MULTIVARIATE STATISTICAL ANALYSIS OF POLLUTION CHARACTERISTICS IN RIVERS OF THE MIDDLE GALMA BASIN, NIGERIA

By

ADIE DONATUS BEGIANPUYE
O.N.D (NWRI, 1987), B.Eng. (A.B.U, 1993), M.Sc (A.B.U, 2002)

Ph.D/Eng/36320/2002-2003

A DISSERTATION SUBMITTED TO THE
POSTGRADUATE SCHOOL,
AHMADU BELLO UNIVERSITY, ZARIA
NIGERIA

IN PARTIAL FULFILLMENT FOR THE AWARD
OF DOCTOR OF PHILOSOPHY IN WATER RESOURCES AND
ENVIRONMENTAL ENGINEERING.

DEPARTMENT OF WATER RESOURCES AND ENVIRONMENTAL
ENGINEERING

AHMADU BELLO UNIVERSITY, ZARIA
NIGERIA

January 2008

DECLARATION

I declare that the work in this dissertation entitled 'Multivariate statistical Analysis of Pollution characteristics of rivers of middle Galma Basin Nigeria' has been performed by me in the Department of Water Resources and Environmental Engineering under the supervision of Prof. C.A Okuofu and Drs. J.A. Otun and A. Ismail.

The information derived from the literature has been duly acknowledged in the text and a list of references provided. No pat of this dissertation was previously presented for another higher degree or diploma at another university,

Donatus B. Adie		11/02/2008
(Ph.D/Eng/36320/2002-2003)	**Signature**	**Date**

CERTIFICATION

This dissertation entitled "Multivariate Statistical Analysis of Pollution Characteristics in Rivers of the Middle Galma Basin, Nigeria" By ADIE Donatus Begianpuye meets the regulations governing the award of the degree of Doctor of Philosophy (Ph.D) of Ahmadu Bello University, Zaria, and is approved for its contribution to knowledge and literary presentation.

Prof. C.A. Okuofu — Date
Chairman, Supervisory Committee

Dr. J. A. Otun — 11/02/2008
Member, Supervisory Committee — Date

Dr. A. Ismail — 11/02/2008
Member, Supervisory Committee — Date

Dr. J. A. Otun — 11/02/2008
Head of Department — Date

POST GRADUATE SCHOOL — Office of the Dean — A. B. U. Zaria Nigeria

Dean, Postgraduate School — 21-5-2008
Date

DEDICATION

This dissertation is dedicated to the memories of the late Dr. Lawson Ivbaralu Odigie and the late Mr. Peter Behakong Adie., the two most influential figures in my academic and social life respectively. May God grant them eternal rest. Amen.

ACKNOWLEDGEMENT

I wish to acknowledge the inspiration given to me on this work by the Almighty God. He makes all things possible in his own time.

I wish to specially praise the efforts of my major supervisor, Prof. Charles Amen Okuofu who has practically nurtured me into a matured mind not only in my academic work but in his care for my family, but also through the several hours we have spent together writing countless job proposals and reports. I have come to understand that what makes a man is the resilience in which he pursues his aspirations. Thank you sir.

I wish to specially thank Dr. Abubakar Ismail, my immediate past Head of Department. He has been quite close and has assisted me on countless occasions both financially, morally and academically. He reposed so much trust in me that has exposed me to countless responsibilities and experience. Dr. Johnson A. Otun, the current Head of Department, has been a guiding light to me in the way he suggests and goads one into action. This thesis would not have seen the light of the day without his input and assistance even in the acquisition of the software. Mr. Maruf A. Ajibike has been my friend since 1985 and I cannot forget all the toils together as we marched along the academic field. He is a close confidant and an adviser. He first broached the idea of using multivariate analysis to me. Engrs. Ralph Iselema, Joshua Iliya, Bode Oyetade and Lawrence Tony Isah have been wonderful in their support.

I wish to acknowledge the patience exhibited by members of my immediate family especially my wife who has always assisted to keep the children away when daddy is busy on the computer in late hours. I wish to also apologize to my children who had

always expressed shock when daddy shouts angrily at their interruptions of his thoughts, which could be atimes very costly.

I wish to acknowledge specially the support from my academic colleagues spread across the globe for their assistance in no small way in the actualization of this thesis: namely Prof, Ityokombol (U.S.A), Dr. Leonard K. Shilgba (ABTI, Yola), Prof. K.J Osinubi, Dr. Idris Abubakar, Dr. A. T. Olowosulu (all of Civil Engineering, A.B.U), Prof. A. M. Oguntuase (Civil Engineering, F.U.T, Akure), Dr. Isaiah A. Oke (Civil Engineering, Ife), Dr. V. I. Ndirika and Prof. O.J. Mudiare (both of Agricultural Enginering, A.B.U). Others are my colleagues in the Department of Water Resources and Environmental Engineering, Faculty of Engineering; all well wishers in the Academic Staff Union of Nigerian Universities, Ahmadu Bello University Chapter and other people close to me.

Finally, I wish to acknowledge the great patience shown by my aged parents who had to endure a little more for some resources to be channeled to this work. I say thank you to all that I may have inadvertently omitted who had made contributions in prayers, pep talks and even criticisms that saw me through this work. I love you all. May God bless and reward every one of you. Amen

ABSTRACT

The physicochcmical characteristics of water in rivers of middle Galma basin were monitored at twelve stations during a period of two years. These stations were located at various points on river Shika, river Kubanni and river Galma. The parameters measured were pH, electrical conductivity, temperature, total dissolved solids, nitrates, phosphates, sulphates, dissolved oxygen, COD and BOD. The data collected was subjected to multivariate statistical analysis techniques (Principal component analysis, Factor analysis, Cluster analysis and Discriminant analysis). Results obtained showed that principal component and factor analysis were able to isolate factors that could be used to effectively monitor the water quality of the basin. Total solids, Chemical Oxygen Demand, Nitrate and Phosphates (in that order), were significant in the analysis, which proved that agricultural activities were a major source of pollution in the area during the wet season and domestic wastes during the dry season periods. Cluster analysis clearly delineated the seasons and sampling stations. The stations on river Kubanni and its tributaries and also the stations on river Shika were near perfectly grouped. Cluster analysis was also applied to the months, where they were clearly divided into rainy and dry months. Discriminant analysis was applied to these clusters to extract stations that were more useful in terms of representative data as well as months that were inappropriately classified. The output results from Discriminant analysis were able to predict all these in a very convincing manner. From the overall results obtained, sampling could be limited to the identified stations S8 and S9 which are Gaskiya Bridge and Tudun Wada Bridge for Kubanni River and station S10 which is located near the Zazzau ginnery on the River Shika. This could provide representative results for the rivers concerned. It therefore showed that multivariate statistical analysis can be used for temporal and spatial data interpretation in the middle Galma basin.

TABLE OF CONTENTS

Title page	i
Declaration	ii
Certification	iii
Dedication	iv
Acknowledgement	v
Abstract	vii
Table of contents	viii
List of Figures	xii
List of Plates	xiii
List of Tables	xiv
List of Appendix	xvii
Chapter One: Introduction	
1.1 General	1
1.2 Statement of the problems	7
1.3 Justification of the study	8
1.4 Research objectives and research benefits	8
1.5 Limitation of the Project	9
1.5 Description of project area	10
1.6.1 Geography	10
1.6.2 Geology	11
1.6.3 Vegetation	11
1.6.4 Drainage basin	12
1.6.5 Settlement pattern	14

3.2 Methods 40

3.2. (a). Parameters monitored 40

3.2. (b). Laboratory analysis 41

(i) pH 41

(ii) Electrical Conductivity 41

(iii) Temperature 42

(iv) Total Dissolved Solids 42

(v) Nitrate – nitrogen 43

(vi) Phosphates 43

(vii) Sulphates 44

(viii) Dissolved Oxygen 44

(ix) Chemical Oxygen Demand 45

(x) Biochemical Oxygen demand 46

(xi) Heavy Metals 46

3.2 (c) Statistical analysis of results 47

(i) Principal component analysis 47

(ii) Discriminant Analysis 49

(iii) Factor Analysis 50

(iv) Cluster analysis 50

3.2 (d). Operation of STATISTICA programme 53

Chapter Four: Results

4.0. General 54

4.1. Cluster Analysis 68

4.2. Factor and Principal component analysis 72

4.3. Discriminant analysis 78

Chapter Five: Discussion 87

5.1. Summary values 87

5.2. Cluster Analysis 87

5.3. Factor and principal component analysis 88

5.4. Discriminant analysis 89

Chapter Six: Conclusion and Recommendation

6.1. Conclusions 92

6.2. Recommendations 93

References: 95

Appendices 103

LIST OF FIGURES

Figure 1.1: Location map of the Galma Basin 17

Figure 3.1: Middle Galma Basin showing sampling points 36

Figure 3.2: Flow chart of FORTRAN programme for Principal Component Analysis 51

Figure 3.3: Flowchart of FORTRAN programme for Discriminant analysis 52

Figure 4.1(a): Plot of 2005 temperature data at each station 54

Figure 4.1(b): Plot of 2006 temperature data at each station 55

Figure 4.2 (a): 2005 pH values at each station 55

Figure 4.2 (b): 2006 pH values at each station 56

Figure 4.3(a): 2005 Electrical conductivity values according to months 56

Figure 4.3(b): 2006 Electrical conductivity values according to months 57

Figure 4.4(a): 2005 total dissolved solids values according to months 57

Figure 4.4(b): 2006 total dissolved solids values according to stations 58

Figure 4.5(a): 2005 Nitrate values according to stations 58

Figure 4.5(b): 2006 Nitrate values according to stations 59

Figure 4.6(a): 2005 Phosphate values according to months 59

Figure 4.6(b): 2006 Phosphate values according to months 60

Figure 4.7(a): 2005 Sulphates values according to stations 60

Figure 4.7(b): 2006 Sulphates values according to stations 61

Figure 4.8(a): 2005 Dissolved Oxygen values according to months 61

Figure 4.8(b): 2006 Dissolved Oxygen values according to months 62

Figure 4.9(a): 2005 Chemical Oxygen Demand values according to stations 62

Figure 4.9(b): 2006 Chemical Oxygen Demand values according to stations 63

Figure 4.10(a): 2005 Biochemical Oxygen Demand values according to stations 63

Figure 4.10(b): 2006 Biochemical Oxygen Demand values according to stations 64

Figure 4.11: Cluster plot for mean values for months 69

Figure 4.12:Cluster plot for mean values for stations 71

Figure 4.13: Scree plot of eigenvalues for stations (cluster – 2) 78

Figure 4.14: Plot of eigenvalues for months-(cluster – 1) 83

Figure 4.15: Plot of eigenvalues for months-(cluster – 2) 85

Figure 4.16: Box and whisker plot of average values of parameter 86

LIST OF PLATES

Plate I: Station 1at Area F as at January 2007 34

Plate II: Station 3 at Hayin Dogo as at January 2007 35

Plate III: Station 8 at Gaskiya Bridge as at January 2007 38

Plate IV: the confluence of Kubanni and Galma Rivers 39

LIST OF TABLES

Table 2.1: List of parameters measured and methods of determination 40

Table 4.1: Mean laboratory values for months 63

Table 4.2: Mean laboratory values for stations 64

Table 4.3: Summary of basic Statistics of the sample parameters 65

Table 4.4: Plotting distance for cluster analysis for sampling stations 69

Table 4.5: Plotting distances for cluster analysis for months 72

Table 4.6: Eigenvalues obtained from station mean analysis 73

Table 4.7: Principal components from station mean analysis 73

Table 4.8: Factors extracted from station mean analysis 74

Table 4.9: Eigenvalues obtained from monthly mean analysis 75

Table 4.10: Principal components for monthly mean values 76

Table 4.11: Factors extracted for monthly mean laboratory results 77

Table 4.12: Chi-squared test with successive roots removed for cluster 1 (stations) 79

Table 4.13: summary of canonical input performance for cluster 1 (stations) 79

Table 4.14: Canonical roots for cluster 1 (stations) 80

Table 4.15: Eigenvalues for cluster 1 (stations) 80

Table 4.16: Chi - squared test with successive roots removed for cluster 2 (stations) 80

Table 4.17 Summary of canonical input performance for cluster 2 (stations) 81

Table 4.18: Canonical roots for cluster 2 (Stations) 81

Table 4.19: Eigenvalues for cluster 2 (stations) 81

Table 4.20: Chi - squared test with successive roots removed for cluster 1 (months) 82

Table 4.21: Summary of canonical input performance for cluster 1 (months) 82

Table 4.22: Canonical roots for cluster (months) 83

Table 4.23: Eigenvalues for cluster 1 (months) 83

Table 4.24: Chi - squared test with successive roots removed for cluster 2 (months) 84

Table 4.25: Summary of canonical input performance for cluster 2 (months) 84

Table 4.26: Canonical roots for cluster 2 (months) 84

Table 4.27: Eigenvalues for cluster 2 (stations) 85

LIST OF APPENDIX

Appendix A: Averaged pH values

Appendix A.1: 2005 pH values

Appendix A.2. 2006 pH values

Appendix B: Averaged Electrical Conductivity values

Appendix B.1: 2005 Electrical conductivity values

Appendix B.2: 2006 Electrical conductivity values

Appendix C: Averaged Temperature values

Appendix C.1: 2005 Temperature values

Appendix C.2: 2006 Temperature values

Appendix D: Averaged Total Dissolved Solids values

Appendix D.1: 2005 Total Dissolved Solids values

Appendix D.2: 2006 Total Dissolved Solids values

Appendix E: Averaged Nitrates values

Appendix E.1: 2005 Nitrates values

Appendix E.2: 2006 Nitrates values

Appendix F: Averaged Phosphates values

Appendix F.1: 2005 Phosphates values

Appendix F.2: 2006 Phosphates values

Appendix G: Averaged Sulphates values

Appendix G.1: 2005 Sulphate values

Appendix G.2: 2006 sulphate values

Appendix H: Averaged Dissolved Oxygen values

Appendix H.1: 2005 Dissolved Oxygen values

Appendix H.2: 2006 Dissolved Oxygen values

Appendix I: Averaged Chemical Oxygen Demand values

Appendix I.1: 2005 Chemical Oxygen demand values

Appendix I.2: 2006 chemical Oxygen demand values

Appendix J: Averaged Biochemical Oxygen Demand values

Appendix J.1 2005 Biochemical Oxygen Demand values

Appendix J.2: 2006 Biochemical Oxygen Demand values

Appendix K: FORTRAN programme for Principal Component Analysis

Appendix L: FORTRAN programme for Discriminant Analysis

Chapter 1 INTRODUCTION

1.1. General

According to the global environment monitoring system (GEMS) survey conducted by United Nations Environment Programme (UNEP) in 1992, which included about 344 water-monitoring stations in 59 countries, it was reported that 10 percent of rivers monitored were polluted. The principal pollutant is domestic sewage. This problem was mainly confined to the developing countries, where counts of bacteria associated with sewage were, in places, millions of times higher than unpolluted waters. Rosegrant et al (2002) stated that the challenges of growing water scarcity are heightened by the increasing costs of developing new water sources, degradation of soil in irrigated areas, depletion of groundwater, water pollution and degradation of water related ecosystems and wasteful use of already developed water supplies, often encouraged by the subsidies and distorted incentives that influence water use.

Bos and Bergkamp (2001) posited that in addition to human consumption, water is integrally linked to the provision and quality of ecosystems. On one hand, water is vital to the survival of ecosystems and the plants and animals that live in them; on the other hand, ecosystems regulate the quantity and quality of water. Wetlands retain water during high rainfall and release it during dry periods, purifying it of many contaminants while forests reduce erosion and sedimentation of rivers and recharge groundwater.

Man's first interest in the quality of the environment was provoked by health and aesthetic considerations. Pollution of the environment, especially water, affects the lives of a great number of people throughout the world especially those living in developing countries. Each year,

hundreds of new chemicals and products are produced and, because many of these substances differ significantly from natural substances, detailed biochemical studies have to be designed to discover any potentially harmful effects. Often when synthetic substances are introduced into the environment, their movement through the ecosystem is difficult to track. Therefore, there is a need to have established procedures to monitor what is in the environment at all times. Many of the synthetic substances are not biodegradable and are of particular concern especially if they can accumulate in organisms including man. Such accumulation has effects which are long term in nature and can cause diseases such as cancers, cardiac disorders and birth defects (Murdoch, 1975).

Wood (1995) stated that the quality of any water body is governed by physicochemical factors and thus, the monitoring of the parameters is vital for both the long and short term analyses. Thus the quality of surface waters is a very sensitive issue. Anthropogenic influences (urban, industrial and agricultural activities), increasing exploitation of water resources, as well as natural processes (changes in precipitation inputs, erosion and weathering of crustal materials), pollute surface waters and impair their use for drinking, industrial and agricultural activities. These influences determine the quality of surface water of a region (Carpenter et al, 1998). According to Chapman (1992), because of the spatial and temporal variation in water quality, a monitoring programme that will provide a representative and reliable estimation of the quality of surface waters is necessary. Thus, monitoring programmes which include frequent water sampling at different sites and the measurement of a large number of biological and physicochemical parameters is usually conducted.

Many chemicals are essential or, at least, beneficial to human health but can be toxic when taken in excess through means such as food or water. An example is nitrate, a nutrient applied in largest quantities for lawn and garden care and crop production. In addition to fertilizer, nitrogen occurs naturally in the soil in organic forms from decaying plant and animal residues. The presence of nitrogen at high levels causes methemoglobinemia with an acute toxic response in sensitive population – children of six months or less. Most humans over one year of age have the ability to rapidly convert methemoglobin back to oxyhemoglobin; hence, the total amount of methemoglobin within the red blood cells remains low inspite of relatively high levels of nitrate/nitrite uptake. In the soil, bacteria convert various forms of nitrogen to nitrate (NO_3). This is desirable as the majority of nitrogen used by plants is absorbed in the nitrate form. However, nitrate is highly leachable and readily moves with water through the soil profile. If there is excessive rainfall or over-irrigation, nitrate will be leached below the plants' root zone and may eventually reach ground water.

Eutrophication has dramatically increased due to excessive input loads of plant nutrients phosphorus and nitrogen (Jolainkai, 1992). This is usually observed as the excessive growth of phytoplankton that turns standing waters and sluggish streams into green (algae bloom).

The dissolved oxygen (DO) content of water is one of the most characteristic quality parameters of surface water bodies. The DO is depleted by oxygen used for multiple chemical, biological and biochemical processes that take place within the water body (deoxygenation). There are also other natural, biological, chemical and physical reactions that that result in increase in DO (reaeration).

In aquatic environment receiving substantial organic waste load the most important consumer of oxygen is the biochemical decomposition of organic matter, that depletes the DO level of water by the amount of oxygen utilized by microorganism for their metabolic activity when decomposing, feeding on organic substances. The most widely used measure of the organic material content of matter is the Biochemical Oxygen demand (BOD), which is the amount of oxygen utilized by microorganisms in a unit volume of water for decomposing organic matter during a specified period of time.

Chemical Oxygen Demand (COD) is the amount of oxygen required to chemically oxidize waste matter. It is complementary to the BOD test. Its only limitation is that it does not differentiate unstable from stable organic matter. Because of its faster rate of determination than the BOD test, it can serve as a first line test for the detection of pollution in water.

Most aquatic organisms are poikilothermic (cold blooded) and so are unable to regulate their core body temperature. Therefore, temperature exerts a major influence on the biological activity and growth of aquatic organisms. Up to a point, the higher the water temperature the greater the biological activity. Fish, insects, zooplankton, phytoplankton and other aquatic species all have preferred temperature ranges. As temperature gets too far above or below this preferred range, the number of individuals of the species decreases until finally there are too few or none. Change in the growth rate of cold-blooded aquatic organisms and many biochemical reaction rate can often be approximated by the Q10 rule which predicts that the growth rate will double if temperature decreases by 10^{o} C (18^{o} F) within their preferred range (Michaud, 1991). Temperature is also important because of its influence on water chemistry. The rate of chemical reactions generally

increases at higher temperature, which in turn affects biological activity. Warm water hold less oxygen than cool water and some compounds are more toxic to aquatic life at higher temperatures.

The pH of a water sample is a measure of the concentration of hydrogen ions. The term pH was derived from the manner in which the hydrogen ion concentration is calculated. It is the negative logarithm of the hydrogen ion (H^+) concentration. Hence at higher pH there are fewer free hydrogen ions. A change of one (1) pH unit reflects a tenfold change in the concentrations of hydrogen ion. Factors that generally affect pH are photosynthesis which reduces carbon dioxide in water and hence acidity and respiration of organic matter which produces carbon dioxide and hence lowers pH. The pH scale ranges from 0 to 14. A pH of 7 is considered to be neutral. Substances with pH of less than 7 are acidic, substances with pH greater than 7 are basic. Values for pH are reported in standard units, usually to one or two decimal places depending on the accuracy of the equipment used.

Total dissolved solids (TDS) comprise inorganic salts and small amounts of organic matter that are dissolved in water. The principal constituents are usually cations of calcium, magnesium, sodium and potassium and the anions of carbonates, bicarbonates, chloride, sulphate and particularly in groundwater, nitrate fro agricultural use. The TDS originates from natural sources, sewage, urban and agricultural runoff and industrial wastewater. The concentration of TDS in water varies owing to different mineral solubilities in different geological zones.

Sulphate is a substance that occurs naturally in water. The sources of sulphate are from the decomposition of organic matter, in wastes from industries from rock blasts and the burning of fossil fuels. Sulphate fertilizers are a major source of sulphate in ambient waters.

Electrical conductivity indicates the degree of turbidity of a water sample. The major effect of turbidity on humans is simply aesthetic. Turbidity also adds real cost to the treatment of surface water supplies used for dinking water since the turbidity must be virtually eliminated for effective disinfection to occur. Particulate matter also provide attachment sites for heavy metals such as cadmium, mercury and lead and many toxic organic contaminants such as polychlorinated biphenyls (PCBs), polychlorinated aromatic hydrocarbons (PAHs), alkyl benzene sulphide (ABS) and many pesticides.

A number of heavy metals are toxic to humans because of their reaction with substances that have biochemical components such as enzymes and protein (Fergusson, 1990). The well reported Mina Mata case in Japan is an example of heavy metal poisoning. Children were born with physical abnormalities as a result of mercury poisoning. This occurred because the population consumed water and fish highly contaminated with mercury. The discharge of such contaminated water was over a long period of time and was not controlled until the effects were seen. Heavy metals such as cadmium used in metals plating can cause kidney and bone mass diseases at high concentration.

Botkin and Keller (1998) observed in their work, that agriculture has both primary and secondary environmental effects. The primary is also called on-site effect and affects an area where the

agricultural practice takes place. A secondary effect or off-site effect is that which affects an environment away from the agricultural site, typically downstream and upwind.

Hydrological processes are the decisive factors that affect the fate of pollutants in aquatic systems. They provide the main transport pathways as well as the media (water) for the most of the chemical and biological processes. Therefore, the fate of pollutants in water is greatly influenced by the all the elements of the hydrological cycle; namely precipitation, infiltration, surface and subsurface runoff, interception storage and evapotranspiration.

1.2 Statement of the problem

In Nigeria the problem of non-availability of potable water in large cities is unprecedented. This is mainly because of population explosion, migration from village to urban areas and the lack of the development of new water sources. The industries do not in most cases conform to the effluent discharge standards. Agricultural practices use fertilizers, pesticides and herbicides which end up in receiving water bodies. The effect of these substances on man, animal and plants cannot be overemphasized. Incidents of water borne diseases are rampant as water abstracted from polluted sources is not properly treated. Water is not tested at the point of abstraction and after treatment to establish its quality. In most cases because of the enormous amount of data that would be required, the authorities would rather not take action and so make subjective rather than objective interpretation of the water quality.

1.3 Justification of the study

In a river catchment and sub-catchment, all flows terminate at the water course. This is so because the various discharges must be subjected to the forces of the hydrological cycle. However, when the various other components of the cycle are satisfied, the rest flows to a receiving water body as surface flow. The sources of flow to the water body could be from known or unknown sources. They can also be intermittent or continuous, accidental or deliberate. The only way to know the impact of these external substances which in some cases may be harmful to man, is to know in what amounts they are present and whether or not they pose any health or environmental hazard. This can only be established if monitoring is carried out along the course of the river and at different times.

River basin authorities in developing countries hardly, if ever, engage in monitoring of river water quality; mainly because of the dilapidation of equipment and insufficient funding of the exercise.

