show that the PH lies within the national water courses standards by WHO for aesthetic quality i.e. 6.5-8.5.

Only a small region around the Ruai Bridge had PH levels outside the recommended levels.

4.2.2 Nitrate(No_3) Distribution

The main source of nitrate in water is from atmosphere, legumes, plant debris and animal excreta. The maximum contaminant level (MCL) for nitrate is given as 45mg/l by the WHO for drinking water. Spatial distribution of nitrate concentrations for the study area is shown in fig.4.2

Only a small region around the Ondiri region has the water with the acceptable range of 45-100mg|l. The rest of the region has nitrate concentration outside the acceptable range.

We can say empirically that the nitrate distribution in the study area are above the prescribed limits and are not portable and hence the water requires to be processed before use.

4.2.3 Total Hardness (TH) Distribution.

Calcium and magnesium mostly cause the hardness of water. Total hardness of water may be divided into two types carbonate or temporary and bicarbonate or permanent hardness. The hardness produced by the carbonates of calcium and magnesium can be virtually removed by boiling the water and is therefore called temporally hardness. The hardness caused by the sulphates and chlorates of calcium and magnesium cannot be removed by boiling and is called permanent hardness.

Total hardness (TH) is the sum of the temporary and permanent hardness. Water that has a hardness of less than75mg|l is considered soft. A hardness of 75 to 150mg|l is not objectionable for most purposes.

The maximum allowable limit of TH for drinking purposes is 500mg|l.The water along Nairobi rivers basin is below the 500mg|l limit hence is desirable as per the WHO limits and is suitable for uses domestically or industrially.

The map of fig.4.3 shows the spatial distribution of TH along the Nairobi River Basin.

4.2.4: Total Dissolved Solids (TDS) Distribution.

The mineral constituents dissolved in water constitute dissolved solids. The concentration of dissolved solids in natural water is usually less than 500mg/l, while water with more than 500mg/l is undesirable for drinking and for many industrial uses.

Water with TDS less than 300mg/l is desirable for dyeing of cloths and manufacture of plastics, pulp paper e.t.c. The total concentration of dissolved minerals is a general indication of the overall suitability of water for many types of uses. Water with high dissolved solid content, would therefore be expected to pose problems like taste, laxative and other associated problems with the individual minerals.

The map of Figure 4.4 shows the spatial distribution of TDS along the Nairobi Rivers. At the upper regions, near the source, the TDS is within the WHO recommended ranges of less than 500mg/l. As we move downstream, the amount of TDS increases gradually and beyond the Museum Hill, the amount of TDS is in the unacceptable ranges; beyond 1000mg/l.

Water with more than 1000mg/l of dissolved solids usually gives disagreeable taste or makes the water unsuitable in other aspects.

TDS can be removed by reversed osmosis, electro dialysis, exchange and solar distillation processes.

4.2.5: Combined Quality Map

Figure 4.5 shows the combined water quality map that was produced by integrating the four maps for PH, TDS, Total Hardness and NO_3^- . The spatial integration was carried out using the QGIS IDW plug in. It can be seen that a large portion of Nairobi river waters is not portable. Only the portion of the river basin near the source and the area surrounded by Karura forest have water which is only portable in the absence of better alternatives.

Therefore it can be seen from the generated maps that most of the Nairobi river waters requires processing before use.

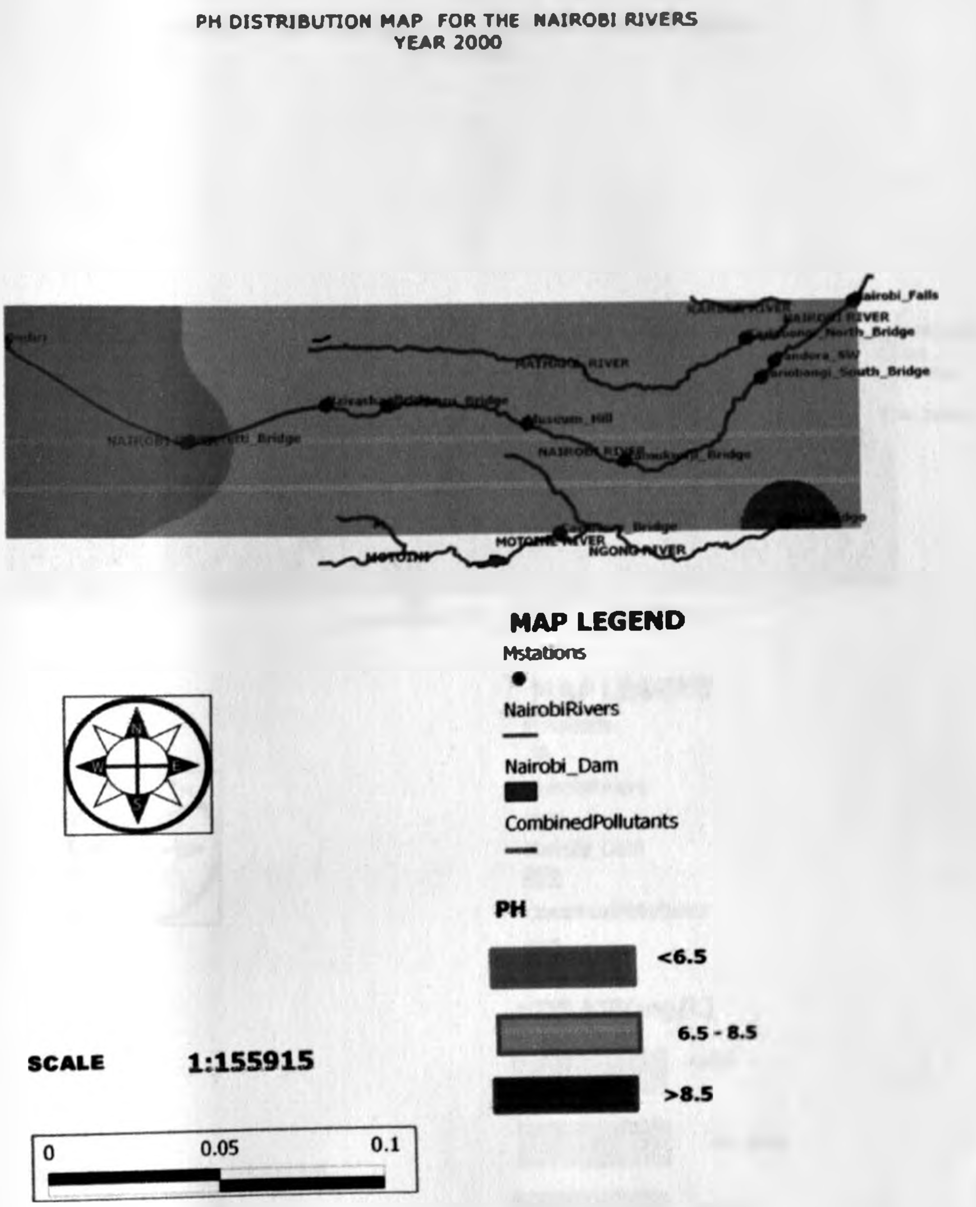


Fig. 4.1 PH Spatial Distribution in the Nairobi River Basin.

NITRATE DISTRIBUTION MAP FOR THE NAIROBI RIVERS
YEAR 2000

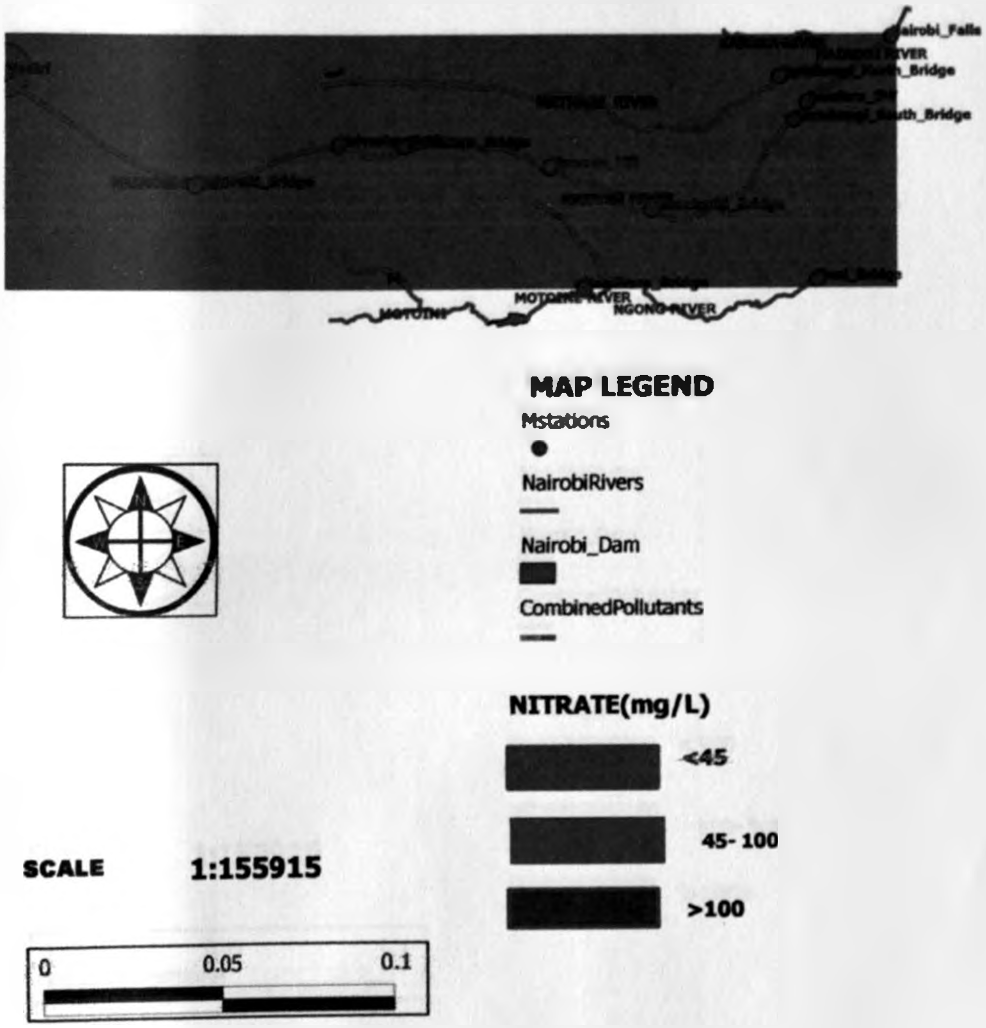


Fig. 4.2 Nitrate Spatial Distribution in the Nairobi River Basin.