Although there are several efforts to generate data in river basins, economic considerations and manpower shortages have compounded the problem in developing countries. There is therefore the need to develop a tool for easy identification of pollution sources and reduction in sampling time at a relatively cheap time..

1.4 Research objectives and research benefits:

This study generally aims to make significant contribution to knowledge on river pollution in Nigeria. The main tasks of this study shall be:

(1) To generate a reliable river monitoring data for the upper Galma basin and compare them with past data.

(2) To identify and test multivariate statistical analysis tools suitable for the interpretation of upper Galma basin pollution data

(3) To provide a suitable framework for objective monitoring and analysis of surface water quality in the upper Galma catchment

The research benefits include the following:

(i) The popularization of the use of the multivariate statistical analysis tool for objective water quality data interpretation

(ii) The improvement on the practices of data handling in water monitoring as would be revealed by model results in view of the usual data scarcity for objective analysis.

(iii) The probable identification of the most important data required for the classification of the pollution status of the catchment.

1.5 Limitation of project

Initial thoughts were to include heavy metals in water quality monitoring and analysis. However, preliminary laboratory results and comparisons with earlier analysis (Adakole, 2000 and Oke, 2000) showed that the heavy metal presence in the flow was within the Federal Environmental Protection Agency (FEPA) guideline values.

Comparative analysis between laboratory results on heavy metal analysis by Garba and Schoncick (2001) showed that the results cannot be relied upon and rather recommended the Center for Energy Research and Training, Ahmadu Bello University, Zaria as the best laboratory with most

reliable research facilities. However, during the course of the project, their units broke down and thus limited the scope of work. Hence the earlier intention of analyzing heavy metals was discarded.

1.6 Description of Project Area

1.6.1. Geography

The Kubanni sub-basin is bounded by latitudes 11° 02' and 11° 12' and longitudes 7° 35' and 7° 45' north of the equator. Most of the area is at elevation 670 meters above mean sea level. The Kubanni River is tributary of the Galma River. The basin forms part of the Galma basin and has an area of 118km^2. The Galma basin is bounded by latitudes 10° 37' and 11° 25' and longitudes 7° 31' and 8° 45' north of the equator. The Galma basin has an area of 6781km^2. The Galma basin is situated mainly in Kaduna State with a small portion of it extending to Kano and Katsina States. The Galma River joins the Kaduna River near Kaduna town. Previous surveys conducted (Ismail, 1996) show that the climatic condition of the area is of semi-arid type. The rainy season lasts from May to October with August being the wettest month. The basin is under-lain by superficial alluvial deposits and laterite (Eigbefo, 1978).

The Galma basin is part of the humid tropics marked by low amount of rainfall with a mean annual rainfall of 1170 mm. Rainfall in the basin decreases northward and begins earlier in the southern part of the basin in March and ends in October. In the Northern part, the rains start around April and end early to mid October. The duration of each storm event is usually short although with moderately high intensity. The mean annual total rainfall depth is about 11 mm with a maximum of 80 mm in August and a minimum of 2.5 mm in October. Though the amount of evaporation in

the area varies greatly from one locality to another, the average value is in the range of 6.5 mm to 10.5 mm.

The wind and relative humidity are affected by the south westerly and north easterly wind patterns. The south westerly wind blows during the wet season (April to September) and the North westerly wind blows in the dry season (October to March). Temperature and sunshine are interwoven. The mean temperatures in this area in the wet and dry seasons are 13°C and 35°C respectively. There is a remarkable drop in temperature in the harmattan season. The harmattan season starts in the month of December and lasts till February. It comes with dust associated with low level anti-cyclones which settles on the basin and other parts of northern Nigeria. The daily mean sunshine values are 4.9 hours for the least in August and about 8 hours between November and February. According to Ogunjobi (2000), monthly mean humidity is lowest in January/February and could be as low as 26%, while the highest is in August/September and could be up to 88%.

1.6.2. Geology

According to Ogunjobi (2000), the Galma basin is underlain by the crystalline rocks of the Basement complex, made up mainly of gneissic complex, meta-sediments and older granite. Except at places where the basement complex outcrops, it is overlain by varied thickness of weathered rocks, laterite soils, alluvium and clays. The southeast part of the basin is dotted in some places with younger granites of Jurassic age which consists mainly of granite with occasional intersperse of gabbros, rhyoalite and syenite

1.6.3. Vegetation

The vegetation of the Galma basin is of the northern Nigerian guinea savannah type, characterized mainly by herbs and grasses with few deciduous widely scattered trees. The herbs and grasses grow very tall in some places, although evergreen trees exist mainly along the perennial rivers and streams, mango trees dot the plains. Because of the seasonal character of rainfall, the colour of vegetation of the basin varies. It is green during the rainy season but becomes yellow or brown in the dry season

1.6.4. Drainage basin

The Galma basin drains Zaria and some part of Kaduna. The Zaria sub-catchment is made up of several small rivers, chief of which are the rivers Saye, Kubanni, Shika, Libarbu, Baki, Kinkiba, Anchau, Danwata and Galma (fig, 1.1). They have other tributaries and natural drains that contribute flows either seasonally or perennially to them. The three major rivers of interest, namely Kubanni, Shika and Galma are used for domestic, industrial and agricultural purposes. These rivers are the main sources of surface water in the Zaria area. River Kubanni has the Kubanni dam located in Ahmadu Bello University, river Shika has the Shika dam and is used for down stream irrigation farming, while river Galma has the Zaria dam on it. Rivers Kubanni and Shika drain into the Galma after their course through Zaria town.

According to Oguejiofor (1994), the Samaru sub-catchment (Kubanni drainage basin) originates in the precincts of Ahmadu Bello Main Campus, South West of its Teaching Hospital as a trench in an undulating agricultural land. It is fed by four or five tributaries. The Kubanni river has a bank width of about 5m and a bank depth of about 2m. The University dam is the main feature on the

Kubanni and is situated on the Main Campus. The river drains the southern zones of the area and receives effluent mainly from domestic activities. The A.B.U Teaching Hospital is situated just north of the river source and effluent from the hospital is expected to enter into it. The watercourse is a sluggish meandering type impounded at the Kubanni dam. Downstream flow is maintained by overflow from the dam, effluents from waste treatment plant, the abattoir and contribution from its tributaries. The Kubanni River is joined at the dam by a highly polluted Samaru stream which studies have shown is a major source of pollution to the dam because of the conveyance of domestic waste water from the Samaru village (Adie, 2001)

River Shika originates at the upland area of Shika village. It is perennial. In it course, it crosses the Sokoto road near NAPRI and is pounded at a downstream point where the Shika dam is located. It meanders between Samaru and Basawa villages and crosses the old Kano road near NARICT and flows into the River Galma. Flow in this river is torrential during the rainy season, but reduces drastically during the dry periods. Bank width and depths are about 5 and 3 meters respectively. However, there are pondages near culverts where man and animal use all year round.

River Galma has its origin from the Jos plateau highlands. It has a total length of about 1100Km. The distance from its source to Zaria is about 700Km. It emerges at Zaria through Hunkuyi, passes behind the north-eastern part of Zaria and behind commercial college after which it is joined by river Shika. From there it passes through agricultural areas as well as receives wastes from industrial and domestic sources. Further downstream, it is joined by River Kubanni. It then flows southwards towards the Zaria Dam and subsequently to join the Kaduna river at a distance of 70Km downstream.

The Bindare stream is a tributary of river Galma which flows in the east-west direction and runs into River Shika within the marshes that also include the confluence of the latter with River Galma. The stream apart from being the major stream that drains Chikaji industrial complex in Zaria, also traverses agricultural land as well as receive domestic wastewater. The Galma receives a lot of industrial effluent before it receives the Kubanni. The Zaria town water supply is abstracted at the point of confluence. Agricultural activities are the major anthropogenic activity in the area and it goes on all year round relying on the river and wetlands for irrigation.

1.6.5. Settlement Pattern

The Kubanni sub-catchment is located within the Galma river basin. The rivers to be studied are either the main rivers or their tributary. Villages and other settlements abound in this area as Zaria is an urban area. The major settlements in the study area are:

(1) Shika which has the National Animal Production and Research Institute (NAPRI), the Ahmadu Bello University Teaching Hospital (ABUTH), a boarding secondary school, poultry farm, fuel stations and very large expanse of arable land.

(2) Samaru which is the location of the main campus of the famous Ahmadu Bello University and its affiliate colleges and centers, several motor mechanic workshops, a medium sized market and several other higher institutions and schools. The campus is more sparsely populated than the densely populated Samaru village

(3) Basawa is a semi-urban settlement which houses a military barracks, the adjoining village, a chemical research institute, large expanse of cultivated land and a military range.

(4) Paladan is a densely populated settlement. Located within it are the famous Nigerian College of Aviation and the Nigerian Institute for Transport Technology.(NITT). There are also in this

area several fuel stations, a market and food and entertainment area. This area is adjoined by Zango, Hanwa, Government Reservation Area and Kwangila. This area contain most of the hotels and motels in the town as well as a motor park

(5) Chikaji and Muchiya area are the location of some of the industries in Zaria. We have a tobacco company, a seed company and several other industrial concerns. There is also the adjoining nucleus village settlement close to the industries.

(6) Sabon Gari and all its appendage places is the central area of the city. It is the commercial area of the city and the most densely populated.

(7) Tudun Wada, Gaskiya, Wusasa and all the adjoining areas are also highly populated. A lot of commercial activities are done here. Higher institutions and hospitals abound in this area.

(8) Zaria city, UPE and old Jos road are the oldest settlement area of the city. The city is highly populated and the seat of the Zaria emirate council as well as a Local Government headquarters. The Jos road area is the location for some additional industries.

The main anthropogenic activity in the area under study that affects the quality of surface water is agriculture. Farming is practiced all year round. In the rainy season crops like maize, millet, guinea corn, yam, potatoes and pepper are cultivated. While in the dry season, vegetables like carrot, cabbage, tomatoes, okra are cultivated. Fertilizers are intensively applied to provide nutrients for the crops. The fertilizers commonly in use include product names like urea, super-phosphate and Nitrogen Phosphate Potassium (NPK). The main components of these fertilizers are nitrogenous, phosphorus, calcium and potassium compounds. Manure is also used to provide nutrients for the crops. Fishing is common, but on a small scale in the parts of the rivers where there is enough water. Industrial activities in the area are of low level. Refuse disposal is a visible problem. Large

quantities of solid wastes (mainly organic) are produced and form enormous heaps near rivers that pass through urban areas. These unplanned dumps cause flooding due to blockage of the poorly built drainages.

Industries are located in two industrial areas – the Chikaji Industrial Area and the Jos road Industrial estate. While the Bindare stream, tributary of River Shika drains the Chikaji Industrial Area, the Jos road Industrial estate is drained by the river Galma after it has received rivers Shika and Kubanni. A notable feature of pollution causing activities is that they operate all year round; particularly the industries and farming operations, since irrigation resources are deployed in the dry season.



Figure 1.1 River Galma Catchment Area

Chapter 2 LITERATURE REVIEW

2.1 Water Quality Monitoring

Water quality monitoring of a river can be used to define the existing conditions, detect trends, and/or establish causes or sources of pollution (Mbarjiorgu, 2003). Singh et al (2004) stated that surface waters are more vulnerable to pollution due to their easy accessibility for disposal of wastewater. Anthropogenic activities introduce a range of inorganic substances, many of which are at concentrations that can be detected in the environment. According to Keller and Wilson (1992) the detection of these substances is enabled by improved sensitivity of analytical techniques and by the increased releases of substances into the environment. Municipal and industrial wastewater discharges constitute a constant polluting source, whereas surface runoff is a seasonal phenomenon that is influenced by the climate of the region under consideration. It is pertinent to prevent and/or control river pollution. This can only be accomplished if there is reliable information on the quality of the water. This data can then be used for effective watershed management and the prediction of pollution potential.

By far the most important effects that pollutants have on the environment are on aquatic life. Excessive nutrients from land used for crop production or waste disposal cause an imbalance in the natural ecosystem and intensify eutrophication in streams and lakes. Pesticides are accumulated in aquatic organisms and eventually man through the food chain. Such accumulation in man is suspected to be the cause of cancers ((Kakulu and Osibanjo, 1992). Synthetic organic materials such as industrial chemicals, detergents and other materials are toxic to man, plants and aquatic life. The effects can be observed in the number of fish kills that occur with the release of such

pollutants into the environment. The chemicals are resistant to ordinary waste treatment processes and may have long-term effects (Vesiling, 1974).

Oluwande (1988) discussed the causes and effects of pollution. He stated that water is polluted when its natural composition is altered. The alteration may produce physical, chemical or biological effect. Physical pollutants include thermal (heat), inert solids and inert liquid chemical and organic pollution components depending on their sources.

Inert liquid pollutants include oil. Oils do not react, dissolve or mix with water easily. The major cause of oily water pollution is from the petroleum industry (Osibanjo, 1988). The main sources are well blow - outs, tanker traffic and refinery operations. Other sources are industrial machinery oily wastes, discarded crankcase oil etc.

Iguisi and Madaki (2003) observed that inspite of the four world conferences on the deteriorating environment and its consequences namely Stockholm - Sweden (1972), Berlin-Germany (1985), Rio de Janeiro-Brazil (1992) and Beijing – China (2005), environmental degradation still persists, especially in developing countries. Effluents from industrial activities are discharged untreated into the rivers. In a study on the effect of polluted water on irrigated crops, Okafor et al (2002) observed that vegetables obtained from farms irrigated with polluted water were contaminated with bacteria of faecal origin. The report concluded that though the pollution was all season, that of the dry season contains more pathogens in number and kind than during the wet season. Oguejiofor (1994) also concluded that the intensity of pollution was high during warm, sunny and

low discharge periods and lower during cold high discharge periods. Adie (2001) found the impact of non - conservative substances in stream pollution is reduced by dilution during flood periods.

Pollution sources are classified as point and non – points. Point sources are discharges at discrete points. The origin of the wastes in the stream is either known or readily identified. Monitoring and control of point source discharges is accomplished in the line of the waste stream, be it pipe or open channel. Unlike point sources, non-point source pollution is generated from various land use areas. According to Cornwin and Wagenet (1996) and Salana et al (1999), 30% to 50% of the earths' surface is believed to be affected by non-point pollution. Wang et al (1999) stated that agricultural activities including the use of pesticides, is the single greatest contributor of non-point source pollutants. Different human activities generate pollutants that accumulate on land. Rainfall and runoffs convey them ultimately to receiving water bodies. Thus, according to Overton and Meadows (1976), the non-point source is simply the storm runoff. Calamari and Chiaudani (1984) stated that the number of possible contaminating substances in a water body is almost infinite; therefore, it is usually difficult to have a unified approach in tackling the factors that affect the fates of such contaminants in the water body.

Dix (1981) defined effluents as being turbid liquids which consist of 99.9% water containing complex compounds of organic and inorganic matters. They occur in the form of visible suspended solids, very finely divided suspended solids, colloidal particles, dissolved compounds and microorganisms such as protozoa, bacteria and viruses. Sixty six percent (66%) of organic matter present in effluents causes objectionable odour and colour due to the anaerobic bacteria reactions that take place within it. Wastewater is made up of both organic and inorganic matter. It is

composed of paper, faeces, urine, soap, detergents, fats, oils, greases and food materials. The inorganic substances present include sand, clay, ammonia gas, ammonium salts derived from the decomposition of urine, metallic salts, nitrates, phosphates etc. The precise composition of the effluent varies according to its origin and its industrial output or input which in some cases can contribute up to fifty percent (50%) by volume of total wastewater.

Kaduna Environmental Protection Agency (KEPA, 2000) stated that the Kaduna Refinery and Petrochemical Company (KRPC) discharges solid, liquid and gaseous waste which includes slop oil emulsions, heat exchange bundle cleaning sludge, cooling tower and\or tank bottom sludge, spent caustic, spent catalysts, miscellaneous process wastes, wastewater, maintenance and material handling waste from its corrugated plate separator. The average rate of effluent discharge from the company into Romi River was estimated at 3600 m^3 /day. This same river is used by the inhabitants of Romi and other adjoining villages for domestic and agricultural activities. Adakole (2000) reported that between 1972 and 1981, Nigeria recorded 957 cases of oil spill incidents resulting in a loss of 1,415,156 million barrels of oil. Chinonyere (2000) also reported that there were certain cases of overload of the Kaduna Refinery waste treatment plant from where polluted water ultimately ended in the Romi River.

James (1993) stated that water pollution monitoring models attempt to simulate the changes in the concentration of pollutants as they move through the environment. The main controlling factors that determine the fate of pollutants are the nature of the pollutants, the nature of the environment and the method of discharge. Egbo (2003) opined that sustainable water resources development is linked with good environmental stewardship. Some of the contaminants are carcinogens (cancer

causing), pentagenic (causing inheritable genetic diseases), tetratogenic (causing birth defects), neurotoxic (destroying nerve tissues) or just plain toxic, with short and long term exposures that may result in untimely death or deleterious health defects. Facts on death due to pollution are unavailable, but the experience from what is happening elsewhere can be used to conclude that the impact could not be different here in Nigeria.

In a state of the art study, Jolankai (1992) stated that the basic objectives of studying the hydrological, chemical and biological processes of contamination transformation and transport in aquatic environment are:

(a) to identify and quantify the state of the aquatic environment in both qualitative and quantitative terms and to reveal its spatial and temporal changes

(b) to quantify the cause and effects relationships. This is the response of various biotic and abiotic components of the aquatic system to the external, natural and anthropogenic inputs, with due regard to the interaction of these components within the water body.

(c) to describe the process mathematically in order to be able to simulate and predict the changes in the state of the aquatic environment resulting from natural processes and man made interventions.

In the same report, he stated that the fate of contaminants in aquatic systems is defined by the following factors:

(i) the type and kind of their external sources

(ii) the transport of the substances via various components of the hydrological cycle

(iii) the chemical , biological and biochemical transformations of these substances from one form to the other.

For a stream with point source pollution, the oxygen sag equation (Streeter – Phelps, 1925) can be used. Researchers have continued to study the effectiveness of this method; and has been observed by O'Connor (1967) that the oxygen sag equation has the following limitations:

(1) The equation is set to evaluate a single point discharging into a river reach with constant parameters (flow, depth and reaction rates), with other parameters assumed to be constant over the river stretch. This assumption can be extended by applying the equation to the system in a piecewise fashion by linking each piece with a mass balance method.

(2) The model assumes plug flow with constant velocity, but shear forces cause longitudinal mixing (dispersion) of the flow constituents

(3) The system is characterized as one dimensional hence river is completely mixed laterally and vertically, hence point sources are assumed to mix instantaneously

(4) Zero order kinetics was originally used to characterize plant activity and sediment Oxygen demand.

Neither of these phenomena is actually that simple, because application of these principles to practical works had always met with limited successes.

However, according to Okuofu (1990), the self-purifying capacities of water bodies are being overtaxed as a result of discharges into them. The result is gross pollution with all its associated problems. Many new organic and chemical substances are produced each year and some of them are persistent and non-degradable. Adie (2001) observed that there were some substances in water that could only transform from one form to another (conservative substances) that are usually difficult to model using the self-purification model.

Some studies have been conducted on pollution in river Kaduna and other rivers within Kaduna State { Beecroft (1985), Beecroft et al (1987), Oguejiofor (1994), Adakole (2000), Adie (2001), Ekanem and Irekpita (2004)}, but not much has been done on the toxicity of the industrial effluents from the various contributing industries in the study area. It is therefore difficult to say if the industrial effluents are the main pollutants or if other sources exist.

In hot climates, Mara (1986) stated that conventional sewage (wastewater) treatment system is recommended. However, Otun et al (2006) observed that less than 15% of industries in northern Nigeria operate a conventional waste treatment plant. Apart from the non-availability, most of the plants are in different states of decay.

Tebbut (1998) suggested that river water is best sampled at several points across a section and at several depths. He modified the thought by the fact that industrial wastewater discharges are intermittent, hence the necessity for an appropriate sampling programme to obtain a true picture of the discharges. It was emphasized that the more polluted a sample, the shorter the time allowed between sampling and analysis in order to avoid significant errors. He also stated that it was not necessary to analyze fish contamination in stream water quality analysis as the measure of such other constituents as Dissolved Oxygen (DO), Biological Oxygen Demand (BOD) and Ammonia would indicate if water was toxic or non-toxic to fish.

Davis et al (2003) were able to prove that brown algae represent an especially efficient and resilient class of bio-absorbents of heavy metals. Ackay et al (2003) hypothesized that trace elements undergo numerous changes during their transport due to dissolutions, precipitation and

absorption phenomena. The trace element concentrations in rivers depend not only on industrial and household waste inputs, but also on the geochemical composition of the area. Wang et al (2003) also stated that pH affects heavy metal uptake by sludge particulates. It was also reported by Rudd et al (1984) that dissolved organic matter (DOM) could form complexes with heavy metals. Hence the presence of DOM reduces the metal uptake.

Ekanem and Irekpita (2004) concluded that the result of their assessment of river systems in Zaria shows that they were already polluted to the extent of major concern. Channels are fast silting up and there are obvious signs of nutrient enrichment, which requires the urgent attention of the authorities concerned.

In view of the spatial and temporal variations in hydrochemistry of rivers, regular monitoring programs are required for reliable estimates of water quality. This results in a huge and complex data matrix comprised of a large number of physico-chemical parameters, which are often difficult to interprete and draw meaningful conclusions (Dixon and Chiswel (1996); Simeonov et al (2003). This large data matrix requires a reliable analysis and interpretation tool.

2.2 Multivariate Statistical Analysis

A good statistical tool can be used to study the temporal and spatial variation of the parameters of pollution and hence aid in interpreting the data. One of such methods is multivariate statistical analysis.

According to Abdi (2003), multivariate data analysis (MVDA) comprises of a set of techniques dedicated to the analysis of data sets with more than one variable. Several of these techniques were

developed only recently partly because they require the computational capabilities of modern computers. When multiple measurements are performed on several samples (observations), MVDA considers multiple variables in combination. The relationships are formalized in a set of structural equations which are the Multivariate models. The observations that are substantially different from the values estimated by the model are rejected. The techniques are not unified in their presentation and so, the choice of the proper techniques for a given problem is often difficult. The methods of multivariate analysis include Principal Component Analysis (PCA), Factor Analysis (FA) Correspondence Analysis (CA), Discriminant Analysis (DA), Cluster Analysis (CA), Multi Linear Regression (MLR), Multivariate Analysis of Variance (MANOVA), Multidimensional Scaling (MDS), Multiple Correspondence Analysis (MCA) and many other variants of them.

Cluster analysis (CA), Factor Analysis (FA), Principal Component Analysis (PCA) and Discriminant Analysis (DA) have widely been used as unbiased methods in analysis of water quality data for drawing meaningful information (Vega et al, 1998; Reghunat et al, 2002; Simeonov et al, 2003). Each researcher focused on not more than one method at a time. Helena et al (2000) stated that the multivariate treatment of data is widely used to characterize and evaluate surface and fresh water quality. It is useful in evidencing temporal and spatial variations caused by natural and anthropogenic factors linked to seasonality. Simeonov et al (2003) stated that the multivariate statistical analysis method is used in classification, modeling and interpretation of large data sets from environmental monitoring programmes; and allows the reduction of the dimensionality of the data and the extraction of information that will be helpful for the water quality assessment and management.

Bates et al (1998) studied the climatic and physical factors that influenced the homogeneity of floods in South Eastern Australia. They found that the question of what constitutes homogeneity in terms of flood and basin characteristics and how it can best be defined is problematic. To be able to establish the homogeneity of the flood data, it was subject to statistical analysis. In a further analysis of their work by the application of multivariate statistical analysis methods of tree-based modelling, canonical variate analysis, cluster analysis and principal component analysis, they were able to establish the relationship between climatic and physical factors by concluding that differences in seasonal pattern of rainfall has little influence on the statistical distribution of flood peaks.

Wengliang (2001) researched into the use of the multivariate analysis methods of linear regression and classification, respectively. He was able to apply the two methods in a secure two party computing (S2C) environment where data were observed by two parties who were not willing to fully share their private observations with others. This research departed significantly from the usual norm of data being sent to a central processing facility. He successfully integrated multivariate techniques into the data treatment and analysis by developing a security model which enabled them solve the problems of data sharing.

Anderson (2003) stated that the general operation of multivariate analysis involves the following:

- The means and variances of the separate measurement (for distributions and for samples) have corresponding relevance
- The different variables depend on each other

- The dependence may involve the covariance (that is the average products of their deviations from their respective means) between them
- When the covariance is standardized by corresponding standard deviations, we obtain the correlation coefficient, which serves as a measure of the degree of dependence

Abdi (2003), explained that the use of multivariate statistical analysis depends on the type of data sets. With one data set, the data tables to be analyzed are made of several measurements collected on a set of units, that is, subjects. In general, the units are placed in rows and the variables in columns. With two data sets, two cases are considered: in the first case, one set of data plays the role of predictors (independent) variables (IVs) and the second set of data corresponds to measurements or dependent variables (DVs). In the second case, the different sets of data correspond to different sets of DVs.

2.2.1 Principal Component Analysis (PCA):

This is a useful technique for the reduction of the dimensionality of a data set consisting of large numbers of interrelated variables, while retaining as much as possible the variability in the data set (Singh et al, 2004). According to Hopke (1985), the PCA is a useful pattern recognition technique that attempts to explain the variance of a large data set and inter-correlated variables with smaller set of independent variables (principal components). Mathematically, the method operates from the covariance matrix, which describes the dispersion of the multiple measured parameters, to obtain the eigenvalues and eigenvectors (An eigenvalues, also called a latent root or characteristic value is a scalar quantity applied to square matrix. The corresponding solution to the given equation is the

characteristic vector known as the eigenvector). According to Wunderlin et al (2001) and Sigh et al (2004), the PCA finds the variables (components), which account for as much of the variance in the multidimensional data as possible. Linear combinations of the original variables and eigenvectors result in new variables, called the principal components (PCs). Wunderlin et al (2001) explained that a further rotation of the axis defined by PCA produces new groups of variables called the varifactors (VFs). The routine for calculation finds the eigenvalues and eigenvectors of the variance –covariance matrix. Variance-covariance is chosen if all variables are measured in the same units. Correlation (normalized variance - covariance) is used if data is measured in different units. Normalization is by dividing all the variables by their standard deviations. The value obtained is the eigenvalues, which gives a measure of the variance accounted for by the corresponding eigenvectors. The percentages of variance accounted for by these components are also given. If most of the variance is accounted for by the first one or two components, then PCA is successful, but if spread more or less evenly among the components, the PCA is not successful. According to Wulder (2005), multi-collinearity (ill -conditioned correlation matrix), is not a problem in PCA as matrix inversion is not required. According to Anderson (2003), the principal components are linear combinations of random or statistical variables which have special properties in terms of variances. They are the characteristic vectors of the covariance matrix. The interpretation of the PCA is the use of the deviations of variables by discarding the linear combinations which have small variances and interprete those that have large variances. Stroud (1995) stated that the application of matrices to technological problems occurs in the form of equation of eigenvalues and eigenvectors which can be interpreted as influence factors in determinant applications.

2.2.2 Correspondence Analysis (CA):

This is a generalization of PCA to contingency tables. The factors of CA give an orthogonal decomposition of the chi – square associated to the table. The rows and columns of the table play a symmetric role and can be represented in the same plot. The routine finds the eigenvalues and eigenvectors of a matrix containing the chi – squared distances between all data points. The eigenvalues, giving a measure of the similarity accounted for by the corresponding eigenvectors, are given for the first four most important eigenvectors (or fewer if there are fewer than four variables). The percentages of similarity accounted for by these components are also given. According to Greenacre (1984), the usefulness of correspondence analysis is that the gain in data inter-probability far exceeds the loss in information.

2.2.3 Discriminant Analysis (DA):

This provides a statistical classification of samples and it is performed with prior knowledge of membership of objects to particular group or clusters such as temporal or spatial grouping of a sample. Wunderlin et al (2001) stated that the main use of DA is to obtain a group membership for a set of predictors. The method consists of finding a transform which gives the maximum ratio of difference between a pair of group multivariate means to the multivariate variance within the two groups. According to Wulder (2005) the underlying assumptions of DA are as follows:

(i) the observations are random samples

(ii) each group is normally distributed

(iii) the variance and covariance matrix for each group is the same

(iv) each of the observations in the initial classification is correctly classified

Further, DA helps in grouping the samples sharing common properties. Wunderlin (2001) noted that Discriminant analysis has been applied successfully to water quality analysis while Abdi (2003) opined that the Discriminant analysis is mathematically equivalent to the multivariate analysis of variance (MANOVA), and is used when a set of independent variables are used o predict the group to which a given unit belongs (normal dependent variables). The tool combines the independent variables in order to create the largest Factor (F), when the groups are used as a fixed factor in an analysis of variance (ANOVA).

2.2.4 Factor Analysis (FA):

Shuxia et al (2003) described factor analysis as an attempt to explain the correlations between observations in terms of the underlying factors which are not directly observable. Factor analyses reduce the attribute space from a large number of variables to a smaller number of factors. The variables with similar characteristics are grouped while some redundant information is taken off. The few factors produced, which might be a linear combination of the original variables is used to explain the observed variance of the original number of variables.

Singh et al (2004) observed that the principal components analysis is a part of factor analysis. According to Abdi (2003), factor analyses are of two types, known as the confirmatory factor analysis (CFA) and the multiple factor analysis (MFA). The CFA allows a generation of models of an underlying explanatory structure, which often expressed as a graph. Then the correlations between the DVs are fitted to this structure. Models are evaluated by comparing how well they fit the data. Variations over CFA are called structural equation modeling (SEM).

In multiple factor analysis, several data tables are combined into one single analysis. The first step is to perform a PCA for each table. Then each data table is normalized by dividing all the entries of the table by the first eigenvalue of its PCA. This transformation - akin to the univariate Z-score – equalizes the weight of each table in the final solution and therefore makes possible the simultaneous analysis of several heterogeneous data tables. According to Wulder (2005) for most forms of Factor analysis, singularity and multi-collinearity are a problem. The solution to this is the deletion of singular multi-collinear variables as required.

2.2.5 Cluster Analysis

This method is used to detect the natural groupings in data. Classification is based on the placing of objects into more or less homogenous groups in a manner that the relationship between groups is revealed. It actually lacks an underlying body of statistical theory and requires decision to be made by the user relating to the calculation of clusters. These decisions have an influence on the results of the classification. Clustering method could be top-down and thus employ logical decision; or bottom-up and thus, result in aggregation. The two key steps within cluster analysis are the measurement of distances between objects and grouping the objects based on the resultant distances. The distances provide for a measure of similarity between objects and may be measured in a variety of ways such as the Euclidean and Manhattan metric distance. The ideal number of groups to establish may be assessed graphically or numerically. Graphically the number of groups is assessed with a dendogram plot. Numerically the number of cases may be assessed on the agglomeration schedule by counting up from the bottom to where a significant break in numbers occurs.

Chapter 3 MATERIALS AND METHODS

3.1 Materials

3.1 (a) Period of Work

In order to monitor the pollution process in the river, the two distinct seasons of the year were taker into consideration. The rainy season in the project area lasts from May to October, while the dry season reaches its peak between January and March. Monthly samples were collected during both in the dry and rainy seasons. Sampling was done on the last Saturday of the month. The samples were collected for a period of two years (January 2005 to December 2006); and the data matrix was averaged to obtain a single set of data for the parameters.

3.1 (b) Sampling Locations

Twelve stations of varying distances ranging from 1 km to 3 km between them along the course o the Kubanni, Shika and Galma rivers were selected as sampling points. These stations correspond tc flow routes and inflow from discharge points (fig. 2). The principle for their selection is that they represented the best point for gaining access to the rivers and also suitable for easy sampling of the current water quality status and have a more progressive pollution load. Samples were collectec from these points as stated in section 3.1 above. The samples were analyzed in the Environmenta Engineering Laboratories of the Department of Water Resources and Environmental Engineering Ahmadu Bello University, Zaria and National Research Institute for Chemical Technology, Zaria.

The description of the locations of the selected sampling stations along the river courses are enumerated below:

Station 1 – Area F

Situated behind Area 'F' Senior Staff quarters of A.B.U .The station is characterized by a rocky substratum over which the river Kubanni flows. The site is located on a major tributary of the river Kubanni. The area is highly cultivated and the station is located on a major cattle path. The situation in this station as at January 2007 is shown in plate I.



Plate I: Station 1in area F as at January 2007

Station 2 – Area BZ

This station is situated behind Area BZ in the senior staff housing area of Ahmadu Bello University main campus. Samples were collected at a bridge that serves as access route to villages behind the

Campus. This station is on the main course of River Kubanni just before the impoundment at the Kubanni dam. The area is highly cultivated and the soil is lateritic in nature even though it is covered by deposited alluvial soil.

Station 3 – Hayin Dogo

This station is situated north of Samaru village. Samples were collected on the bridge linking Samaru village to Basawa Village and near the Basawa barracks. The area is highly cultivated with the site serving as a water hole for cattle. The station is one of the tributaries that drain into River Shika. The situation on this station as at January 2007 is shown in plate II.



Plate II: Station 3 at Hayin Dogo as at January 2007

Station 4 – Basawa.

This station is situated on the bridge linking Palladan to Basawa village. The area is highly cultivated with some small scale industrial activities close by such as a bakery, poultry and a ranch. The army cantonment and several academic institutions are located in this vicinity.

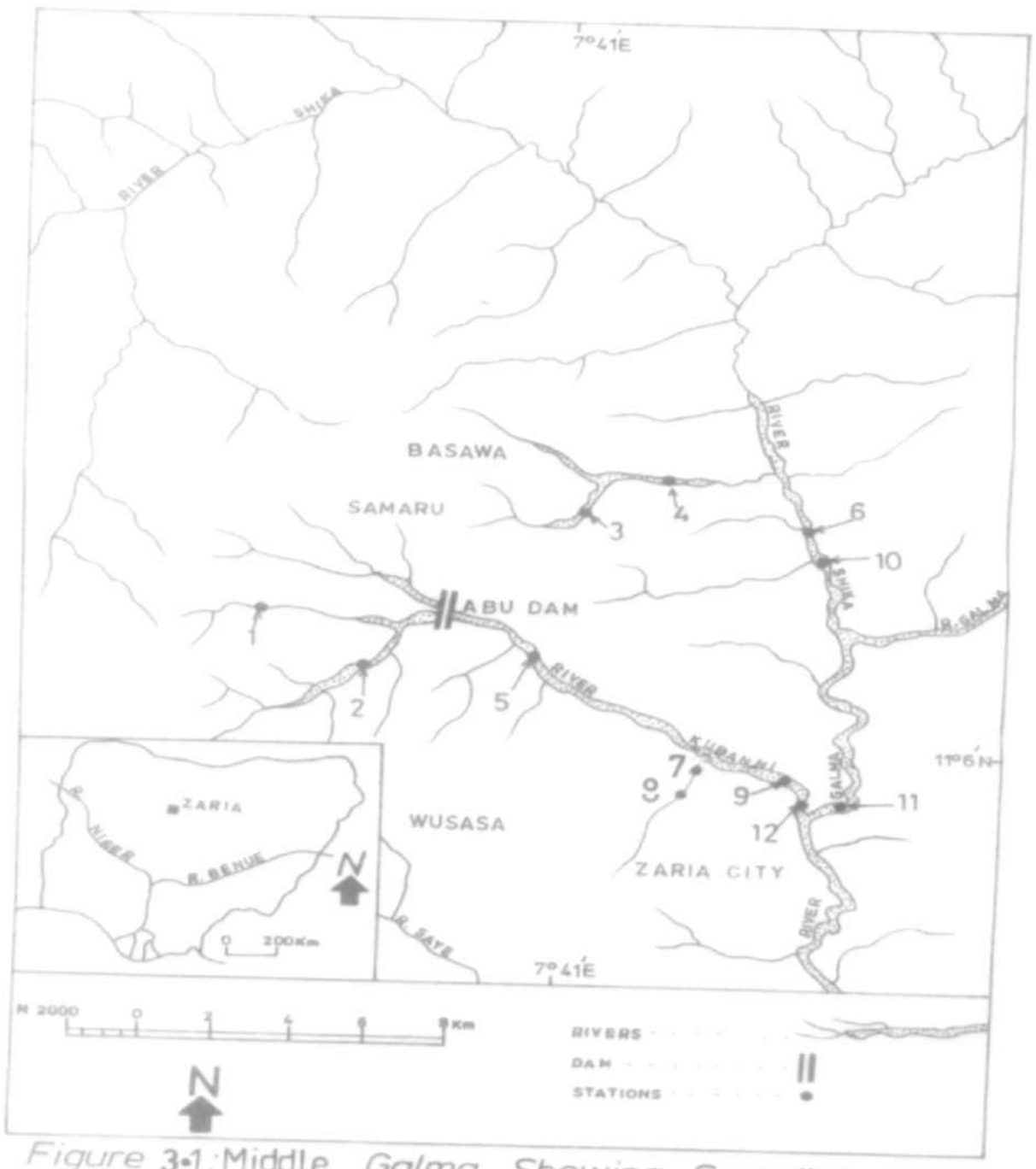


Figure 3-1: Middle Galma Showing Sampling Points

Station 5 - Zango

The station is situated downstream of the Zaria abattoir. The area is highly cultivated and the station is on River Kubanni downstream of the dam. Upstream, it receives effluent from the Ahmadu Bello University waste stabilization pond and domestic drainages from houses in Jama'a village and offices within the area.

Station 6 – Kings Road

The station is situated on a bridge on Kings Road, Sabon Gari, Zaria leading to Commercial College and is on Bindare Stream, a tributary of River Shika and receives effluent from the contiguous domestic area and the Chikaji Industrial Area. The area is highly populated with some agricultural activities. The station location serves as a refuse dump also and is located in a major irrigation area.

Station 7 – Fire Service

The station is situated on the bridge on Kaduna road .The station is on a tributary of River Kubanni. This stream originates from the Kufena hills in Wusasa and drains the Zaria city settlement which is a very densely populated enclave within the ancient walls. The area near the station is intensely cultivated.

Station 8 – Gaskiya Bridge

The station is situated on the road linking Agoro to Gaskiya in Tudun Wada, Zaria. The station is down stream of station 7 and located on a tributary of River Kubanni. The area around the station is densely populated. Close to the station are refuse dumps. The stream in this stretch receives effluent from domestic and commercial activities (mechanic workshops and a slaughterhouse).

Station 11 – Galma

This station is situated close to the bridge on old Jos road. The area is highly cultivated all year round and contains the Zaria east industrial area. Samples were collected just before River Kubanni joins the River Galma. This station receives effluent from all the industries in the industrial estate. The confluence of the two rivers is shown in plate IV.

Station 12 – Kubanni

The station is situated at old Jos road just before draining into River Galma. The area is highly cultivated and populated with a refuse dump close to the bank of the river. Intake point for Zaria water works is located at the immediate downstream section of the confluence of rivers Kubanni and Galma.



Plate IV: the confluence of Kubanni and Galma Rivers

3.2 Methods

3.2 (a) Parameters Monitored

An electronic sampler was used to collect the sample from below the surface of the river. This was discharged into a one and a half (1.5) liter polyethylene container. This operation was repeated from the first to the last station mindful of the time of flow from one station to the other in some cases.

The list of the parameters monitored was influenced by the nature of the discharges from the major pollution source. Listed in Table 3.1 are the parameters measured and the equipment for field and laboratory analysis. The laboratory analysis was carried out as recommended in the Standard Methods (APHA, 1985)

Table 3.1: List of parameters measured and methods of determination

S/No	Water quality parameter	Method of determination
1	Temperature	Thermometer
2	Dissolved Oxygen	Modified Winkler method
3	pH	Corning EEL digital pH meter
4	Electrical conductivity	Beckam laboratory conductivity meter
5	Biochemical Oxygen Demand	BOD bottle method
6	Chemical Oxygen Demand	Acidified dichromate method
7	Nitrate Nitrogen	Phenoldisulphonic acid method
8	Phosphate	Vanadomolybdophosphoric Acid method
9	Sulphate	Turbidmetric method.
10	Total Dissolved Solids (TDS)	Gravimetry using evaporation

3.2 (b) Laboratory Analysis:

Samples were collected monthly for 24 months. Temperature, pH and Dissolved Oxygen were measured immediately the samples got to the laboratory. Electrical Conductivity (EC), Total Dissolved Solids (TDS), Nitrates, Phosphates, Sulphates, Chemical Oxygen Demand (COD) and Biochemical Oxygen Demand (BOD) were analyzed thus:

(i) pH: The value was determined electronically using a pH meter with a glass electrode and a buffer solution of pH 4 and 9 respectively.

(1) the pH meter was washed for 15 – 20 minutes. The electrode was removed from the distilled water and wiped dry.

(2) the pH meter was standardized using the first buffer solution at the specified temperature (25°C)

(3) the pH meter was switched off and the electrode from the second buffer was removed and rinsed thoroughly and wiped dry with a clean tissue paper.

(4) the electrode was reinserted into the sample and power restored on the meter. The ph reading was then taken on the calibrated scale.

(5) the electrode was rinsed and was stored in distilled water.

The average values for pH are shown in Appendix A. Also the 2005 values are shown in Appendix A.1, while the 2006 values are shown in Appendix A.2

(ii) Electrical Conductivity (EC): According to Tebbut (1998) and Standard Methods (APHA, 1985), the EC can be measured using the Bekam laboratory meter.

(1) the cell was rinsed with distilled water

(2) the sample was then poured into the conductivity cell with a constant value (k=1)

(3) the cell was then rinsed with the sample while ensuring that no air bubbles remains in the cell

(4) the temperature of the sample in each cell was taken with a thermometer. The measured temperature was used to adjust the electrolytic conductivity meter.

(5) an electrode was inserted into the cell and connected to he meter. The electrical conductivity of the sample was then read off the meter

The average values for Electrical conductivity are shown in Appendix B. Also the 2005 values are shown in Appendix B.1, while the 2006 values are shown in Appendix B.2

(iii) Temperature: The sample temperature was measure in the field with a thermometer. The average values for temperature are shown in Appendix C. Also the 2005 values are shown in Appendix C.1, while the 2006 values are shown in Appendix C.2

(iv) Total Dissolved Solids: Gravimetric method by the use of an oven evaporating dish was used.

(1) the evaporating dish was ignited at 105 °C and the cooled to room temperature in a dessicator.

(2) the evaporating dish was then weighed in a balance and recoded as W_1.

(3) 100 mL of the sample was measured and filtered using double filter paper. The filtrate was poured into and evaporating dish. The dish was then placed in a water bath to dry at a temperature of 105 °C.

(4) the dish was then cooled in a dessicator to room temperature and weighed as W_2

The final weight of total dissolved solids (W) was calculated as in equation 1

$$W(mg/L) = (w_1 - w_2)x100 \qquad 3.1$$

The average values for total dissolved solids are shown in Appendix D. Also the 2005 values are shown in Appendix D.1, while the 2006 values are shown in Appendix D.2

(v) Nitrate-Nitrogen: 50ml of sample was evaporated to dryness and cooled. The residue was dissolved with 2ml phenoldisulphonic acid reagent using a glass rod. The residue was diluted and transferred to an Erlenmeyer flask. 10ml of 12mM KOH was added and the mixture made up to 100ml with distilled water. Absorbance was measured at 410nm and compared with the nitrate standard. Distilled water was used as blank. Nitrate - N was calculated using:

$$Nitrate - N(mg/L) = \frac{Nx1000}{SampleVolume(ml)} \qquad 3.2$$

N = Concentration from standard curve.

The average values for Nitrates are shown in Appendix E. Also the 2005 values are shown in Appendix E.1, while the 2006 values are shown in Appendix E.2

(vi) Phosphate: 35ml of the sample was placed in 50ml volumetric flask and 10ml Vanadate – Molybdate reagent added and made up to mark with distilled water. Absorbance was measured at 400nm using distilled water as blank and KH_2PO_4 as standard.

Phosphate was calculated as:

$$Phosphate - P(mg/L) = \frac{Px1000}{SampleVolume(ml)} \qquad 3.3$$

Where, P = concentration from standard curve.

The average values for phosphates are shown in Appendix F. Also the 2005 values are shown in Appendix F.1, while the 2006 values are shown in Appendix F.2

(vii) Sulphates: 100ml of sample was placed in a conical flask and 20ml buffer solution added and stirred. A spoonful of $BaCl_2$ crystals was added and mixed. Absorbance is measured after 60 seconds at 420nm using distilled water as blank and Na_2SO_4 as standard.

Sulphate was calculated as:

$$Sulphate(mg/L) = \frac{Sx1000}{SampleVolume(ml)} \qquad 3.4$$

S = concentration from standard curve.

The average values for sulphates are shown in Appendix G. Also the 2005 values are shown in Appendix G.1, while the 2006 values are shown in Appendix G.2

(viii) Dissolved Oxygen (DO): The DO was measured in the laboratory. A 250 ml of sample was dispensed in the BOD bottle. 2ml of manganese Sulphate solution was added. 2 ml of alkali iodide-azide reagent was also added. The bottle was stoppered to exclude air bubbles and then mixed by inverting the bottle several times. When the precipitate had settled, a clear supernatant was left above. The bottle was then shaken again. When the final settling produced 100 ml of clear supernatant, 2 ml of Sulphuric acid (H_2SO_4) was added. The bottle was re-stoppered and mixed by gentle inversion. A 200ml of the mixture was removed for titration. The extracted 200 ml was titrated with 0.025N Sodium Sulphate ($Na_2S_2O_3 .5H_2O$) to pale straw colour. 1 ml of starch solution was then added. Titration was continued to the first disappearance of the blue color.

The DO content was calculated thus;

$$DO(mg/L) = \frac{TitreValue(ml)x0.2x100}{200} \qquad 3.5$$

The average values for Dissolved oxygen are shown in Appendix H. Also the 2005 values are shown in Appendix H.1, while the 2006 values are shown in Appendix H.2

(ix) Chemical Oxygen Demand (COD): 400 mg of Mercury Sulphate ($HgSO_4$) was pipetted into a reflux flask. 20 ml of sample was added and mixed. 10ml of standard potassium dichromate ($K_2Cr_2O_7$) solution was added. Glass beads heated for one (1) hour was added. The flask was then attached to the reflux condenser. 30 ml of concentrated H2SO4 containing Silver Sulphate (Ag_2SO_4) was added through the open end of the condenser and mixed thoroughly while adding the acid. The mixture was cooled and the condenser washed with 25 ml of distilled water. The mixture was diluted to 150 ml with distilled water and then cooled to a warm temperature. Three (3) drops of ferine indicator was added. The mixture was then titrated with ferrous ammonium Sulphate ($Fe(NH_4)_2(SO_4)_3$) by a sharp change from blue to green and then to reddish brown. A blank was refluxed in the same manner. The blank consisted of 20 ml distilled water, with same reagents.

The COD was calculated as:

$$COD(mg/L) = \frac{(a-b).Nx8000}{ml.of\ sample} \qquad 3.6$$

where,

a = ml of ($Fe(NH_4)_2(SO_4)_3$) used for blank

b = ml of ($Fe(NH_4)_2(SO_4)_3$) used for sample

N = normality of ml of ($Fe(NH_4)_2(SO_4)_3$) = 0.25

The average values for chemical Oxygen demand are shown in Appendix I. Also the 2005 values are shown in Appendix I.1, while the 2006 values are shown in Appendix I.2

(x) Biochemical Oxygen Demand: In this test, the BOD_5 was determined, so as to be able to remove the effect of the oxygen demand exerted by nitrification. The sample used for DO measurement was incubated for 5 days. The sample was treated as in the dissolved oxygen analysis

above after the Fifth (5^{th}) day. The difference between the DO measurement and the results after 5 days gave the BOD.

$$BOD_5 = DO.(Initial) - DO(After.5.days) \qquad 3.7$$

The average values for biochemical oxygen demand are shown in Appendix J. Also the 2005 values are shown in Appendix J.1, while the 2006 values are shown in Appendix J.2

(xi) Heavy Metals Analysis: Methods for determining the concentrations of heavy metals vary in complexity according to the interfering substances that may be present (APHA, 1998). Metals are determined typically by (i) flame atomic absorption (ii) electothermal atomic absorption (iii) inductively coupled plasma (ICP) or mass spectrometry. The various classes of metals are:

(a) Dissolved metals which are present in unacidified samples that pass through a 0.45µm membrane filter

(b) Suspended metals which e present in unacidified samples that retain on a 0.45µm membrane filter

(c) Total metals which are a total of dissolved and suspended metals of concentration determined on an unfiltered sample after digestion,

(d) Acid extractable metals which are metals in solution after an unfiltered sample is treated with hot dilute acid

3.2 (c) Statistical Analyses of Results

In this research, the four selected methods of statistical analysis are principal component analysis, factor analysis, cluster analysis and Discriminant analysis. These methods fit the spatial patterns of

rainfall and stream flow measurements in the tropics. They are also amenable for use in the data scarce conditions that are prevalent in third world countries.