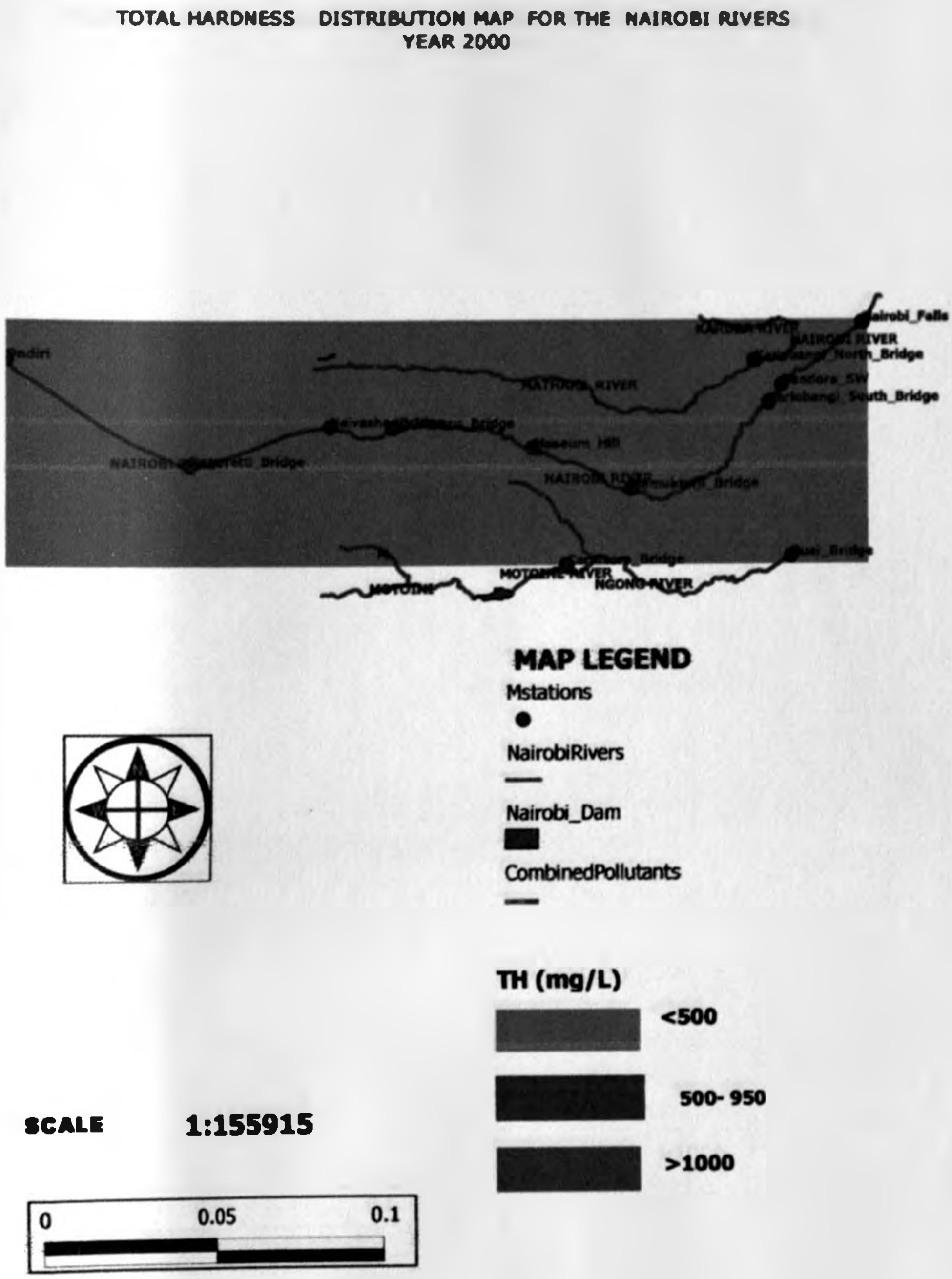


Fig. 4.3 TH spatial Distribution in the Nairobi River basin.

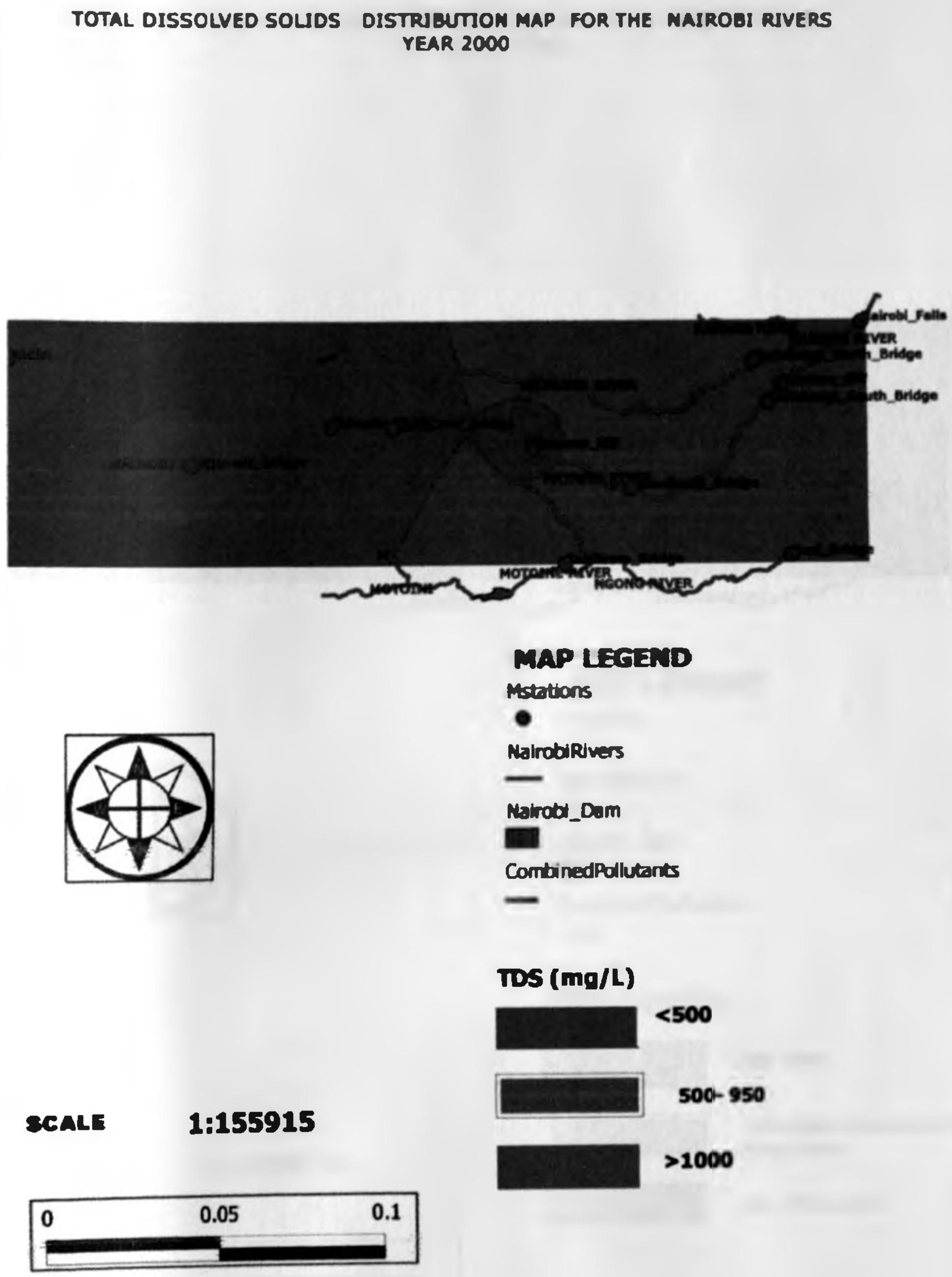


Fig. 4.4 Q TDS Spatial Distribution in The Nairobi River Basin.