(i) Principal Component Analysis (PCA):

Data obtained from water quality laboratory analysis be treated by computing factors such as means and standard deviations for each of the monitored parameters. The steps to solving PCA problems are expressed thus:

$$A = \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})' \qquad 3.8$$

And an estimate of $\sum$ (the population covariance matrix) is expressed as

$$S = \frac{A}{n} \qquad 3.9$$

Where

S = the sample covariance matrix

n = N – 1

The sample data is converted to Z scores in other to remove the influence of measurements carried out in different units. With equation 3.10 it is possible to perform a linear combination of the variables.

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{S_j} \qquad 3.10$$

where

i = 1,2,.....p is the rows

j = 1,2,.....n is the columns

We use the iterative procedure to find the first principal component by computing:

$$Z^{(j)} = SZ^{(j-1)} \qquad 3.11$$

Each successive vector is compared by computing a variance coefficient given by equation (3.12):

$$r_i^{(j)} = \frac{Z^{(j)}}{Z_i^{(j-1)}} \qquad 3.12$$

Each of the values of $r_1^{(j)}$ calculated is an approximation of the value of l_1, which is the largest root of S. The first set of values calculated is S_1 and the matrix of values of S_1 is $b^{(1)}$. This is multiplied by its transpose ($b^{(1)\prime}$).

Then S_2 is calculated by further iteration as follows:

$$S_2 = S_1 - l_1 b^{(1)} b^{(1)\prime} \qquad 3.13$$

The subsequent principal component is found from

$$S_3 = S_2 - l_1 b^{(2)} b^{(2)\prime} \qquad 3.14$$

Application of matrices to technological problems occur in the form of equation 3.15

$$A.X = \lambda.X \qquad 3.15$$

For a non trivial solution, $X \neq 0$ the values of λ are called the eigenvalues of matrix A and the solution to equation (16) is called the eigenvectors.

Therefore,

$$|A.X - \lambda I| = 0 \qquad 3.16$$

The unit matrix I is introduced since we can only subtract a matrix from another matrix. $|A.X - \lambda I|$ is called the classic determinant, while $|A.X - \lambda I| = 0$ is called the characteristic equation. Expanding the determinant of equation (3.16) gives a polynomial of degree A, while the solution of the characteristic equation gives the values of the eigenvalues (λ) of matrix A.

These equations have been written into a programme by STATISTICA, a programme that incorporates basic statistics and multivariate analysis. The FORTRAN programme listing of principal component analysis is shown in appendix K. The flowchart for principal component analysis is shown in figure 3.2.

(ii) Discriminant Analysis (DA):

$$f(Gi) = ki\sum_{j=1}^{n} Wij + Pij$$

where;

i is the number of groups (G)

ki is the constant inherent to each group

n is the set of parameters used to classify a set of data into a given group

Wij is the weight coefficient assigned by DA to a given selected parameter

Pij the analytical value of the selected parameter.

In this research work two groups for temporal (2 seasons) and twelve groups for spatial (twelve sampling stations) were selected. The station (spatial) and the season (temporal) were the grouping (dependent) variables, while all the measured parameters constituted the independent variables. A FORTRAN programme listing of the Discriminant analysis is shown in appendix L. The flowchart is shown in figure 3.3.

(iii) Factor Analysis (FA):

Factor analysis has no mathematical basis. It works in conjunction with the principal component analysis. The factor analysis reduces the contribution made by less significant variables in order to

simplify even more of the data coming from principal component analysis. This is achieved by rotating the axis defined by the principal component analysis according to such well defined rules as transformation;. in which case data is rewritten with properties lacking in the original data. Unlike the principal component analysis which assumes that all the variance is common, facto analysis assumes there is some unique variance in the data. This is interpreted by the values of eigenvalues obtained.

(iv) Cluster Analysis

The Euclidean distance (between) rows is a robust and widely applicable measure. Distance is converted to similarity by changing the sign. The Euclidean distance is expressed as equation (2.18)

$$\text{Euclidean Distance}_{jk} = \sqrt{\sum_{i=1}^{n}\left(X_{ij} - X_{ik}\right)^2} \qquad 3.18$$

Where;

X_{ij} = Variate

X_{ik} = Mean of variates

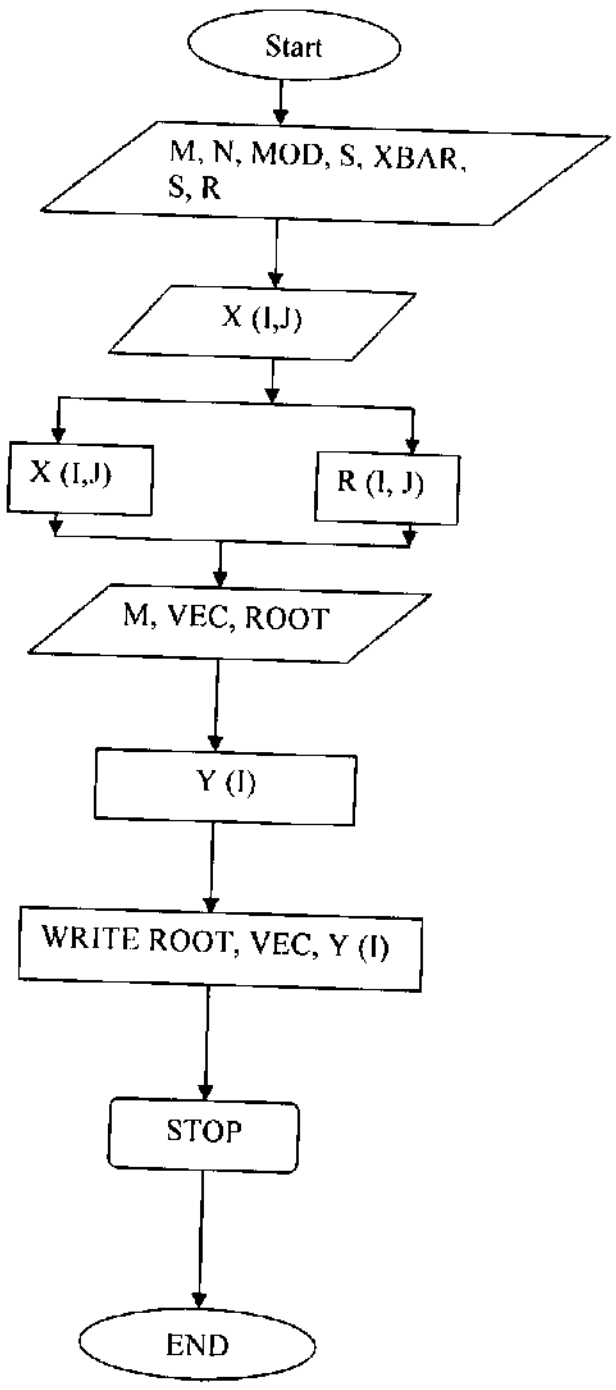


Figure 3.2: Flow chart of FORTRAN programme for Principal Component Analysis

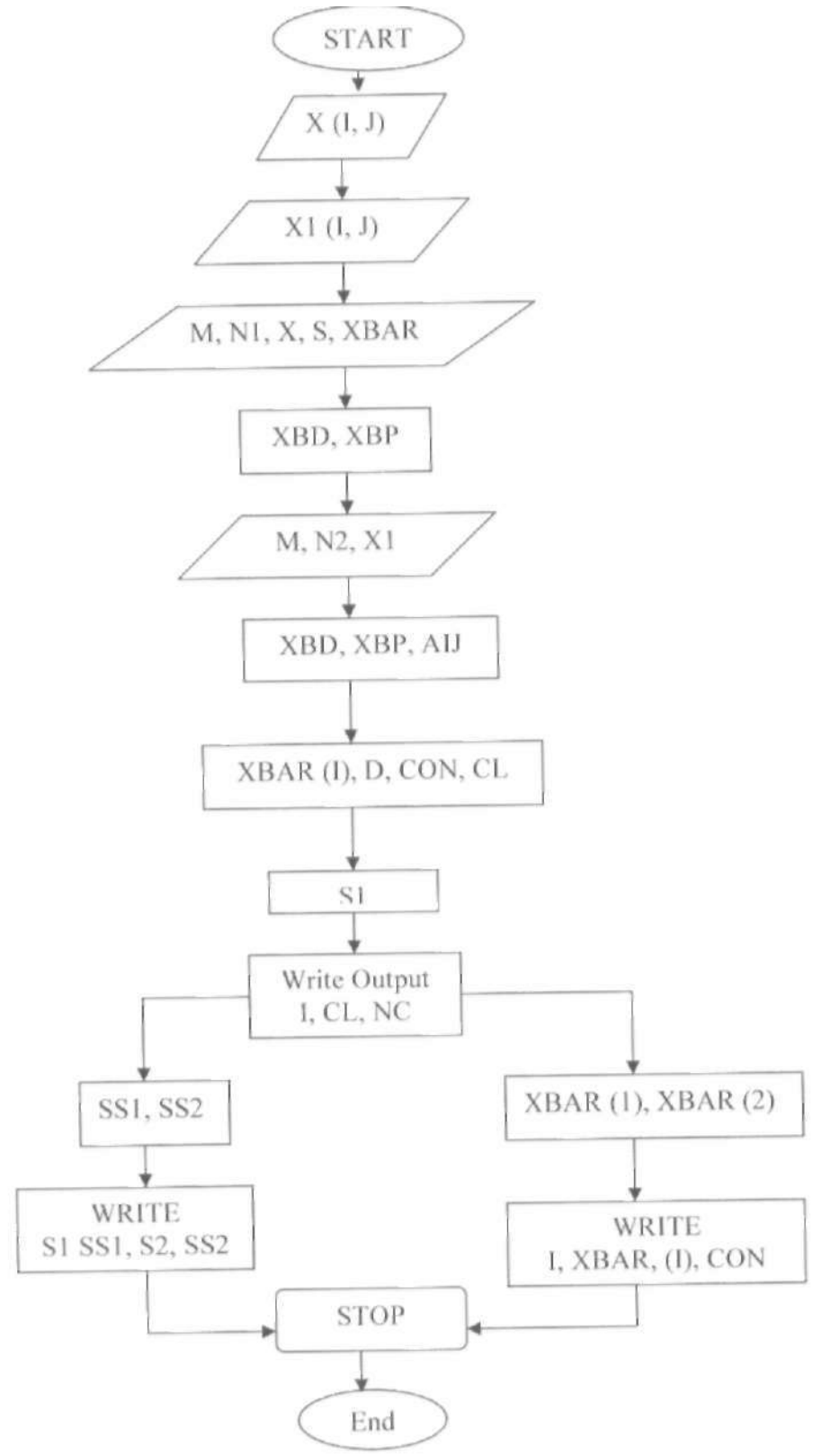


Figure 3.3: Flowchart of FORTRAN programme for Discriminant analysis

3.2 (d) Operation of the STATISTICA programme

STATISTICA stores its data in a disk file optimized for quick access and storage efficiency. The data are organized into cases and variables, where cases are equivalent to records in data base management programme (or rows of spread sheets) and variables are the equivalent columns of a spread sheet. Therefore each case consists of a set of values or variables. Case names are used as labels only (and not text values), thus they can contain all printable characters (the same as column labels. The STATISTICA programme has its own peculiar operational menu. It also has a data file manager, option for analysis type and a switch between analyses. Data is stored in spreadsheet only and can be opened from the STATISTICA menu by a specific command. A dialogue box appears on the operation of any of the methods of analysis which require specific instructions for the required analysis to be carried out. Result interpretations are based on the operators experience and understanding of the multivariate analysis techniques.

4.0 GENERAL

The graphs of the values of the parameters obtained at each station and in each month was plotted to show the time and space variants as well as to justify the averaging of this data for analysis. Figure 4.1(a) and 4.1 (b) shows the values for pH in 2005 and 2006 respectively, while figure 4.2 (a) and 4.2 (b) shows the plot for Electrical conductivity for 2005 and 2006 respectively. Figure 4.3(a) and 4.3(b) is for 2005 and 2006 temperature. Figure 4.4(a) and 4.4(b0 is for total dissolved solids in 2006 and 2007. Others are figures 4.5(a) and 4.5(b) for 2005 and 2006 Nitrates. 2005 and 2006 Phosphates are plotted in figure 4.6(a) and 4.6 (b). Similarly, the sulphate values are plotted in figures 4.7(a) and 4.7(b) for 2005 and 2006. The plot of 2005 and 2006 dissolved oxygen is plotted in figures 4.8(a) and 4.8(b), Chemical oxygen demand and biochemical oxygen demand are plotted in figures 4.9(a) and 4.9(b); and figures 4.10(a) and 4.10(b) for 2005 and 2006 respectively.

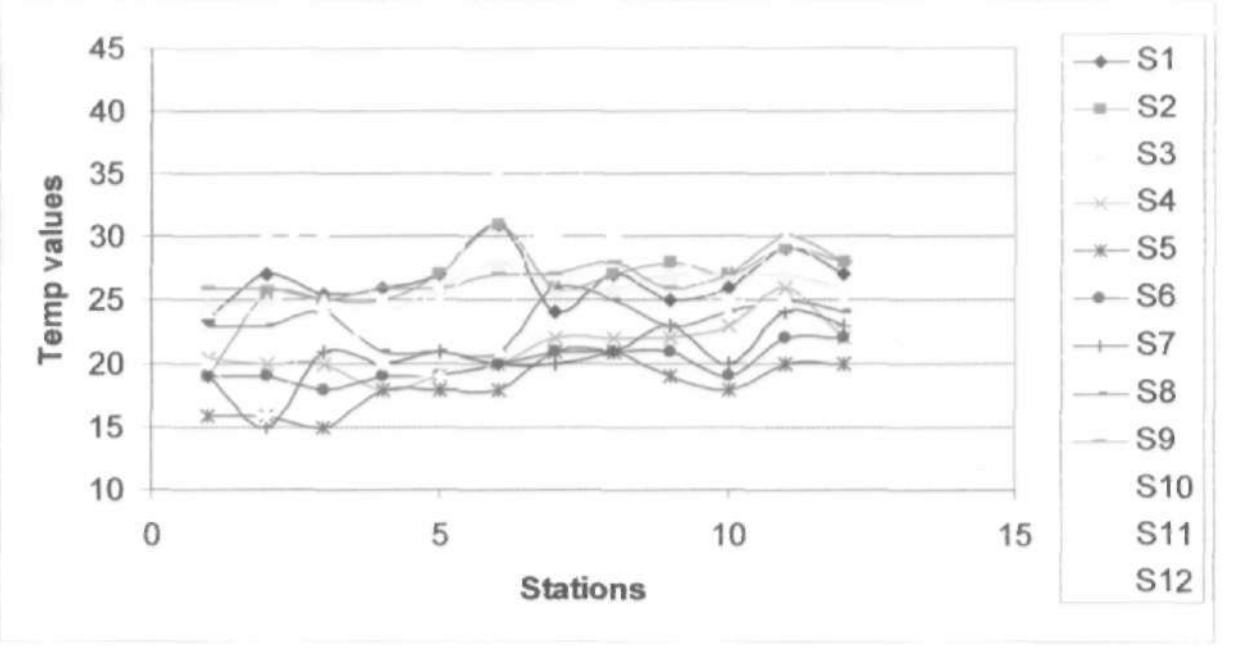


Figure 4.1(a): Plot of 2005 temperature data at each station

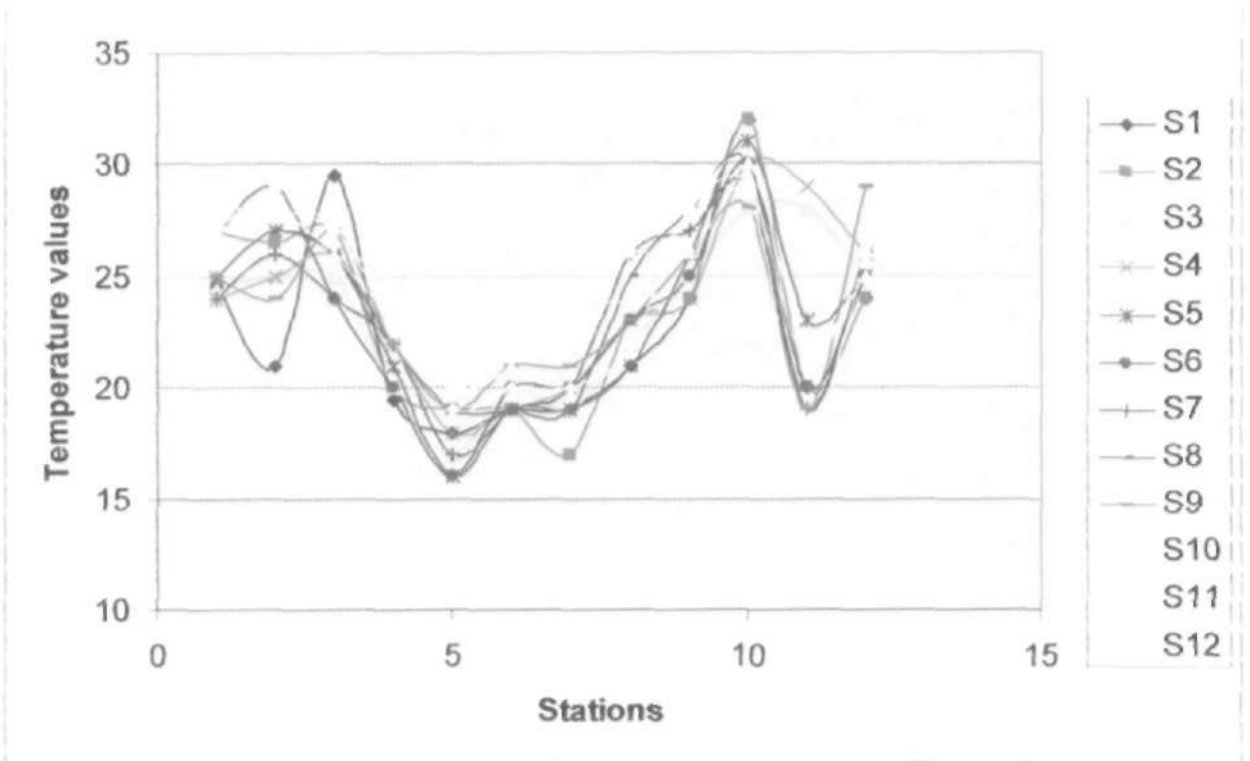


Figure 4.1 (b): Plot of 2006 temperature data at each station

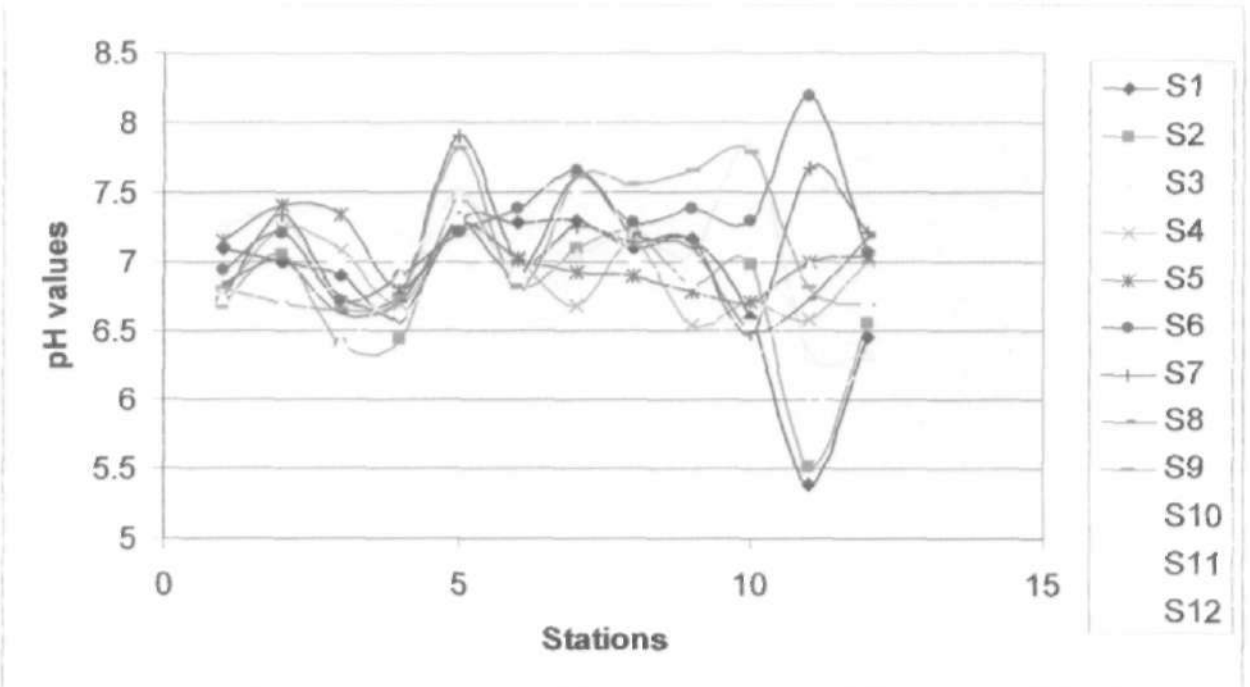


Figure 4.2 (a): 2005 pH values at each station

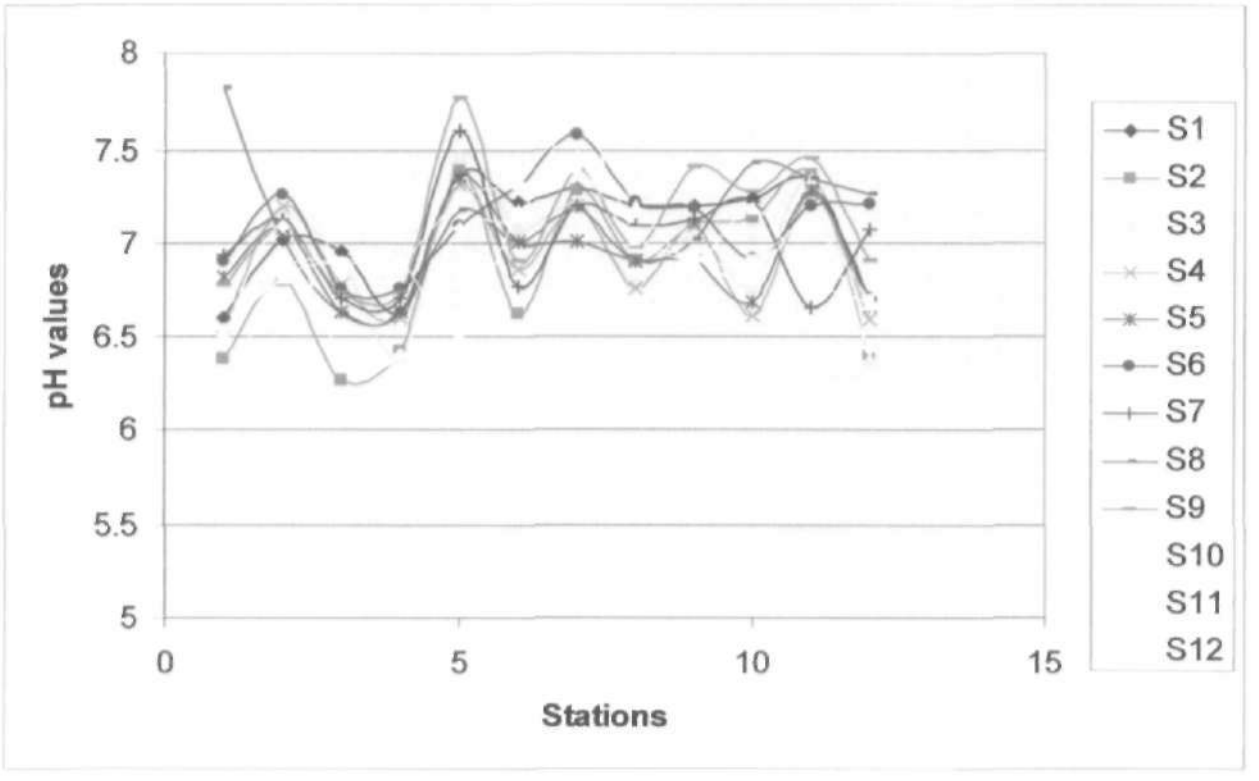


Figure 4.2 (b): 2005 pH values at each station

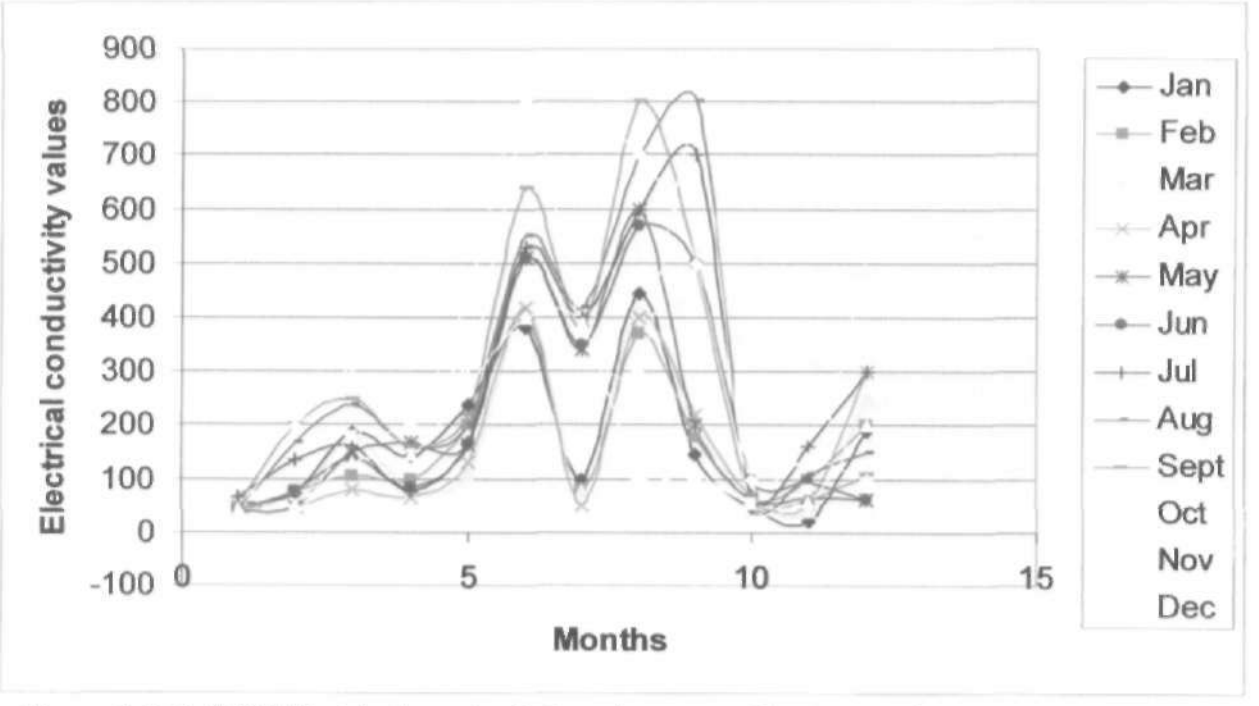


Figure 4.3(a): 2005 Electrical conductivity values according to months

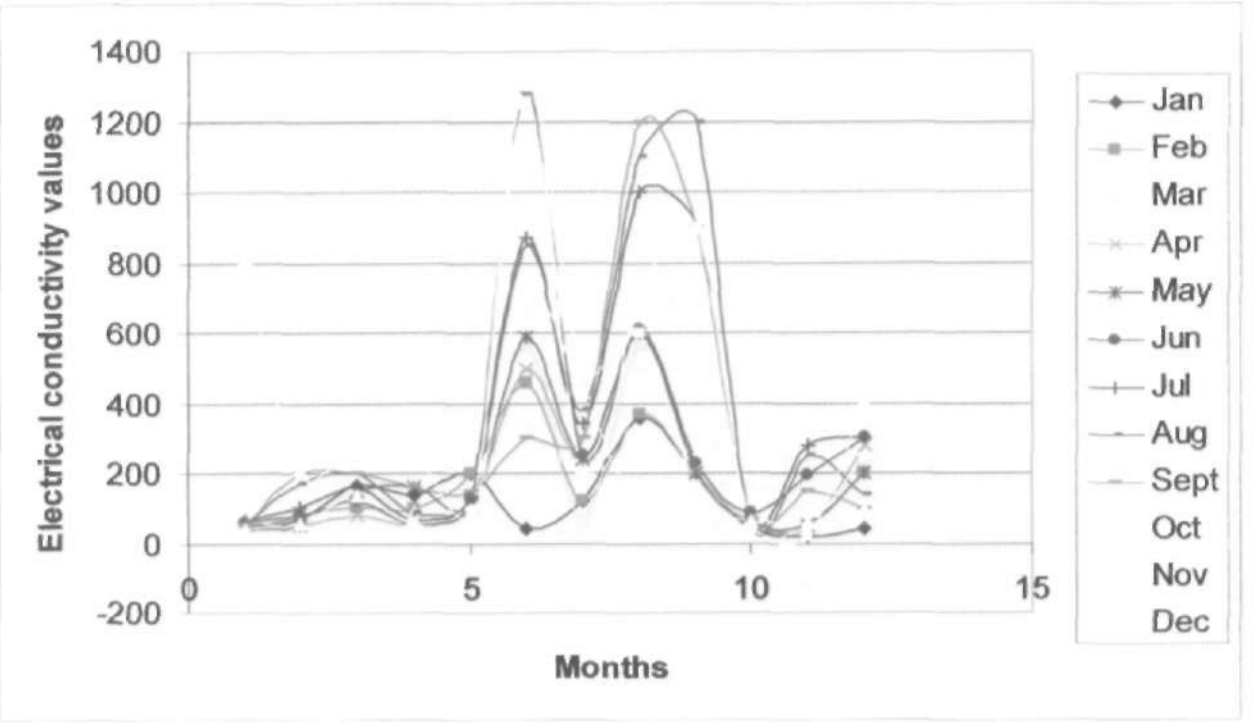


Figure 4.3(b): 2006 Electrical conductivity values according to months

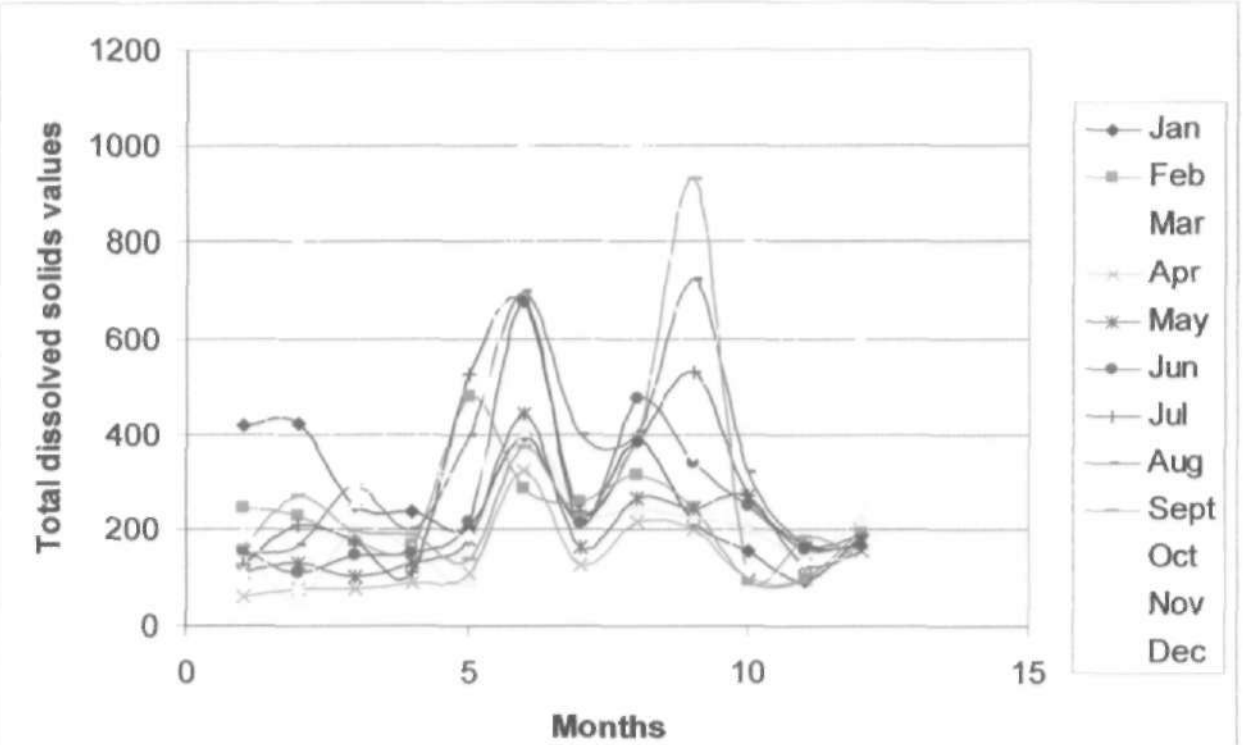


Figure 4.4(a): 2005 total dissolve solids values according to months

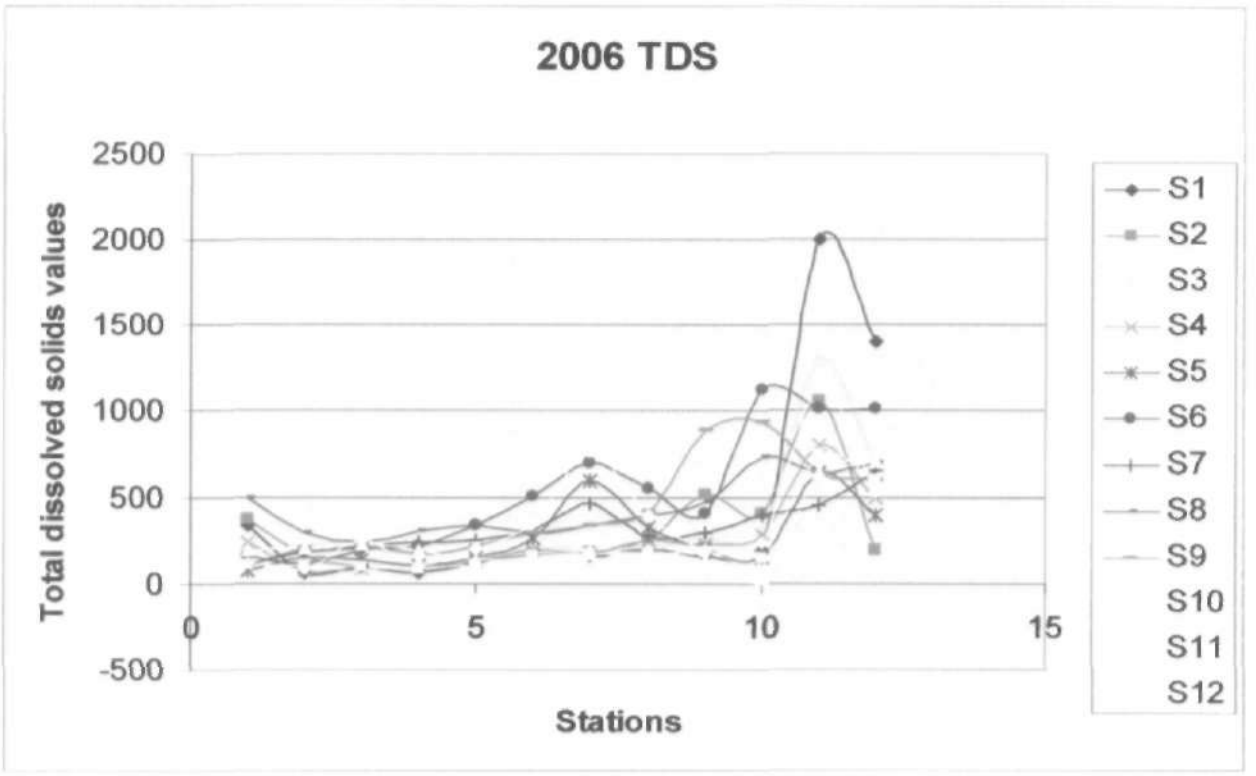


Figure 4.4(b): 2006 total dissolve solids values according to stations

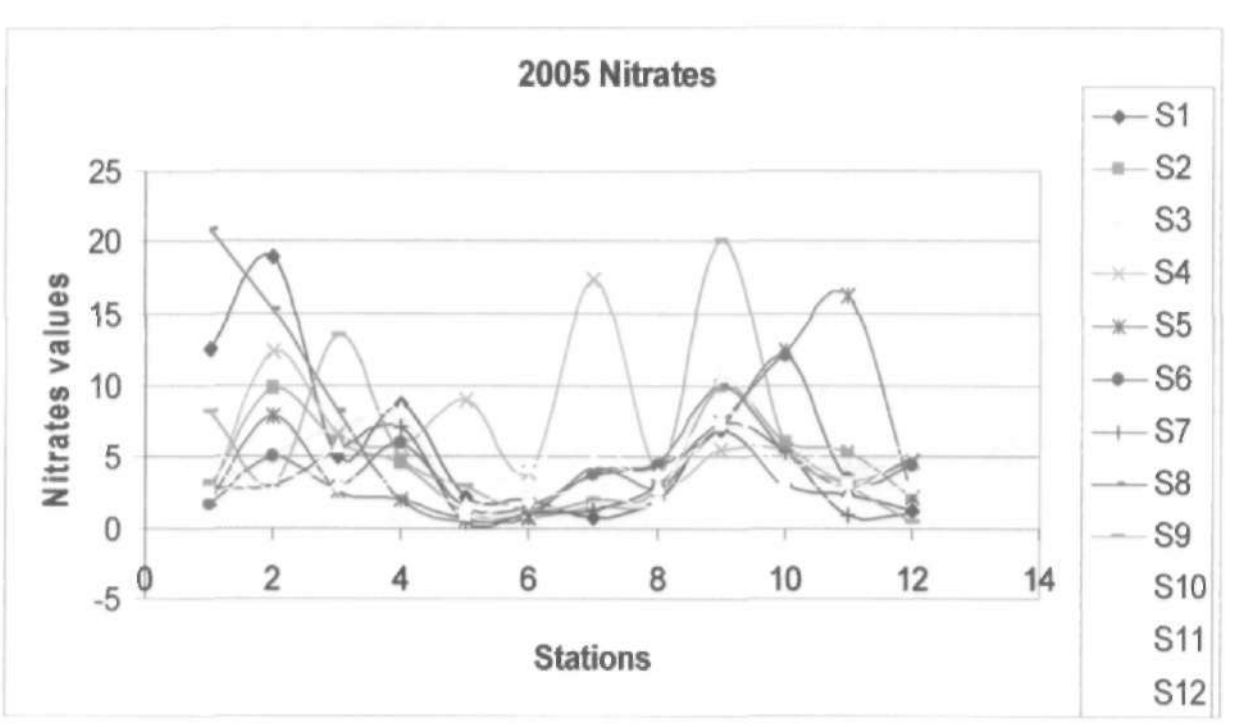


Figure 4.5(a): 2005 Nitrate values according to stations

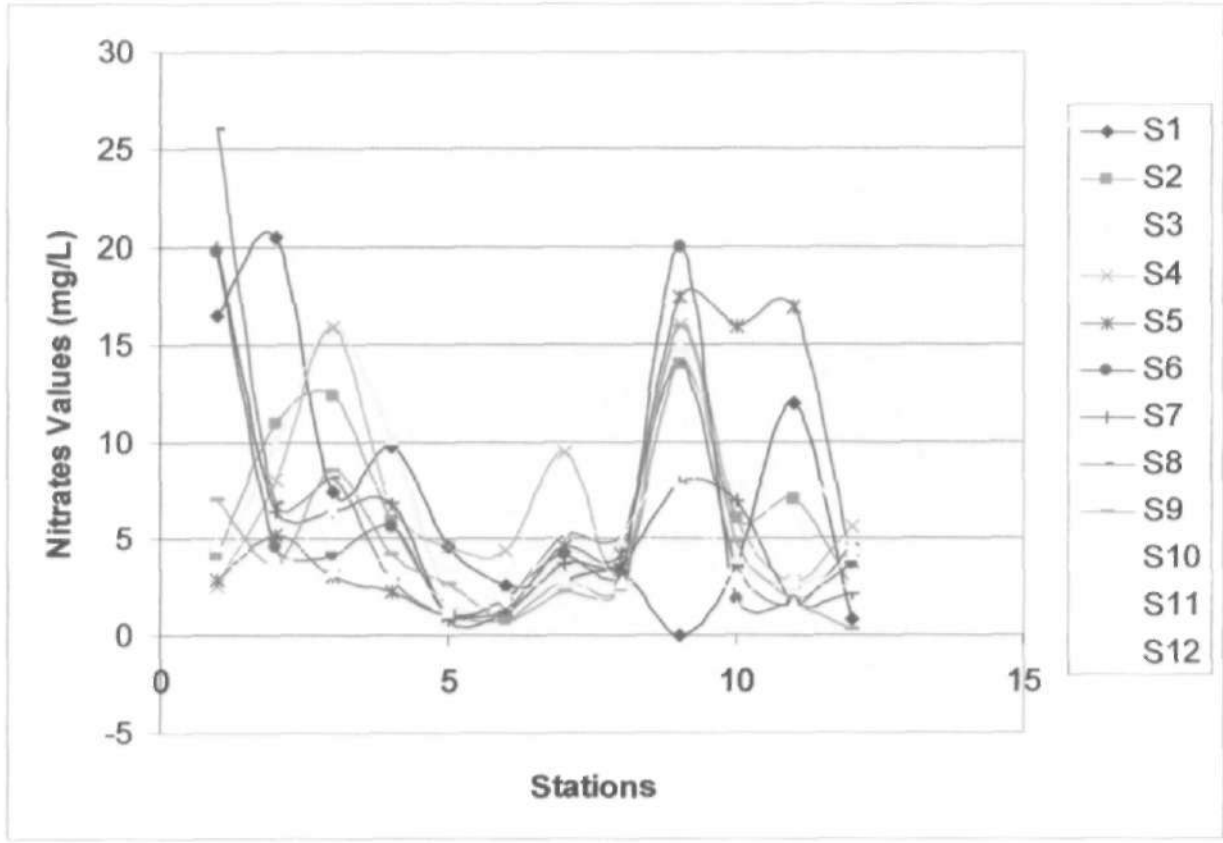


Figure 4.5(b): 2006 Nitrate values according to stations

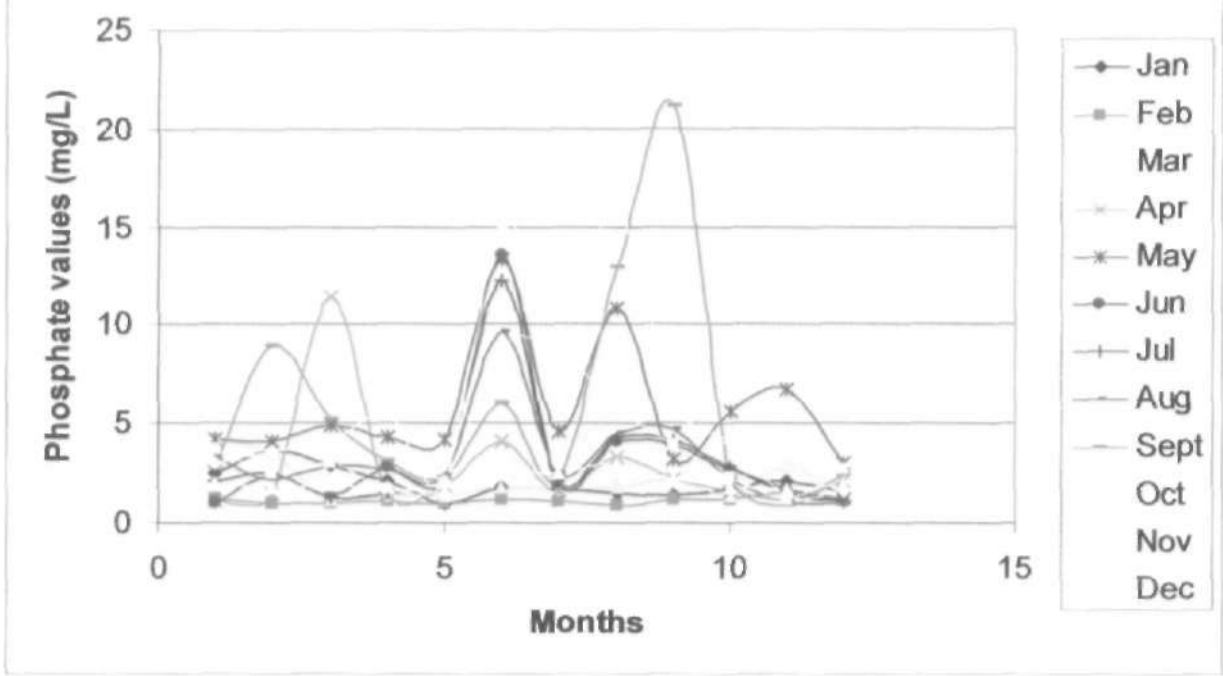


Figure 4.6(a): 2005 Phosphate values according to months

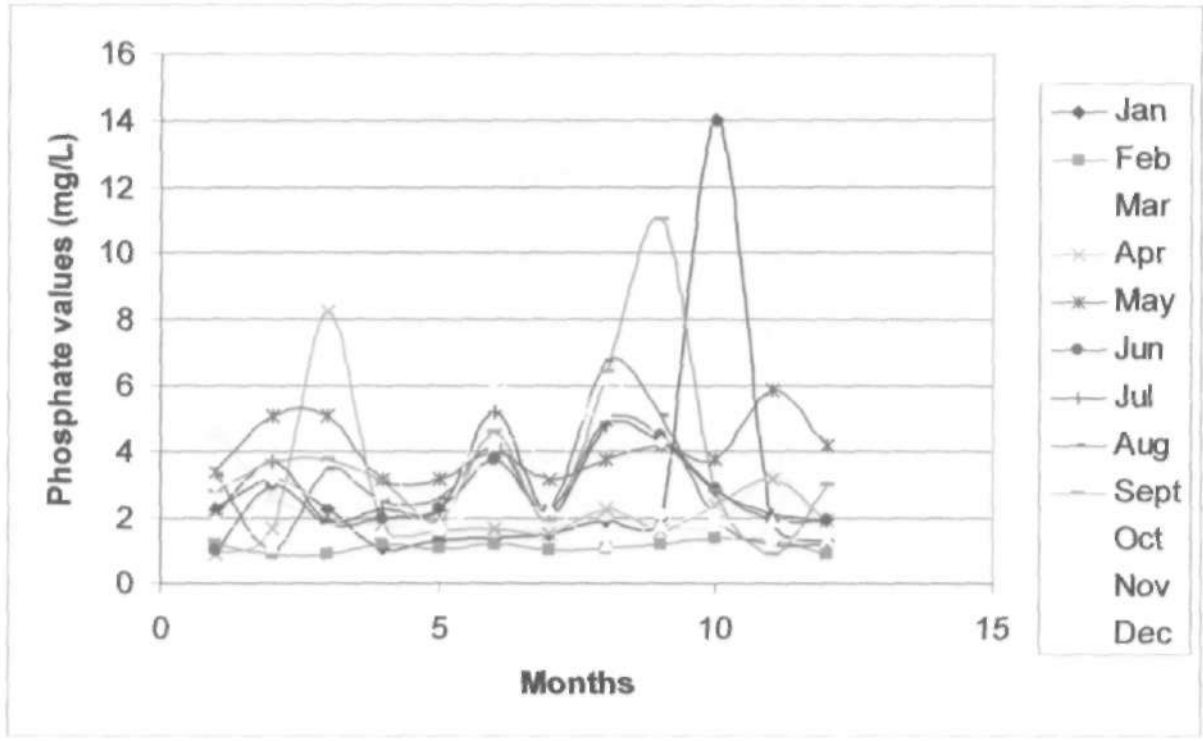


Figure 4.6(b): 2006 Phosphate values according to months

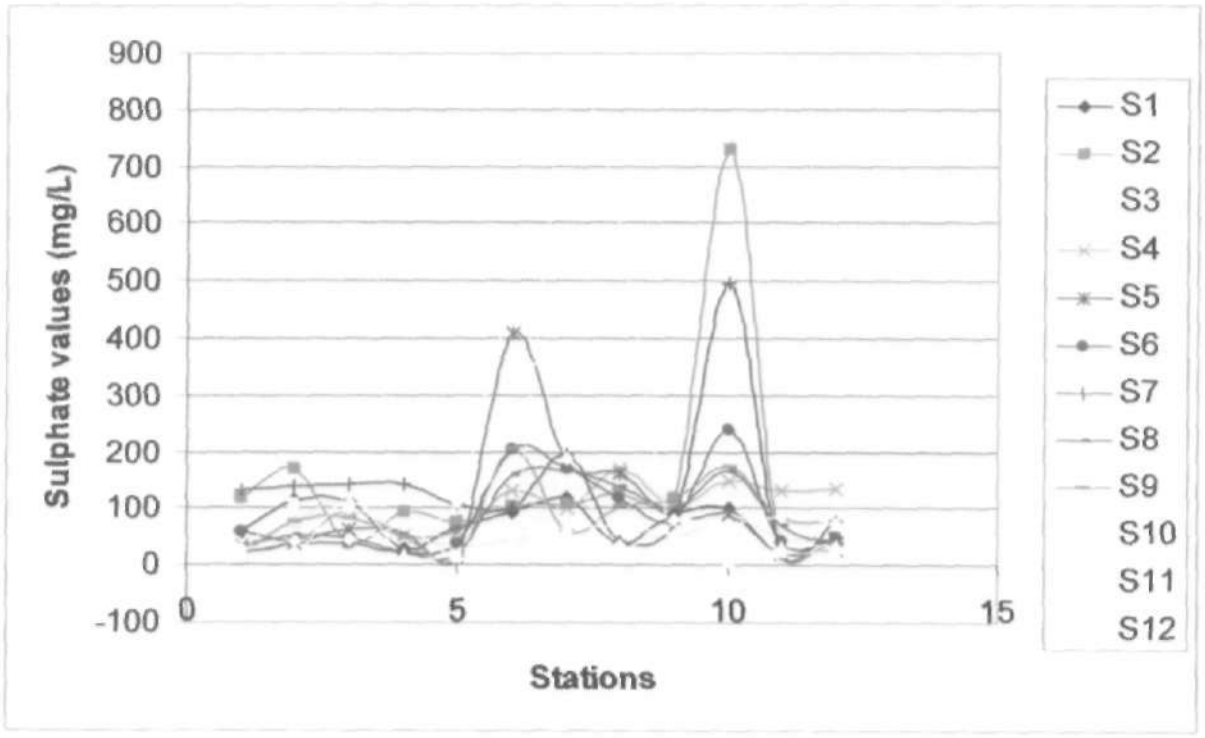


Figure 4.7(a): 2005 Sulphates values according to stations

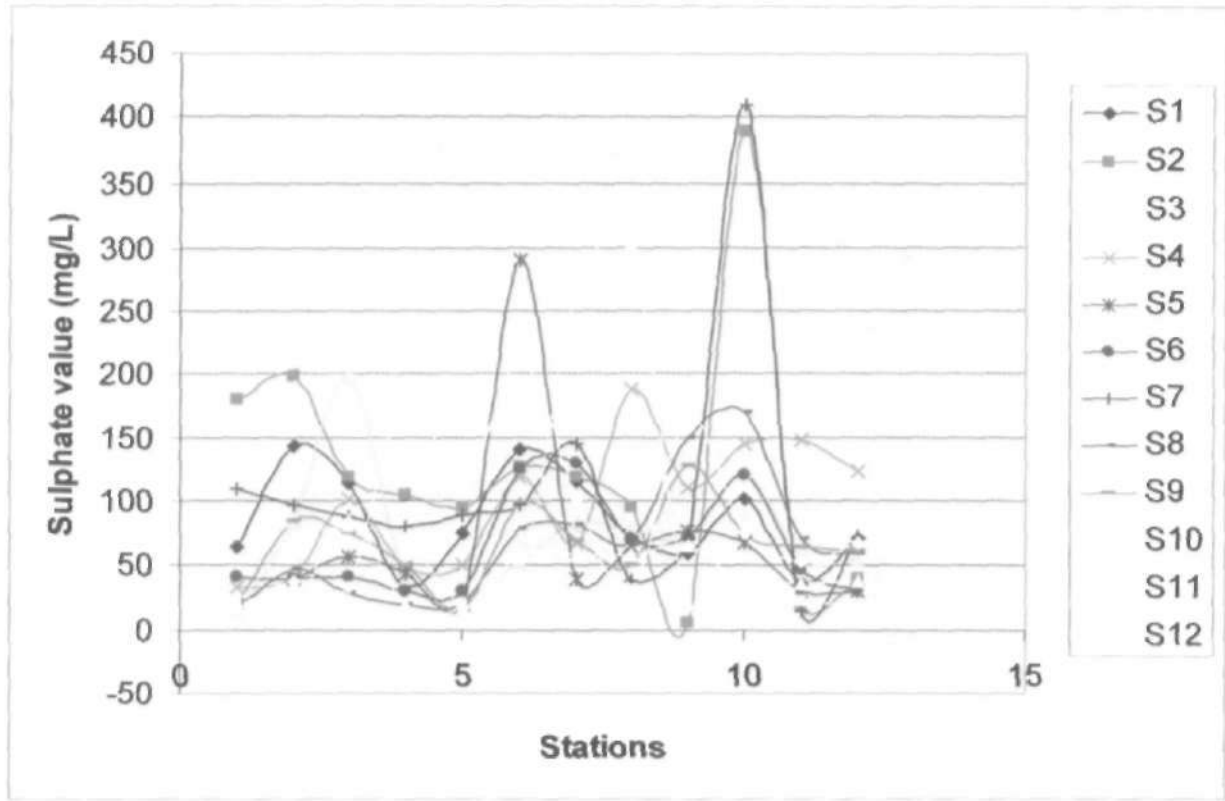


Figure 4.7(b): 2006 Sulphate values according to stations

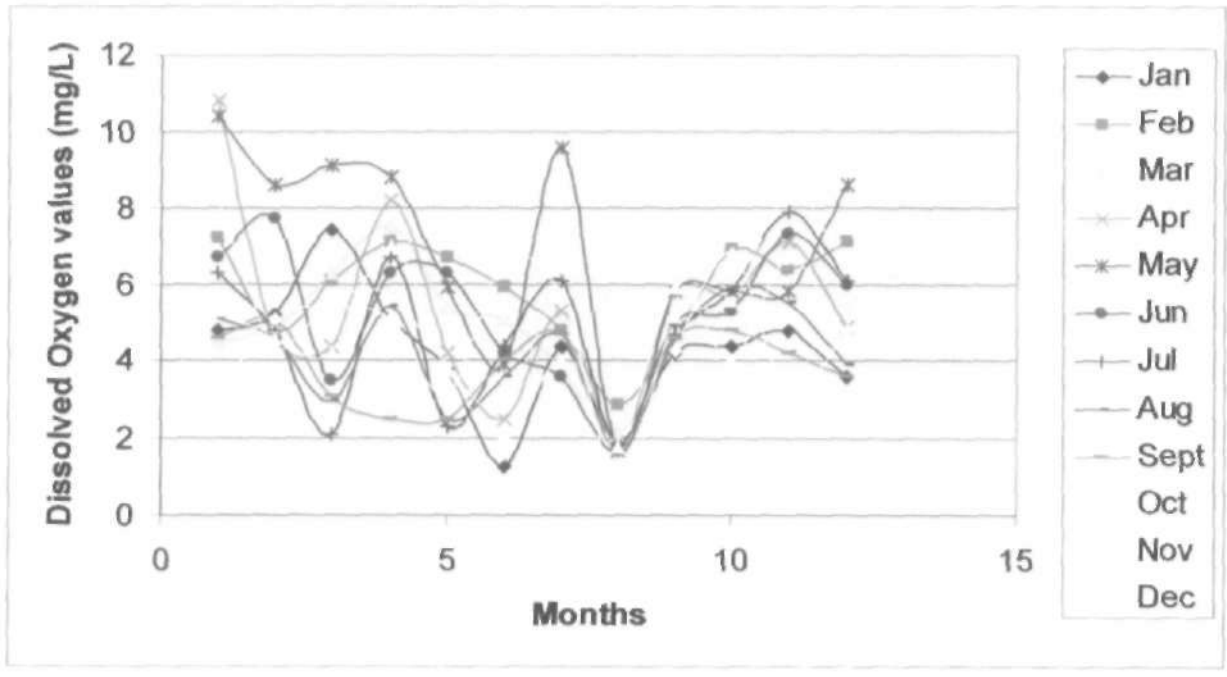


Figure 4.8(a): 2005 Dissolved Oxygen values according to months

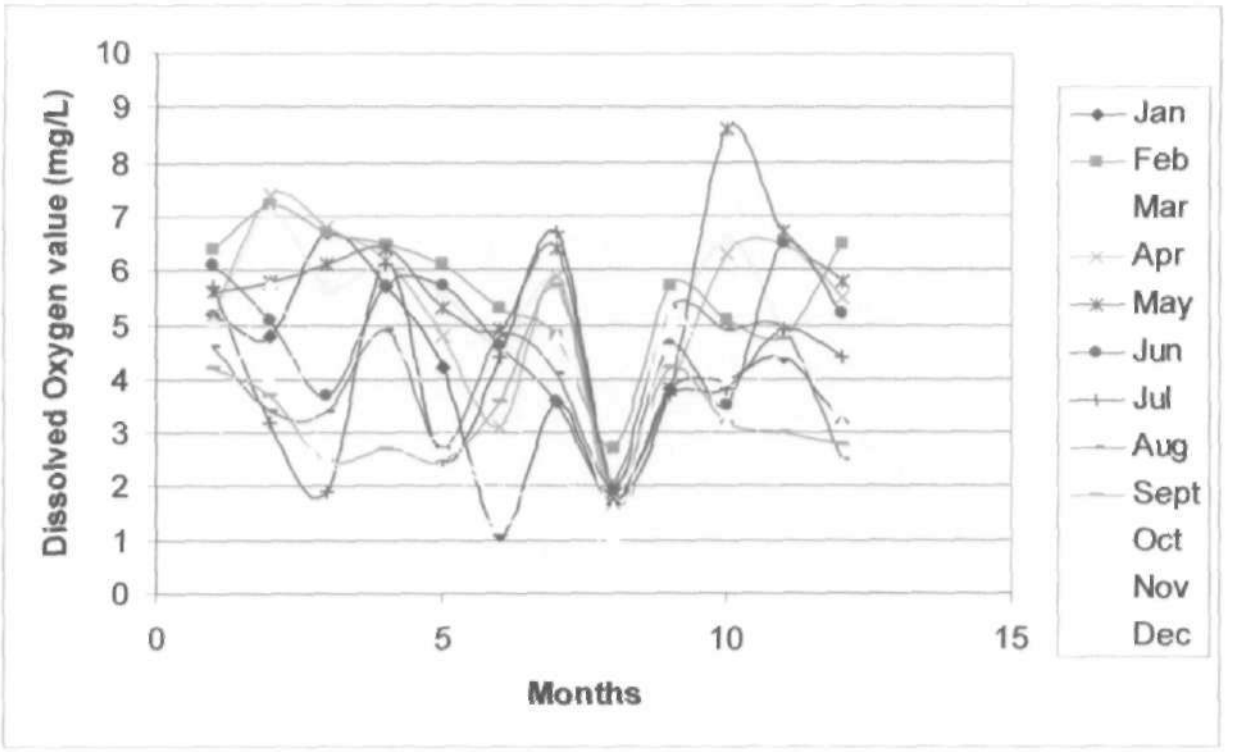


Figure 4.8(b): 2006 Dissolved Oxygen values according to months