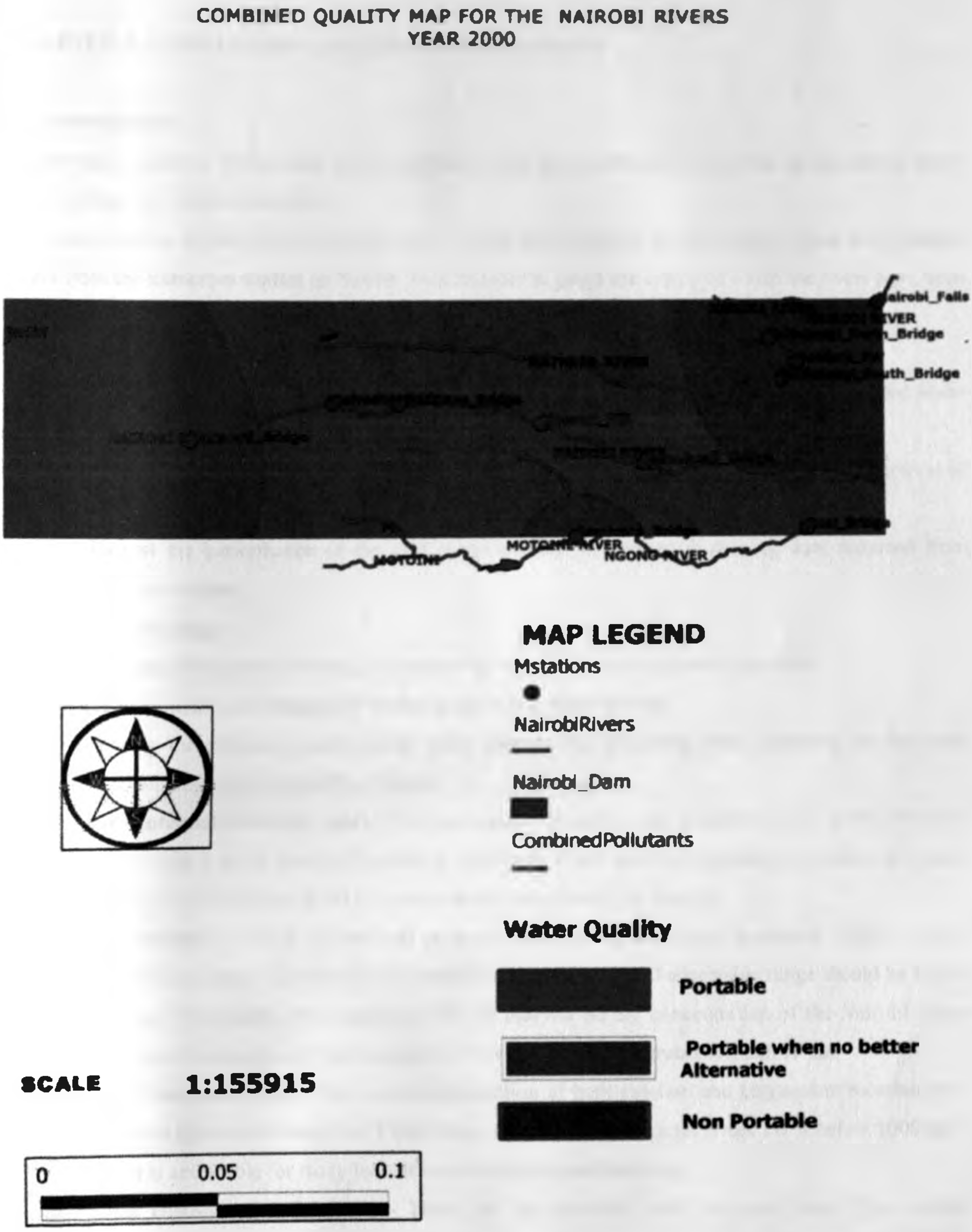


Fig. 4.5 Water Quality zone map for the Nairobi River basin.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.0: Introduction.

The primary objective of the study was to develop a GIS based software tool for use in monitoring water quality along the Nairobi river basin.

This was aimed at tackling the research problem, which was to collate the information about the pollution levels from the numerous studies on Nairobi river in order to gauge the extent to which the rivers have been degraded.

The following were the specific objectives;

1) To develop a GIS database to store digital map of the Nairobi River basin and associated water quality data.
2) To customize the GIS application for water quality monitoring, for easy analysis and retrieval of information.
3) To test the performance of the GIS water quality monitoring tool using data acquired from previous studies.

5.1 Empirical Findings.

We shall synthesis the empirical findings to answer the study's four (4) research questions:

5.1.1 What Parameters determine the Water quality in a River Basin?

After reviewing the previous studies from water experts the following were identified as the main parameters which determine the quality of water:

a) **PH (Potential Hydrogen ions) -** This is measure of acidity and alkalinity of the water. Its value ranges from 1 to 14. Levels of acidity or alkalinity affect aerobic respiration and hence the water quality. The acceptable WHO PH levels should be in the 6.5-8.5 range.

b) **Nitrates (NO_3)** - These are the final products of the biodegradation of Ammonia. High levels of ammonium nitrates indicate efficient biodegradation. The WHO acceptable range should be below 100mg/l. The results of our analysis showed that the nitrate concentration of the Nairobi rivers was above the acceptable level and hence the water requires reprocessing before use.

c) **Total Hardness (TH) -** This is the concentration of both calcium and magnesium bicarbonates. The WHO acceptable range for TH is below 1000mg/l. The Nairobi rivers TH is below 1000mg/l. which is acceptable for many industrial and domestic applications.

d) **Total Dissolved Solids (TDS) -** These are the dissolved solids in ionic form. These solids determine whether the water can be used for either domestic or industrial applications. The WHO recommends a TDS of 1000mg/l and below. TDS of the Nairobi Rivers is acceptable for the upper regions near the source whereas it rises to unacceptable levels at the lower regions.

5.1.2: What is the structure of the database for storing water quality data?

The four water quality parameters PH, NO_3^-, TH and TDS where used as the fields of the relational database for storing water quality parameters. Microsoft's MS Access database management system was used to create the database tables, their relationships and entry forms. The database was linked to the GIS software during the analysis stage. This was successfully accomplished as per the first objective.

5.1.3 What GIS is best suited to analyse and generate water quality maps?

There was the choice of using commercial GIS software, the most popular being Arc View or one of the open source GIS; represented by QGIS.

Commercial GIS software have expensive licenses, nevertheless they have a lot of support and extensive documentation of every aspect of usage. This is freely available once the license is acquired. On the other hand, free source GIS is freely downloadable but one has to rely on documentation from volunteers. Due to cost considerations and the fact that QGIS is very popular in the academic field, it was selected as the GIS for implementing this research project.