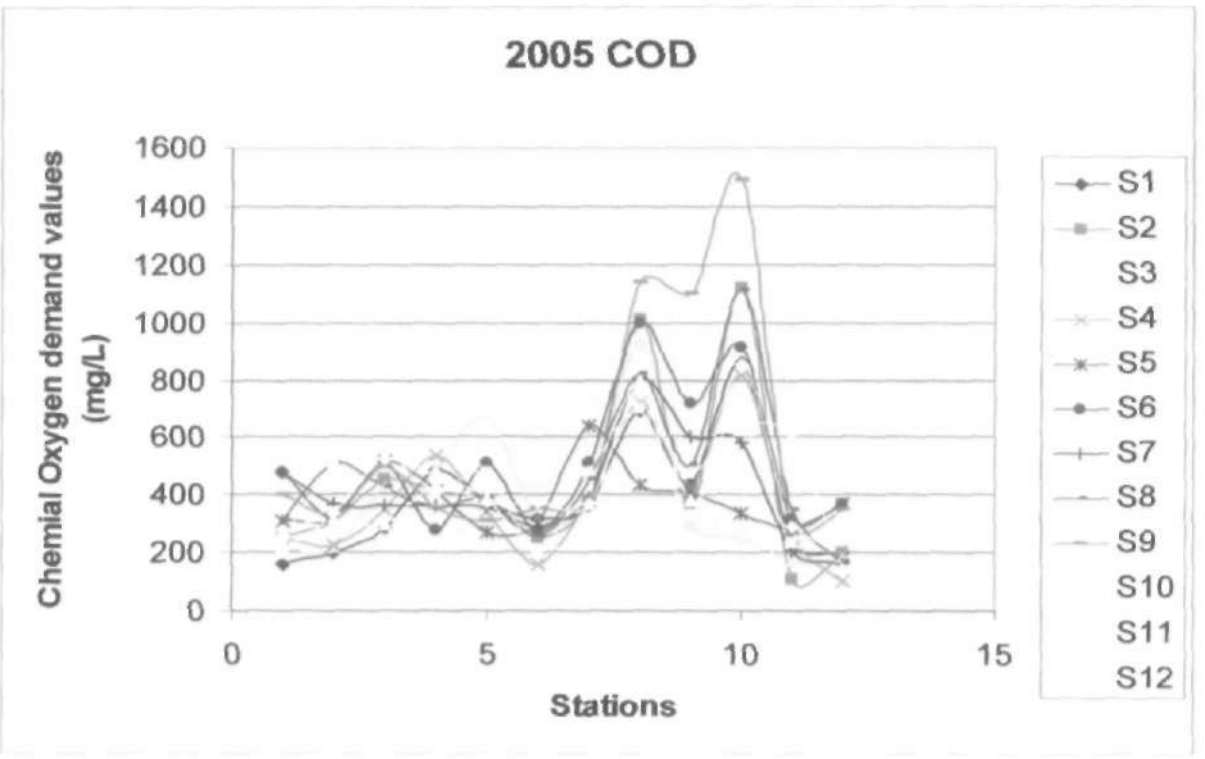


Figure 4.9(a): 2005 Chemical Oxygen Demand values according to stations

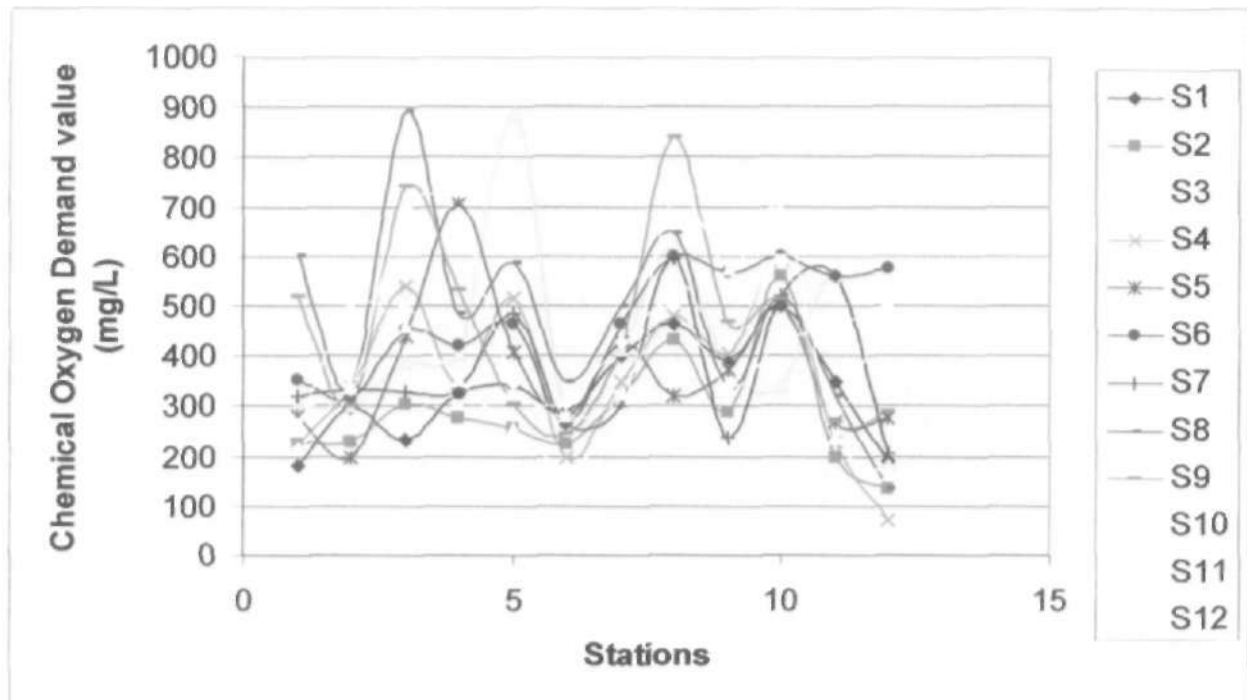


Figure 4.9(b): 2006 Chemical Oxygen Demand values according to stations

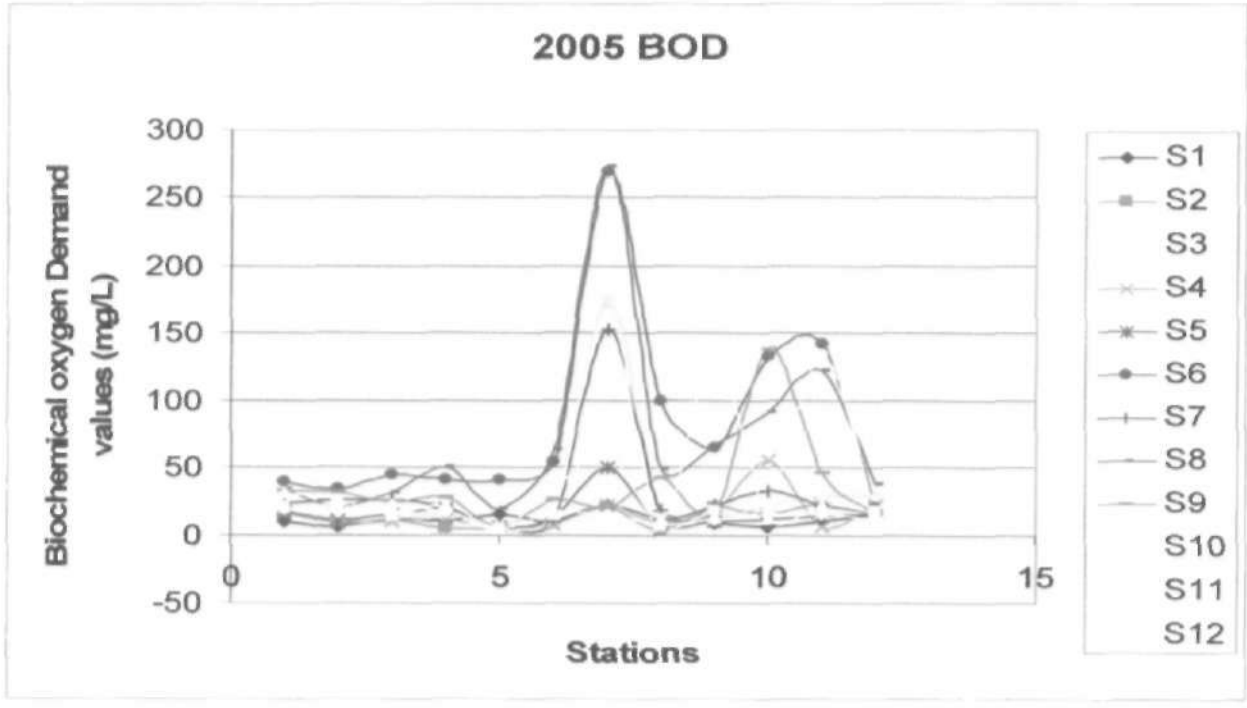


Figure 4.10(a): 2005 Biochemical Oxygen Demand values according to stations

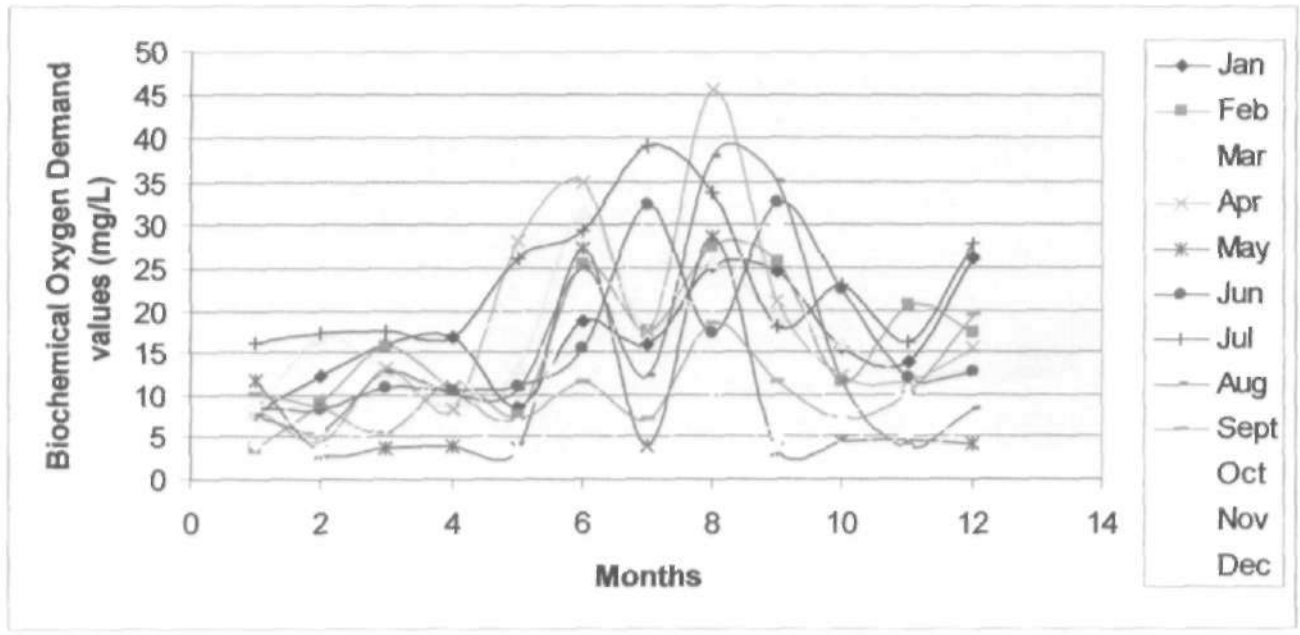


Figure 4.10(b): 2006 Biochemical Oxygen Demand values according Months

Table 4.1 which has average values for stations and table 4.1 which has average value for months were used for Cluster analysis, Factor analysis and Principal components analysis respectively. Discriminant analysis was based on the Clusters obtained. The values wee averaged on the basis that for each year samples were taken progressively from station one to station twelve. The climatic influence at each sampling run is the same at the time of sample collection and therefore any extraneous effect would be uniform. Distances between successive station is not more than 3 km which made it possible to conclude that uniform condition exist in them. The Clusters approximated to rainy and dry season, in the case of monthly clustering. In the case of sampling stations, they stations on River Shika and Kubanni were clearly identified. There were two stations located on River Galma. Though these were not identified as a Cluster, Discriminant analysis pointed them out as not belonging to any group. Table 4.1 contains the basic statistical data on the parameters. The parameters obtained were compared with the World Health Organization (WHO, 1993) and Federal environmental Protection Agency (FEPA, 1991) standards for river water.

Table 4.1: Mean values of Laboratory Results by Stations

PARAMETERS	STATIONS											
	1	2	3	4	5	6	7	8	9	10	11	12
pH	6.96	6.79	6.99	6.95	7.00	7.23	7.07	7.00	7.08	6.88	6.96	6.93
Conductivity (µs/m)	61.8	96.0	171.0	123.5	172.5	651.7	284.0	597.3	442.3	62.9	83.8	205.8
Temperature (^{o}C)	22.6	22.7	23.2	23.0	23.2	24.2	25.0	25.8	24.5	24.4	25.9	25.5
Total Solids (mg/L)	292.6	333.0	285.0	239.0	271.5	563.3	351.8	461.8	433.7	333.0	209.7	245.1
Nitrate (mg/L)	6.6	5.1	5.0	6.9	5.9	6.0	4.0	6.7	5.4	4.5	3.1	4.5
Phosphate (mg/L)	2.2	2.9	3.1	2.4	2.0	5.5	2.5	4.7	4.3	2.5	2.0	1.7
Sulphates (mg/L)	78.7	178	66.5	100.8	82	78.3	122.0	76.5	79	132.4	82.4	35.6
Dissolved Oxygen (mg/L)	5.7	5.2	4.6	5.6	3.6	3.7	5.0	1.7	4.7	5.1	5.6	4.8
COD (mg/L)	358.3	363.3	398.0	368.3	396.2	485.3	393.6	470.7	519.0	440.7	414.7	395.0
BOD (mg/L)	10.4	10.8	20.5	13.5	13.8	61.2	25.7	52.1	27.7	12.6	16.0	27.0

Table 4.2: Mean values of Laboratory Results by Months

PARAMETERS	MONTHS											
	JAN	FEB	MAR	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC
pH	6.74	7.06	6.78	6.68	7.41	7.04	7.4	7.05	7.07	6.92	6.95	6.74
Conductivity	165.3	163.8	175.4	157.5	327.9	254.2	299.6	311.7	316.3	390	317.3	199
Temperature	26.2	26.3	26.0	21.5	18.2	20	20.7	23.0	26.5	26.5	28.5	25.8
Total Solids	244.3	196.6	185.2	145.3	194.3	260	330.0	329.2	301.2	497.3	784.7	641.7
Nitrate	7.7	7.3	6.9	5.0	2.2	1.9	4.3	3,3	10.9	5.5	4.8	3.6
Phosphate	1.9	1.1	2.2	2.6	4.9	3.0	3.0	3.1	4.9	4.6	2.7	1.9
Sulphate	46.8	74.4	84.7	58.8	47.1	141.5	146.8	107.5	113.6	187.3	50	54.1
Dissolved Oxygen	4.1	5.8	5.2	5.4	6.5	5.0	4.5	4.0	3.6	3.6	4.4	3.9
COD	301.3	332.2	430.3	422	412	273.3	406.7	691.3	419.7	668	319.7	248.3
BOD	20.3	17.5	18.6	20	9.6	20	72	19.1	22.3	33.8	25.2	24.3

Table 4.3: Summary of Basic Statistics of the Laboratory Results

Parameter	Graph Symbols	Mean	Standard Deviation	Minimum	Maximum	WHO Standards	FEPA Standards
pH	X_1	6.99	0.24	6.32	7.41	6 – 9	6 – 9
Conductivity (µs/cm)	X_2	255.3	206	10	1100	-	1000
Temperature (°C)	X_3	24.1	3.3	16	32.7	32.5 – 37	40
Total Dissolved Solids (mg/L)	X_4	335	104.3	84	1050	-	2000
Nitrates (mg/L)	X_5	5.3	1.2	0.8	23.6	10	20
Phosphates (mg/L)	X_6	3.0	1.2	0.7	18.6	5	5
Sulphates (mg/L)	X_7	92.7	36.7	3.6	860	250	500
Dissolved Oxygen (mg/L)	X_8	4.6	0.9	1.0	8.4	4	4
COD (mg/L)	X_9	416.9	51.4	88	1000	500	200
BOD (mg/L)	X_{10}	24.3	16.5	4.8	249.6	20	30

4.1. Cluster Analysis:

4.1.1. Analysis of parameter average values in sampling Stations

(i) Input Data: Table 4.1 contains the average values of laboratory results for each station of survey. The table is the input data for the clustering analysis for station results.

- Clustering Method: Variables as contained in each column
- Amalgamation (linkage) rule: Single linkage
- Distance measure: Euclidean distances
- Missing data: Substituted by means
- Selection of variables for analysis: Select all

Joining Results

- Number of Variables 12
- Number of cases identified 847
- Rectangular branches are used for joining

Result obtained from the clustering of data in Table 4.1 following the data analysis procedure listed above is shown in Figure 4.11 below.

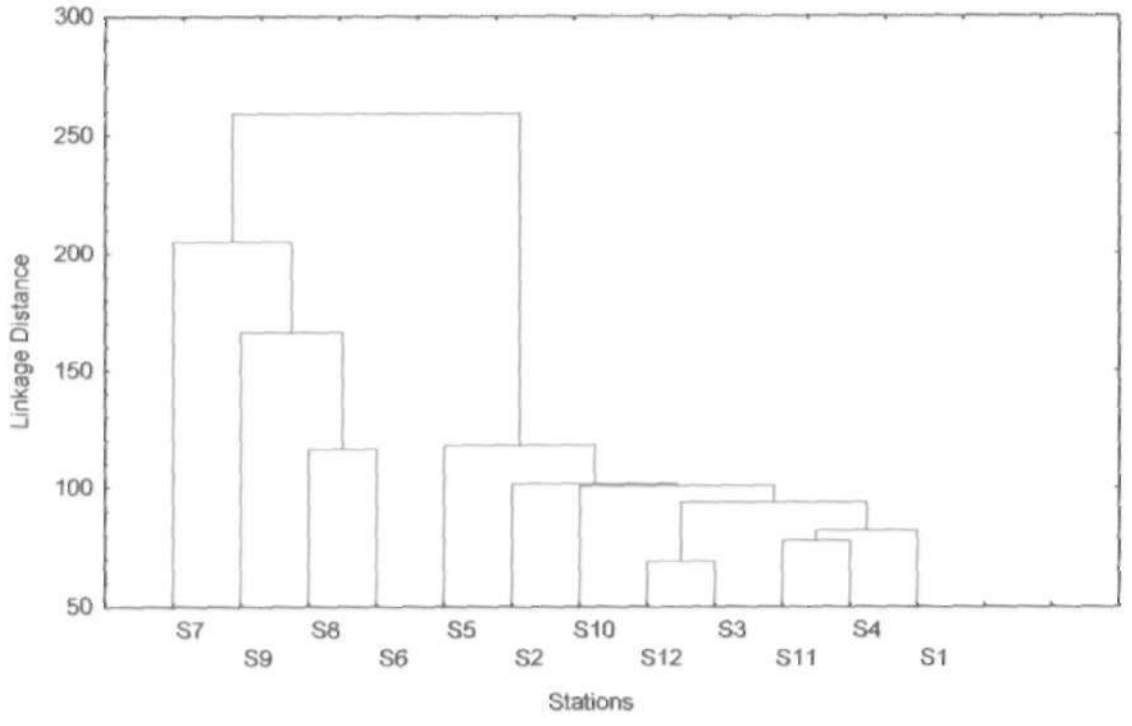


Figure 4.11: Cluster value Plot for average Laboratory values for stations

Figure 4.11 shows the cluster groups obtained. The clusters could be taken as two separate and distinct groups as follows:

- **Cluster 1: S7, S9, S8 and S6**
- **Cluster 2: S5, S2, S10, S12, S3, S11, S4 and S1**

Table 4.4 below shows the Euclidean and linkage distance values as obtained by the analysis.

Table 4.4 Plotting distances for cluster analysis for sampling stations

Stations	Linkage Distance	Euclidean distances	
S7	206	1	2.75
S9	167	2	3.48
S8	115	3	3.97

S6	260	1.9	6.07
S5	119	5	7.19
S2	101	6	8.39
S10	102	7	9.83
S12	70	8	9
S3	93	8.53	11.24
S11	78	10	11
S4	81	10.56	11.97
S1	82	11.97	-

4.1.2. Analysis of Average values for months

(ii) Input Data: Table 4.2 contains the average values of laboratory results for each month of survey. The table is the input data for the clustering analysis for monthly results.

Clustering Method:

- Variables used are as contained in each column
- The amalgamation (linkage) rule used was Single linkage method
- Joining distance was measured as Euclidean distances
- Missing data were substituted by the mean of the data
- All the variables were selected for the analysis

Joining Results

- The number of variables outputted was 12
- The number of cases identified was 602
- Rectangular branches were used for joining the identified value points

Result obtained from the clustering of data in Table 4.2 following the data analysis procedure liste above is shown in Figure 4.12 below.

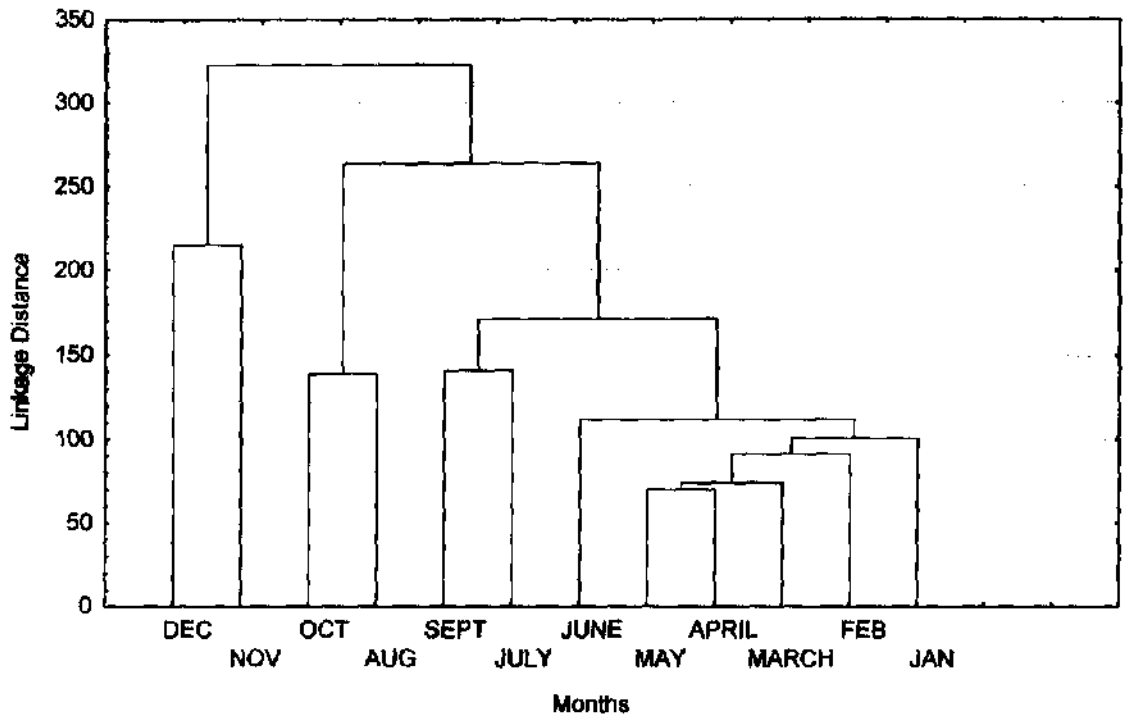


Figure 4.12: Cluster Value Plot for average Laboratory values for months

Table 4.5 below shows the Euclidean and linkage distance values as obtained by the analysis. Figure 4.12 shows the cluster groups obtained. The clusters could be taken as two separate an distinct groups as follows:

- **Cluster 1: Dec. Nov. Oct., Aug. Sept and July**
- **Cluster 2: June. May. April. March. Feb and Jan.**

Table 4.5 Plotting distances for cluster analysis for months

Months	Linkage Distance	Euclidean distances	
Dec	214	1	1.97
Nov	323	1.46	5.36
Oct	138	3.03	3.98
Aug	264	3.53	7.24
Sept	142	5	5.97
July	171	5.52	9.03
June	111	7	11.05
May	72	8	9
April	73	8.53	9.97
March	91	9.24	10.73
Feb	100	10.16	12
Jan	100	12	-

4.2. Factor and Principal Component Analysis

4.2.1 Analysis of parameter average values for sampling stations

(i) Input data: The data in Table 4.1 represents average values for laboratory results obtained according to the sampling stations. This result was used for the factor and principal component analysis respectively.

- Missing data were case wise deleted
- 847 cases were processed (selected). This represented the full array of data
- 847 valid cases were accepted for analysis
- Correlation matrix was computed for 12 variables
- The maximum number of factors that were expected is 4

- The minimum acceptable eigenvalues is 1.000

Correlation matrix is ill-conditioned and the determinant of the correlation matrix is equal to zero. The maximum number of principal components that can be extracted will be equal to the number of positive eigenvalues for the matrix.

From the analysis of the data the following results were obtained:

- Only one eigenvalues was significant with a value of 1 and above. This is shown in Table 4.6.

Table 4.6: Eigenvalues obtained from station mean results analysis

Value	Eigenvalue	% total Variance	Cumulative value	Cumulative %
1	10.9319	91.0975	10.9319	91.0995

Table 4.7 shows the principal components extracted according to sampling stations and their significance. Column 1 shows the stations, factor score coefficient, communalities and factor loading describe the relationship between individual laboratory values. But because the analysis is basically for the parameters of significance in the station sampling runs these values are all close. This explains that there is a relationship between the stations.

Table 4.7: Principal Components for station mean laboratory results

Stations	Factor Score Coefficient	Communalities	Factor loading
S1	0.08729	0.9106	0.9543
S2	0.08619	0.8878	0.9422
S3	0.0906	0.9812	0.9906
S4	0.09005	0.9692	0.9845

S5	0.09075	0.9841	0.9920
S6	0.08144	0.7926	0.8903
S7	0.08745	0.9143	0.9562
S8	0.08192	0.8019	0.8955
S9	0.08851	0.9361	0.9675
S10	0.08645	0.8931	0.9450
S11	0.08624	0.8889	0.9428
S12	0.0902	0.9721	0.9860

Because there is only one eigenvalue, only one set of factors is extracted from the analysis. This basically explains the factors in the study that are significant, while the principal component shows the stations that are significant. Table 4.8 shows the factors against the study parameters. The highlighted factors and parameters are taken as the most significant.

Table 4.8: Factors extracted for station mean laboratory results

Parameters	Factors
pH	-0.7875
Conductivity	0.7833
Temperature	-0.6544
Total dissolved solids	**1.6366**
Nitrates	**-0.8027**
Phosphates	**-0.8213**
Sulphates	-0.1572
DO	**-0.8057**
BOD	-0.6743
COD	**2.2831**

4.2.2 Analysis of parameter average values for months

(ii) Input data: The data in Table 4.2 represents average values for laboratory results obtained according to each sampling months. This result was used for the factor and principal component analysis respectively.

- Missing data are casewise deleted
- 14 cases were processed (selected), 10 before weighting
- 10 valid cases were accepted
- Correlation matrix was computed for 12 variables
- Maximum number of factors that are expected is 4
- Minimum acceptable eigenvalues is 1.000

Correlation matrix is ill-conditioned and the determinant of the correlation matrix is equal to zero. The maximum number of principal components that can be extracted will be equal to the number of positive eigenvalues for the matrix.

From the analysis of the data the following results were obtained:

- Two eigenvalues are significant in terms of the above established criteria. This is shown in Table 4.9.

Table 4.9: Eigenvalues obtained from monthly laboratory mean values results analysis

Value	Eigenvalues	% total Variance	Cumulative value	Cumulative %
1	10.7219	89.3494	10.7219	89.3494
2	1.0097	8.4138	11.7316	97.7632

Table 4.10 shows the principal components extracted according to months and their significance. It should be noted that unlike in Table 4.7 where only one eigenvalues was acceptable, in this case there are two eigenvalues. Hence the principal component analysis gave values of factor score, communalities and factor loadings according to each of the two eigenvalues.

Table 4.10: Principal Components for monthly mean values

Months	Factor Score Coefficient		Communalities		Factor loading	
	Factor1	Factor2	Facto1	Factor2	Factor1	Factor2
Jan	0.9136	-0.7924	0.9596	0.9660	0.9796	-0.0800
Feb	0.0926	0.1044	0.9864	0.9975	0.9932	0.1054
March	0.0889	0.2873	0.9082	0.9923	0.9530	0.2901
April	0.0877	0.3136	0.8843	0.9885	0.9404	0.3227
May	0.0909	0.1859	0.9505	0.9857	0.9749	0.1877
June	0.0862	-0.1195	0.8547	0.8693	0.9245	-0.1207
July	0.9199	-0.7049	0.9729	0.9780	0.9864	-0.0712
Aug	0.0904	0.2164	0.0513	0.9991	0.9754	0.2185
Sept	0.0925	0.0154	0.9844	0.9846	0.9922	0.01554
Oct	0.0926	0.1039	0.9876	0.9986	0.9938	0.1049
Nov	0.0740	-0.5811	0.6300	0.9838	0.7938	-0.5948
Dec	0.0753	-0.5743	0.6519	0.9881	0.8074	-0.5786

A further analysis of the principal components obtained in Table 4.10 reveals two sets of factors. These factors show the relative significance of each parameter measured according to months. The significance is measured according to the numerical values obtained irrespective of whether positive or negative. Table 4.11 shows the values of the factors extracted for the monthly laboratory results.

Table 4.11: Factors extracted for monthly mean laboratory results

Parameters	Factors Analysis	
	Factor 1	Factor 2
pH	0.6632	0.0709
Conductivity	0.8386	-0.2507
Temperature	-0.6081	0.0875
Total solids	**1.232**	**-2.3823**
Nitrates	-0.7173	0.1202
Phosphates	**-0.7335**	**0.7050**
Sulphates	-0.4950	0.2901
DO	-0.7337	0.1098
BOD	-0.5924	0.0832
COD	**2.0098**	**1.7666**

The Scree plot shown in Figure 4.13 shows that of the four eigenvalues selected only two are greater than 1.000. Therefore, number of factors extracted is 2. The eigenvalues are 10.7219 and 1.0097. This basically explains the factors in the study that are significant, while the principal component shows the months that are significant.

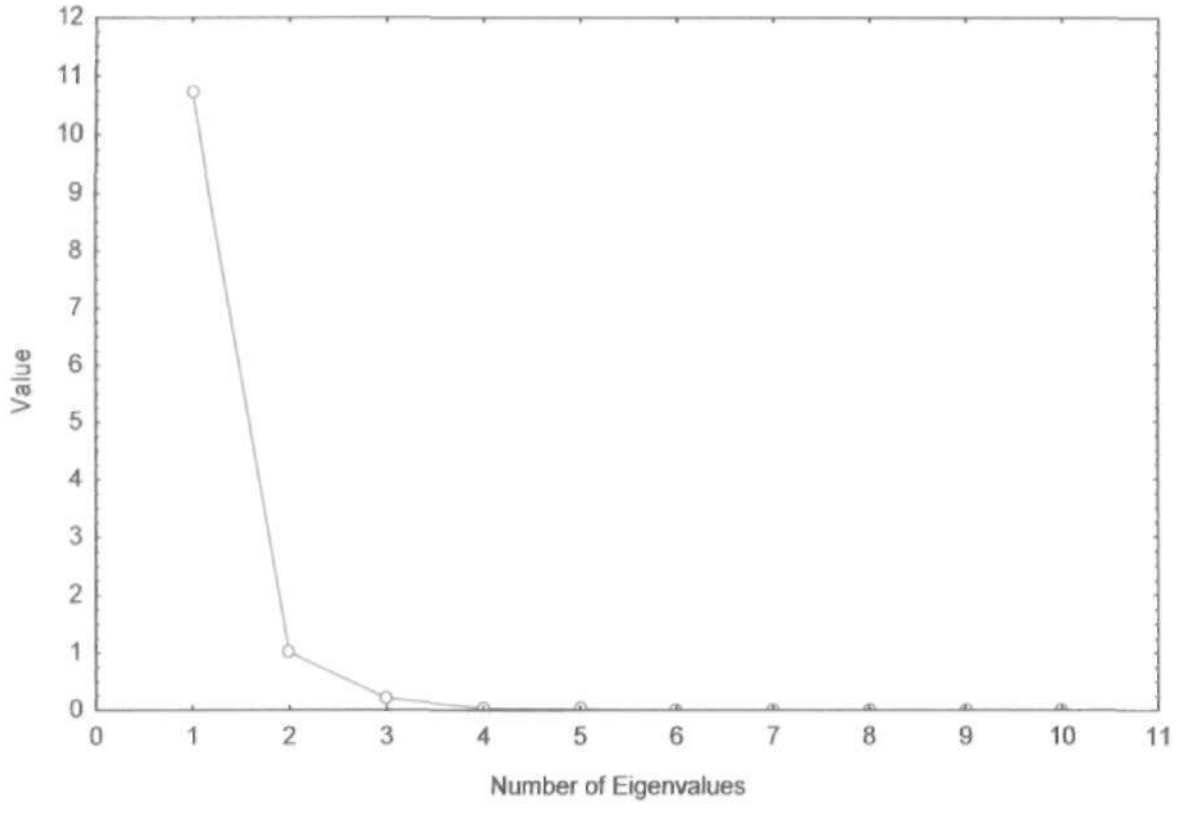


Figure 4.13: Scree plot for eigenvalues for mean values for months

4.3. Discriminant Analysis

When performing a multiple group Discriminant Analysis (Canonical Analysis) we do not have to specify how to combine groups so as to form different Discriminant functions. The programme automatically determines optimal combination of variables so that first function (root) provides the most overall discrimination between groups, the second, second most and so on. Their contribution to discrimination within groups will not overlap. The maximum number of roots computed is equal to the number of groups minus 1.