QGIS proved to be a good choice and although scattered, its documentation was freely available in the internet. Whereas a particular application was lacking or inadequate, a plug in would be available somewhere else in the net. By using a number of these plug ins (or software application extensions) we were able to customize the QGIS to generate a software tool for water for water quality monitoring and hence attaining our second objective.

5.1.4: What is the best format for displaying the geospatial pollution levels along a river basin?

The conventional methods of representing data are by use of graphs, bar charts, or pie charts among others. Since GIS is concerned with map making, the best method was to present our water quality data in a map form. First we had to turn the discrete data corresponding to the different monitoring points along the Nairobi river basin into contiguous data. This was done by use of inverse distance weighting (IDW) interpolation. This resulted into generation of water quality maps. The resulting maps gave similar results as predicted using other analysis methods. This way the GIS tool was tested and found to be quite accurate in predicting the water quality of the rivers, thus attaining the final objective of the projective.

5.2: Recommendations.

The following are key recommendations aimed at improving the GIS tool:

- A plug in subroutine should be incorporated to offer graphical menu to select the different tool functions, using a pointing device.
- PostgreSQL database could be used instead of MS Access. This is better suited for a distributed environment with remote multi access.

- Use of a Global positioning (GPS) device to identify the exact coordinates of the sampling points. This would give a more accurate coordinates than those acquired using Google maps.

5.3 Further Research Work.

In this research most of the data was historical, having been collected by other researchers. Also the data was from selected monitoring points along the Nairobi river basin. The coordinates of these monitoring points were established using Google maps.

We recommend that further research should be on the development of **a satellite and GIS tool** to assess water quality on the Nairobi Rivers. The research should focus on the use of Satellite remote sensing imagery to assess water quality. GIS software should be customized to perform the necessary processing.

The present research findings demonstrated the potential of using GIS software in water quality monitoring. More work can be done to extend its use in the management of other aspects of the environment like wildlife, air pollution, underground water management; just to mention a few..

REFERENCES

1) Alidi, A.S. (1998). '*A goal programming model for an integrated solid waste management system*' Arabian J. Sci. Eng., 23(1B), 3-16.

2) Atkinson, P.M. and Tate, N.J. (Eds.) (1999). Advances in remote sensing and GIS analysis. Chichester ; New York : Wiley.

3) Beck, M.B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resouces.* 23, 1393-1442

4) Brezonik, P.L., Kloiber. S.M., Olmanson, L.G. and Bauer, M.E. (2002). *Satellite and GIS Tools to Assess Lake Quality*, Water Resources Center, University of Minnesota.

5) Briggs, D.J. (1995). Environmental statistics for environmental policy: genealogy and data quality. J. Environ. Management. 44,39-54.

6) Chang, N.B. and Wei, Y.L. (1999). Strategic planning of recycling drop-off stations and collection network by multi-objective programming. *Environ. Manage.* 24(2), 247-263.

7) Chen, H.W. and Chang, N.B. (1998). Water pollution control in the river basin by fuzzy genetic algorithm-based multi-objective programming modeling. *Water Sci. Technol.*, 37(8), 55-62.

8) Chen, J.C., Chang, N.B., Chang, Y.C. and Lee, M.T. (2003). Mitigating the impacts of combined sewer overflow to an urban river system via web-based share-vision modeling. *J. Civ. Eng. Environ. Syst.*

9) Cunningham, W.P. and Saigo, B.W. (1999). *Environmental Science a Global Concern (Fifth Edition)*, McGraw-Hill.

10) Flynn, K.M. (1999). *Water Quality and Geographic Information Systems: The Alabama Watershed Demonstration project*, Au- burn University.

11) Goksel, C. (1998). Monitoring of a water basin area in Istanbul using remote sensing data. *Water Sci. Technol.*, 38, 209-216.

12) http://www.un.org/esa/sustdev/documents/docs.htm

13) http://www.un.org/esa/sustdev/documents/docs.htm.

14) Huang, G.H. (1996). IPWM: an interval parameter water quality

15) Huang, G.H., Guo, H.C. and Xu, Y.L. (1996). *Integrated Environmental Planning for Sustainable Development in Lake Erhai Basin with A Diagnostic Study for Local Environment Concerns*, United Nations Environment Programme, Nairobi, Kenya.

16) Issaias, I. (2000), Environmental Impact of Urbanization on Water Resources-A case study on Nairobi Dam. MSc. Thesis, Imperial College of Science, Technology and Medicine, Univ. London, London, UK.

17) Kazmi, A.A. and Hansen, I.S. (1997). Numerical models in water quality management: a case study for the Yamuna River (India). *Water Sci Technol.*, 36, 193-199.

18) Li, S. and Chen, G.H. (1994). Modeling the organic removal and oxygen consumption by biofilms in an open-channel flow. *Water Sci Technol.*, 30, 53-62.

19) Masliev, I. and Somlyody, L. (1994). Probabilistic methods for uncertainty analysis and parameter estimation for dissolved oxygen models. *Water Sci. Technol.*, 30, 99-108.

20) Miller, H.J. and Han, J. (2000). Discovering geographic knowledge in rich environment. *SIGKKD Exploration*, pp. 105-108.

21) Mwanthi, M W., L.O. Nyabola and E. Tenambergen (1997). Solid Waste Management in Nairobi City: Knowledge and Attitudes. Journal of Environmental Health, 60.

22) Natumanya E., F. Kansiime and F. L. Mwanuzi. 2009. Assessment of Nutrient Loading and Retentation along Nsooba stream and Lubigi Wetland. Kampala, Uganda.

23) Nyikuri, R. N. (1994). Reducing pollution along Nairobi River, Final year Civil Engineering project, Department of Civil Engineering, Univ. Nairobi, Kenya.

24) Ohayo-Mitoko, G. (1996). Concentrations of Heavy Metals, Organ chlorine Pesticides, Organic and Microbial Pollution in the Nairobi River and its Tributaries, in Dissemination Workshop on "Concentrations of Heavy Metals, Organochlorine pesticides, Organic and Microbial Pollution in the Nairobi River and its Tributaries", Sustainable Development Consultants.

25) Singh, R B. (1995). *Global Environmental Change Perspective of Remote Sensing and Geographical Information System*, A.A. Balkema, Rotterdam.

26) Soncini-Sessa, R., Luca, V. and Enrico, W. (1999). T woLe: a software tool for planning and management of water reservoir networks. *Hydrol. Sci. J.*, 44, 619-632

27) Tobler, W.R. (1994). Bidimentional regression, *Geographic Analysis*, pp. 187-212

28) Vijayan, G., Suresh, N.N., Subramanian, S.R. and Thamizhselvi S.N.T. (1999). Management of water resources for quality and quantity. *J. Indian Water Works Assoc.*, 31, 43-50.

APPENDICES

Appendix 1: SQL Programs

a) A SQL Program to create a VIEW for Displaying Selected Data from Selected Tables

```
CREATE VIEW Pollayer AS
SELECT Year,Code,Name,Lat,Long,PH,TH,TDS,NO3
FROM SampleStationDescrption,SampleStationData
WHERE
SampleStationDescrption.Code= SampleStationData.STID
```

b) A SQL statement to display Data in a VIEW

```
SELECT * FROM Pollayer
```

Appendix 2: Test Data

Selected parameters of Nairobi River water: Year 2001

Parameter/Site	January	April	July	October
Water temp. ° C/Sampling site				
Kikuyu	22.4	18.6	14.8	16.5
Kawangware	23.8	19.5	16.5	17.7
Chiromo	27.4	20.1	16.5	18.2
Eastleigh	32.1	22.9	18.9	21.5
Njiru	23.3	22.6	18.4	20.7
Fourteen Falls	22.8	24.3	19.2	22.4
PH/Sampling site				
Kikuyu	5.6	6.2	6.4	6.0
Kawangware	4.8	7.0	7.4	7.3
Chiromo	7.4	6.9	7.3	7.3
Eastleigh	7.2	6.8	7.1	7.1
Njiru	6.1	7.3	7.4	7.4
Fourteen Falls	5.3	7.0	7.3	7.3
Dissolved Oxygen $\mu g/l$ Sampling site				
Kikuyu	5700.0	3800.0	500.0	1300.0
Kawangware	4810.0	8000.0	9100.0	6100.0
Chiromo	7400.0	7900.0	7200.0	5900.0
Eastleigh	5000.0	900.0	300.0	900.0
Njiru	3700.0	4000.0	3200.0	2800.0
Fourteen Falls	1200.0	4300.0	4400.0	3200.0

Alkalinity μg/l/Sampling site				
Kikuyu	69200.0	120000.0	60000.0	80000.
Kawangware	72000.0	110000.0	85000.0	80000.
Chiromo	80000.0	100000.0	70000.0	75000.
Eastleigh	150000.0	180000.0	80000.0	90000.
Njiru	220000.0	173000.0	75000.0	75000.
Fourteen Falls	86000.0	68000.0	35000.0	65000.
Phosphates μg/l/Sampling site				
Kikuyu	22000.0	11000.0	2500.0	3500.
Kawangware	23000.0	12000.0	1500.0	3300.
Chiromo	24000.0	14000.0	1400.0	4000.
Eastleigh	27000.0	12000.0	9000.0	9600.
Njiru	24600.0	10000.0	3700.0	5300.
Fourteen Falls	31250.0	9000.0	2000.0	5600.

(*Continued*)

Parameter/Site	January	April	July	October
Nitrates μg/l/Sampling site				
Kikuyu	6200.0	6000.0	27200.0	1700.0
Kawangware	40200.0	22800.0	29500.0	520600.0
Chiromo	40600.0	39500.0	61000.0	63500.0
Eastleigh	30600.0	13300.0	31500.0	8300.0
Njiru	8100.0	12300.0	28000.0	6700.0
Fourteen Falls	30700.0	21800.0	32000.0	156700.0
BOD μg/l/Sampling site				
Kikuyu			33000.0	20720.
Kawangware			37200.0	29030.
Chiromo			114010.0	178360.
Eastleigh			183000.0	204020.
Njiru			171000.0	226420.
Fourteen Falls			30000.0	46400.
Hardness μg/l/Sampling site				
Kikuyu	195000.0	85000.0	98000.0	65000.
Kawangware	270000.0	135000.0	122000.0	80000.
Chiromo	305000.0	100000.0	195000.0	78000.
Eastleigh	320000.0	90000.0	202000.0	80000.
Njiru	410000.0	95000.0	170000.0	80000.
Fourteen Falls	210000.0	55000.0	85000.0	73000.

Appendix 3: Test Data

Table . Analysis of variance (ANOVA, Tukey HSD for unequal N) among the three groups of sites as defined by PCA along Nairobi River.

	Upper Stream (1)		Mid Stream (2)					Lower Stream (3)								Groups p Values		
Sites	1S	2R	3R	4R	5R	6T	7R	8R	9T	10R	11R	12T	13R	14P	15R	1 × 2	2 × 3	1 × 3
Altitude (m)	2028	2024	1824	1761	1693	1693	1690	1680	1627	1639	1557	1545	1524	1524	1503	0.003*	0.004*	0.000
pH	5.3	6.4	7.2	7.1	7.4	7.7	7.5	6.9	6.8	7.7	7.5	8.7	8	7.8	7.7	0.02*	0.77	0.01*
Dissolved oxygen ($mg\ l^{-1}$)	8.5	7.8	6	3.9	2.6	1.7	2.2	0.2	0.2	0.2	0.2	0.2	0.2	0.8	1.0	0.32	0.0004*	0.002*
Temperature (°C)	19.1	19.5	19.8	18.8	19	18.1	19.2	20.5	22.3	21.7	22.2	19.2	20.6	22.3	21.1	0.93	0.005*	0.01*
Hardness ($mg\ l^{-1}$)	77	51	73	99	98	133	99	79	125	51	95	143	98	130	115	0.31	1.0	0.3
Alkalinity ($mg\ l^{-1}$)	49	44	84	73	78	195	72	178	172	186	190	262	233	259	222	0.07	0.002*	0.0006*
Total dissolved ($mg\ l^{-1}$)	479	448	843	995	966	1023	940	1213	1344	1133	1310	1845	1391	1605	1495	0.0005	0.001*	0.000
Conductivity ($\mu S\ cm^{-1}$)	688	648	1213	1442	1440	1533	1417	1790	1968	1680	1935	3023	2038	2363	2203	0.001*	0.003*	0.0002*
Total suspended ($mg\ l^{-1}$)	19	30	98	320	67	16	66	513	799	1123	815	240	1212	29	302	0.59	0.08	0.06
Turbidity (NTU)	1.2	1.7	2.6	4.4	2.9	1.1	2.6	14.4	16	27.3	21	7.9	17.1	1.2	5.8	0.79	0.04*	0.08
Water discharge ($m^3 s^{-1}$)	0	0.04	0.15	0.33	0.25	0.02	0.26	0.44	0.55	0.53	1.84	0.57	2.24	0.32	2.72	0.04*	0.08	0.003
Chemical Oxygen demand ($mg\ l^{-1}$)	10	51	32	90	64	16	67	186	280	269	176	131	179	94	128	0.47	0.01*	0.01*
Nitrate ($\mu g\ l^{-1}$)	94	184	2253	3804	4077	3577	4413	838	201	184	94	161	158	2280	301	0.005*	0.001*	0.6
Nitrite ($\mu g\ l^{-1}$)	71	21	281	1312	794	1037	1017	61	561	41	431	41	141	1994	176	0.05*	0.15	0.4
Orthophosphate ($\mu g\ l^{-1}$)	9	7	140	122	167	167	178	1178	4083	4133	4117	2097	3238	5083	3065	0.0002*	0.0002*	0.0002*
Silicate ($mg\ l^{-1}$)	27	26	27	20	25	16	24	20	17	18	18	18	19	25	18	0.44	0.25	0.09

Marked (*) indicate significant difference at $p < 0.05$.

Appendix 4: Research Budget

Table 9.1: Research Budget

No.	Activity/Item	Estimated Cost (Kshs)
1.	Transportation ,Collecting Data (6Weeks)	15,000
2.	Photocopying, Binding and cost of digitized maps (or cost of Digitizing)	15,000
3.	Phone calls and Internet Costs	5,000
4.	Buying a laptop and appropriate software	60,000
5.	Data Processing (Research assistant)	15,000
6.	Miscellaneous and Contingencies	10,000
Total Cost		**120,000**

Appendix 5: Project schedule

Table 5.2: Schedule of activities

ACTIVITY	Duration	Sept 2011	Oct 2011	Nov 2011	Dec 2011	Jan 2012	Feb 2012	Mar 2012	April 2012
Proposal Writing		15							
Proposal Presentation: **Milestone one**			24	4					
Detailed Literature Review				12				4	
Data Collection					5	30			
Progress Presentation:- **Milestone Two**						30	10		
Data analysis							1	10	
Report writing		15							15
Final presentation: - **Milestone Three**								12 23	
Submission									16