The input data for discriminate analysis was obtained from the indicated clusters shown in Figures.4.11 and 4.12 respectively. For both the analysis of mean laboratory results for stations and

months, two distinct clusters were indicated. Discriminant analysis was thus performed on the basis of these groups of clusters as follows:

4.3.1Analysis of mean laboratory results for sampling stations:

Cluster 1: S7, S9, S8 and S6

- Input is variables as indicated by the cluster 1 above while cases are the various parameters.
- Missing data are mean substituted.
- Variables are chosen as follows: First list S7, S9, Second list S8, S2

Output

Table 4.12: Chi-Square test with successive roots removed for cluster 1 (stations)

	Canonical R	Canonical R. Sqr.	Chi. Sqr.	Df	P	Lambda Prime
0	0.9788	0.9581	21.1643	4	0.0003	0.0385
1	0.2818	0.0794	0.5378	1	0.4633	0.9206

Table 4.13: Summary of canonical input performance for cluster 1 (stations)

	Left side	Right Side
No. of Variables	2	2
Variance Extracted	100%	100%
Total redundancy	91.3691%	94.9425%
Variables 1	S7	S8
Variable 2	S9	S6

Table 4.14: Canonical Roots for cluster 1 (stations)

Set	Variables	Root 1	Root 2
Right side	S8	-1.6265	12.2805
	S2	0.6298	-12.3717
Left side	S7	0.1333	-3.5664
	S9	-1.1272	3.3862

Table 4.15: Eigenvalues for cluster 1 (stations)

	Root 1	Root 2
Variance Extracted	0.9582	0.07941

Cluster 2: S5, S2, S10, S12, S3, S11, S4 and S1

Input

- Variables as indicated by the cluster 2 above while cases are the various parameters.
- Missing data are mean substituted.
- Variables are chosen as follows: First list S5, S2, S10, S12 and second list S3, S11, S4, S1

Output

Table 4.16: Chi-Square test with successive roots removed for cluster 2 (stations)

	Canonical R	Canonical R. Sqr.	Chi. Sqr.	Df	P	Lambda Prime
0	0.9999	0.9999	123.14	16	0.0000	0.000
1	0.9995	0.9989	66.7981	9	0.000	0.000
2	0.9939	0.9878	36.02	4	0.0000	0.0003
3	0.9861	0.9725	16.16	1	0.0006	0.0275

Table 4.17: Summary of canonical input performance for cluster 2 (stations)

	Left side	Right Side
No. of Variables	4	4
Variance Extracted	100%	100%
Total Redundancy	99.628	99.979
Variables 1	S5	S3
Variable 2	S2	S11
Variable 3w	S10	S4
Variable 4	S12	S1

Table 4.18: Canonical Roots for cluster 2 (stations)

Set	Variables	Root 1	Root 2	Root 3	Root 4
Right side	S3	-0.4124	1.5185	3.4154	12.115
	S11	-0.1760	3.9828	2.8495	-1.6950
	S4	0.0148	-8.6129	-3.1481	-13.4353
	S1	-0.437	3.1845	-3.0887	3.0663
Left side	S5	-0.4745	-2.7931	3.2101	-1.3456
	S2	0.4638	-4.7518	-2.4010	-3.9670
	S10	-0.9645	7.1589	-0.6759	4.3507
	S12	-0.0461	0.4721	-0.4278	1.6567

Table 4.19: Eigenvalues for cluster 2 (stations)

	Root 1	Root 2	Root 3	Root 4
Variance Extracted	0.9999	0.9989	0.9879	0.9725

4.3.2 Analysis of parameter average values for Months

Cluster 1: Dec, Nov, Oct., Aug, Sept and July

Input

- Variables as indicated by the cluster 1 above while cases are the various parameters.
- Missing data are mean substituted.
- Variables are chosen as follows: First list M8, M9, M7 and Second list M12, M11, M10

Output

Table 4.20: Chi-Square test with successive roots removed for cluster 1 (months)

	Canonical R	Canonical R. Sqr.	Chi. Sqr.	Df	P	Lambda Prime
0	0.9998	0.9997	68.5199	9	0.0000	0.00004
1	0.9913	0.9827	22.3519	4	0.0017	0.0172
2	0.0812	0.0066	0.0364	1	0.8488	0.9934

Table 4.21: Summary of canonical input performance for cluster 1 (months)

	Left side	Right Side
No. of Variables	3	3
Variance Extracted	100%	100%
Total Redundancy	93.21%	47.963%
Variables 1	M12	M8
Variable 2	M11	M9
Variable 3w	M10	M7

Table 4.22: Canonical Roots for cluster 1 (months)

Set	Variables	Root 1	Root 2	Root 3	Root 4
Right side	M8	0.4762	3.9656	2.2141	12.115
	M9	-0.8457	1.3966	0.2496	-1.6950
	M7	-0.5457	-4.6818	-1.3729	-13.4353
Left side	M12	-2.8231	-7.1394	3.2101	-16.4879
	M11	2.4785	5.6354	-2.4010	-16.5936
	M10	-0.6735	1.4377	-0.6759	0.2774

Table 4.23: Eigenvalues for cluster 1 (months)

	Root 1	Root 2	Root 3
Variance Extracted	0.9998	0.9827	0.0065

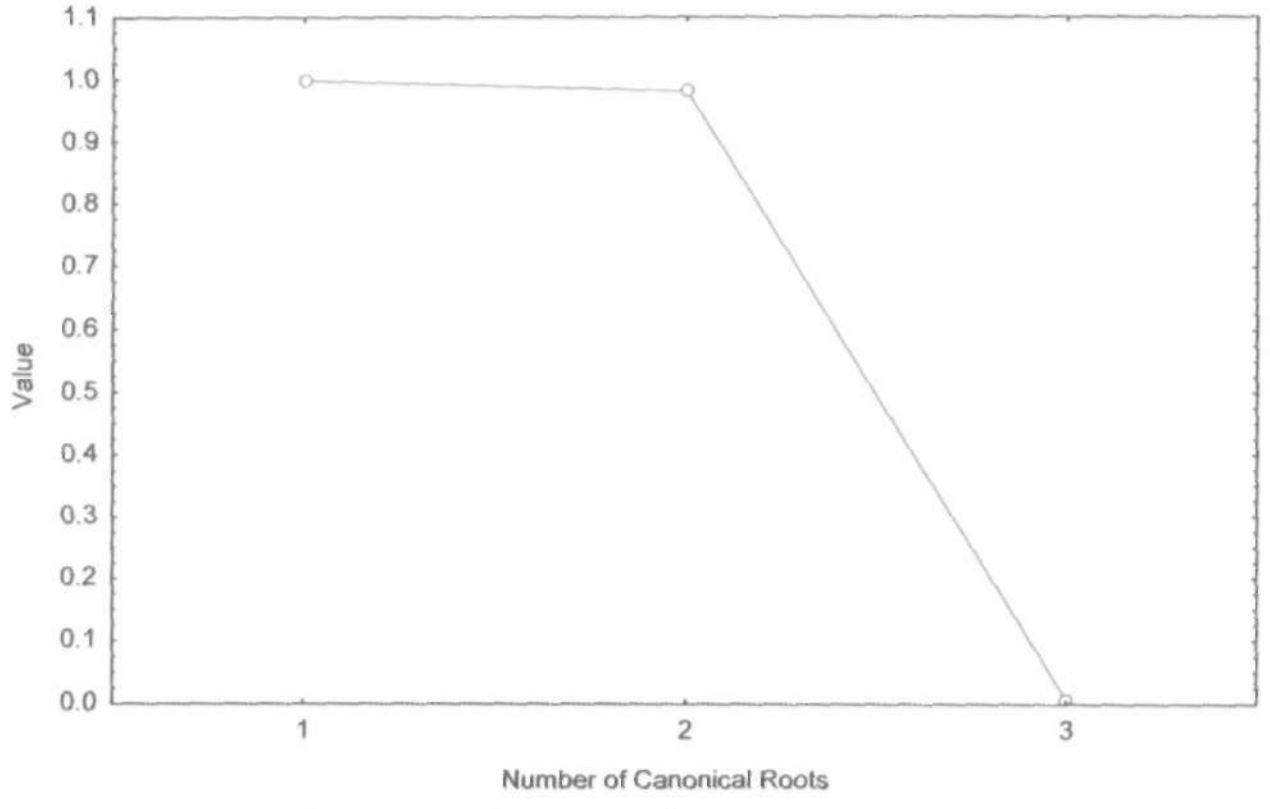


Figure 4.14: Plot of eigenvalues for cluster -1 for months

Cluster 2: June, May, April, March, Feb and Jan.

Input

- Variables as indicated by the cluster 1 above while cases are the various parameters.
- Missing data are mean substituted.
- Variables are chosen as follows: First list M6, M5, M4 and Second list M4, M2,M1

Output

Table 4.24: Chi-Square test with successive roots removed for cluster 2 (months)

	Canonical R	Canonical R. Sqr.	Chi. Sqr.	Df	P	Lamda Prime
0	0.9999	0.9998	65.8248	9	0.0000	0.0001
1	0.9825	0.9652	19.1138	4	0.0008	0.0310
2	0.3315	0.1099	0.6401	1	0.4237	0.8901

Table 4.25: Summary of canonical input performance for cluster 2 (months)

	Left side	Right Side
No. of Variables	3	3
Variance Extracted	100%	100%
Total Redundancy	97.7686	97.1600
Variables 1	M6	M3
Variable 2	M5	M2
Variable 3w	M4	M1

Table 4.26: Canonical Roots for cluster 2 (months)

Set	Variables	Root 1	Root 2	Root 3
Right side	M3	-1.2376	-6.2739	-4.0496
	M2	0.2564	7.1537	6.6998
	M1	-0.0161	-0.8668	-3.1311

Left side	M6	-0.0353	0.8757	1.8924
	M5	0.1539	3.5356	-6.9506
	M4	-1.1237	-4.1782	5.3380

Table 4.27: Eigenvalues for cluster 2 (months)

	Root 1	Root 2	Root 3
Variance Extracted	0.9998	0.9652	0.1099

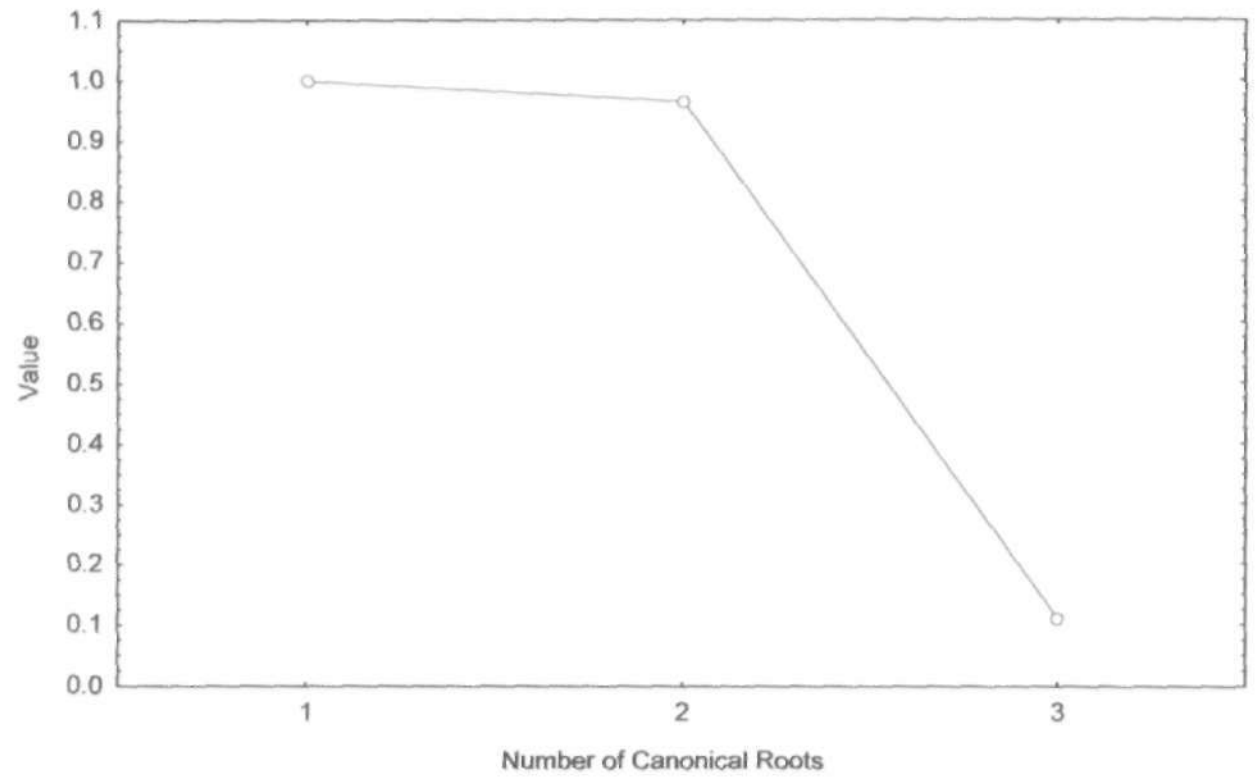


Figure 4.15: Plot of Eigenvalues for cluster 2 for months

The basic statistic values contained in Table 4.3 was obtained on the analysis of the entire data set. This was automatically pictorially represented by the box and whisker plot as shown in Figure 4.16 below. The top and down arrows indicate the maximum and minimum values for the parameter. The mid box shows the mean value, while the rectangle shows the height of the 25^{th} and 75^{th} percentile values.

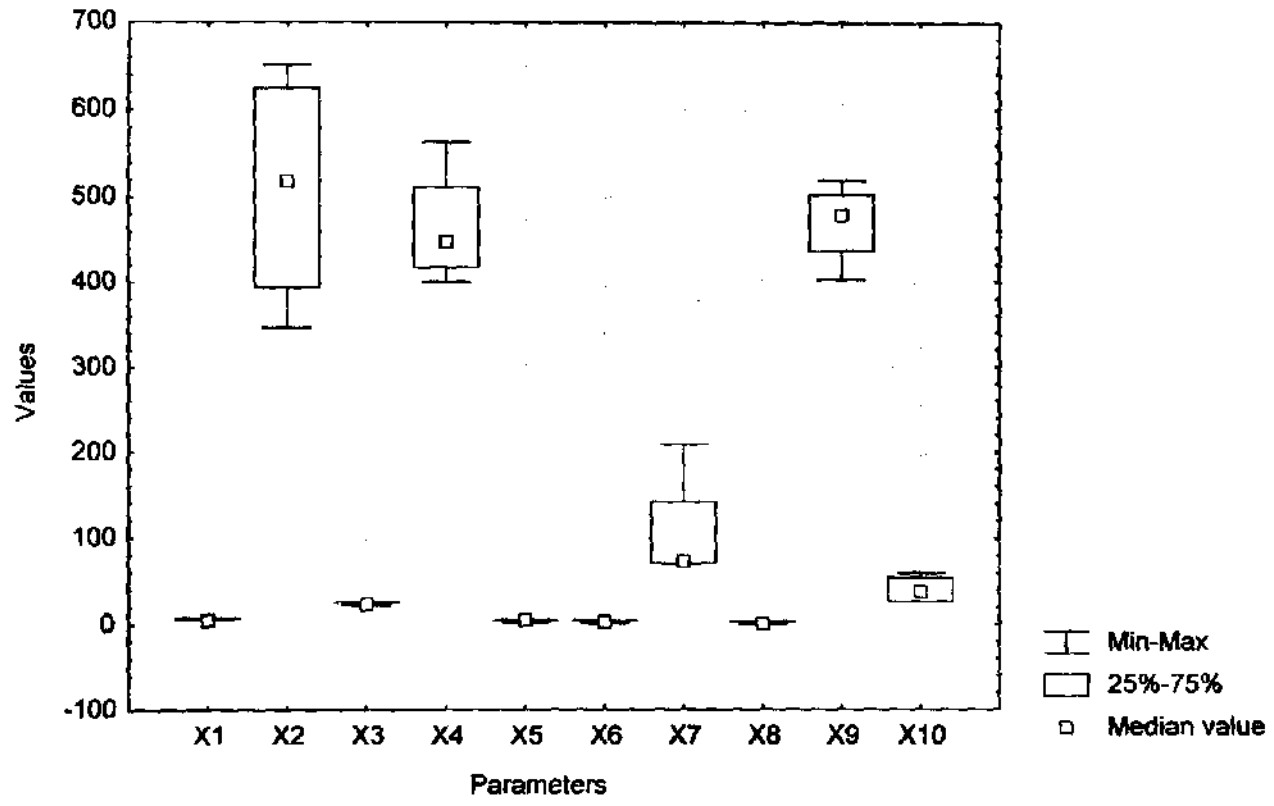


Fig. 4.16: Box and whisker plot of Average values for the parameters

Chapter 5 DISCUSSIONS

5.1 Summary values

Table 4.3 shows a comparison of the laboratory results obtained with the WHO and FEPA standards. The result shows that in terms of the average values only Dissolved Oxygen and Chemical Oxygen Demand were exceeded. However, in terms of maximum values obtained Nitrates, Phosphates, Sulphates, Chemical Oxygen Demand had values above the ones recommended. Figure 4.16 show the box and whisker plot of these values. This is a pictorial expression of the average values in Table 4.3.

5.2 Cluster Analysis

Cluster analysis was applied to the data to detect similarity of groups between sampling stations and months. Euclidean squared distance method of analysis was applied. From the results, two statistically significant cluster sets were formed for the sampling station and months respectively (Figure 4.11 and 4.12).

For the sampling stations, the two cluster groups obtained was interpreted as approximating to the stations located on each arm of the two major rivers studied ; River Kubanni (S5, S2, S10, S12, S3, S11, S4 and S1) and Shika (S7, S9, S8 and S6).

Similarly for months, the two cluster groups approximated to the two distinct season of the year; Wet (Dec, Nov, October, September, August and July) and Dry (June, May, April, March, Feb and Jan.) respectively. The apparent overlap is based on the non-isolation of the usual dry spell that exists within the wet months.

From the analysis above it can be deduced that Cluster analysis makes it possible to design a future spatial sampling strategy in an optimal way. It convincingly offers a reliable grouping of surface waters in the upper Galma basin into flow paths of similar characteristics. It is also possible to determine the sampling station that can be used for rapid water quality assessment studies by selecting a representative sampling station from each cluster bearing in mind the peculiar on-site conditions.

5.3 Factor Analysis and Principal Component Analysis

Factor and principal component analysis was performed on the mean values of the station (Table 4.1) and months (Table 4.2) respectively. This analysis was performed so as to compare the compositional patterns between the analyzed water samples and thus identify factors that influence each principal component of the data set.

The first analysis was performed on the data obtained based on sampling stations. Only one (1) eigenvalue was obtained representing a value of 91.1% of the total variance (Table 4.3). From the factor loading, all the stations are significant as they have a value ranging from 0.89 to 0.9 (Table 4.4). But in terms of factors extracted, total solids and chemical oxygen demand are highest. This is followed by dissolved Oxygen, Nitrates and Phosphates. The third most significant factors are pH and electrical conductivity.

In the second analysis based on months, two (2) eigenvalues are identified. The first eigenvalues accounts for 89.3% of the variance, while the second eigenvalues accounts for 8.45. The two

together account for 97.8% 0f the total variance (Table 4.6). Because of this, two factor loadings were calculated according to the significance imposed by the two eigenvalues. The results show that the loadings imposed by eigenvalues number one (1) is significant and that all months have high factor scores on the first principal components (Table 4.7), however results based on factor score shows that for the two eigenvalues, total solids and chemical oxygen demand still ranked highest, followed by conductivity ,then nitrates and phosphates in that order.

The result suggests that pollution in the upper Galma basin is mainly from erosive forces which accounts for high total solids and conductivity. The upper catchment being mainly agrarian and the use of chemical and organic fertilizer being rampant suggest the high values of Chemical Oxygen Demand, Nitrates and Phosphates. Therefore pollution can conclusively be said to arise from non-point sources, especially as most of the existing industries are non- functional.

5.4 Discriminant Analysis

Discriminant analysis was carried out based on the sample parameter Clustered groups for months and station. The analysis was used so as to maintain a standard step approach rather than a step-wise approach.

The interpretations of the data in Discriminant analysis were based on the following criteria:

- Wilks' Lambda is a reflectance of a variables importance. The smaller the value the more important the variable.
- Eigenvalues greater than one (1) are significant
- Percentage of variance greater than 5 is significant

- Cumulative percentage to approximately 75% is significant
- A correlation value of greater than 0.6 is significant
- The f and p values also provide a measure of importance of a variable

Discriminant analysis for parameter cluster values for sampling stations

(A) Cluster 1:

From Table 4.12, root 1 is identified as important. The stations exhibited similar redundancy with stations S7 and S9 having a value of 91.94% and S8 and S6 with value of 94.94%. Using root 1 as a pivot, we identify the most significant sampling stations in this cluster as S8 and S9 with a eigenvalue of 0.9582 (Table 4.15).

These two stations are therefore identified as being representative of the sampling stations in this cluster. They are therefore regarded as the most polluted sampling points.

(B) Cluster 2

From Table 4.16, roots 1 and 2 are most significant, followed by root 3. Root 1 identified S10 as significant, while root 2 identified S12 as significant. From the values obtained in the root analysis, stations S3, S11, S4 and S1 agree significantly, hence they have the highest redundancy. In the eigenvalue analysis, roots 1, 2 and 3 have close values and thus confirm our earlier judgment of their significance (Table 4.19).

This analysis indicates that station 12 and 1 were inappropriately classified. Station 10 was identified as constituting a major pollution point, hence its significance. The result supports the classification of stations S3, S4, S10 and S11 because they lie on the same river stretch.

Discriminant analysis for parameter cluster values for months

(A) Cluster 1:

From Table 4.20, root 1 is significant because it has the lowest lambda value. The redundancies for left and right sides of analysis are 93.21% and 47.95% respectively (Table 4.21). Root 1 in Table 4.22 shows that M11 and M12 stand out and hence are considered inappropriate in the cluster. (Table 4.23)

The interpretation of this analysis means that month M11 and M12 were wrongly classified as rainy months while in actual fact they are dry months.

(B) Cluster 2

From Table 4.24, roots 1 and 2 are significant. However, because root 1 has the lowest lambda prime value, it was used for the interpretation of the result. Table 4.25 shows the redundancies on the left side as 97.77% and on the right side as97.16%. Table 4.26 identifies using root 1 identifies M3 in the left side and M4 in the right side as significant.

This analysis shows that M3 and M4 were wrongly classified as dry months while in actual fact they are wet months.

Chapter 6 CONCLUSION AND RECOMMENDATION

6.1 CONCLUSIONS

The multivariate statistical analysis tools of Cluster analysis, factor analysis, principal component analysis and Discriminant analysis were applied to twelve water sampling point in the upper Galma basin. Samples were collected over a period of twenty four months and systematically prepared to fit into excel packages that would form the input data for the multivariate statistical analysis. All statistical calculations were done using STATISTICA package.

Cluster analysis was able to sort out the data into dry and wet season on month basis, as well as according to sampling stations along rivers Kubanni and Shika respectively.

The clusters obtained provided a basis for the factor and principal component analyses. The factor and principal component analyses were applied to clustered groups to find out the parameters of interest in the basin. Factor analysis identified Total Solids, Chemical Oxygen Demand, Electrical Conductivity, Nitrates and Phosphates as the significant parameters. This means that these parameters exert a noticeable influence on the water quality and consistently exceeded desired limits. Therefore, it was easy to conclude that most of the pollution was identified as from non-point sources of which leaching from agricultural land could be fingered.

Discriminant analysis was used as a confirmatory test for the deliberate overlap allowed in the cluster groups. This is because it was possible to isolate a third cluster which would have made a

distinctive seasonal demarcation. Discriminant analysis was able to isolate these non-conformal months and stations from the cluster. This has confirmed the success of the data analysis.

From the overall results obtained, sampling could be limited to the identified stations S8 and S9 which are Gaskiya Bridge and Tudun Wada Bridge for Kubanni River and station S10 which is located near the Zazzau ginnery on the River Shika. This could provide representative results for the rivers concerned.

6.2 RECOMMENDATIONS

Among the prerequisites for decision-making, the following are important:

- political willingness
- cultural attitudes and
- administrative capacity

Consequently there is need to create public awareness on the deleterious effects of pollution.

Based on the result obtained from the cluster, factor and principal component analysis and discriminant analysis, the following recommendations are made:

(1) That automatic sampler should be installed at Gaskiya Bridge, Tudun Wada Bridge and at the bridge before the Zazzau ginnery for constant water quality sample collection. Otherwise, aside from specific sampling studies, data collected at these stations could indicate the degree of river pollution

(2) That a data bank on water quality be urgently set up by either the government or a conglomerate of the industrial concerns in the area, so that these data will be available to researchers.

(3) That adequate extension education be given to farmers on how to protect their farms from leaching and soil erosive forces.

(4) Though this research points to non-point sources as sources of pollution, it still important that effluent discharges from the few industries still in operation be intensified to prevent any escalation in the untreated effluent discharges into the river.

(5) That farming activities be restricted from the river banks from where the impact of the leached chemical fertilizers and detached soil mass reaches the rivers quickly.

(6) That this research should be use as a basis for embarking of mass pollution studies in the various river basins in Nigeria to ascertain their state of pollution and identify representative stations for continuous sampling

(7) That multivariate statistical analysis technique should be incorporated into the syllabus of our undergraduate and postgraduate applied statistics courses. At the moment this course is not being taught. It is a compulsory course in universities in most developed countries.

REFERENCES

Abdi H. (2003): Multi-variate Analysis. In: Lewis-Beck M., Bryan A., Futing T. (Ed.); Encyclopaedia of Social Sciences Research Methods, Thousand Oaks, Dallas.

Ackay A., Dguz A. and Korapire C. (2003): Study of Heavy metal pollution and speciation in Buyak Menderes and Gediz River sediments. Water Research No. 317 pp 812 – 822.

Adakole J.A. (2000): The effects of Domestic, Agricultural and industrial effluents on the water quality of Bindare Stream, Zaria. PhD dissertation, Department of Biological Sciences, A.B.U, Zaria.

Adie D.B. (2001): Modelling of Water quality and Flow in a Semi-urban Watershed. M.Sc thesis, Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria.

Anderson T.W (2003): An Introduction to multivariate Statistical Analysis. Wiley Series in Probability and Statistics.

Bates C.B, Rohman A., Mein R.G., and Weinman E.P (1998): climatic factors that influence the homogeneity of regional floods in southern Australia. Water Resources Research Vol.24, no. 122, pp3363-3381.

Beecroft G.A (1985): Water Pollution. In Water Resources Master Plan of Kaduna State by Parkman Nig Ltd for FEPA, Nigeria

Beecroft G.A, Yammama A.G., Kooi P.J. and Van-Straten G. (1987): Kaduna River Pollution Monitoring, the 1986 exercise. Nigerian Journal of Engineering, Ahmadu Bello University, Zaria.

Bos E. and Bergkamp G. (2001): Water and the Environment. In 2020 focus 9; overcoming water scarcity and quality constraints. R.S. Meinzen-Dick and M. W. Rosegrant, eds. Washington D.C.: International Food Policy Research Institute.

Botkin D.B and Keller E.A (1998): Environmental science. John Wiley and Sons, Inc., U.S.A

Calamari D. and Chiaudiani A. (1984): Problems and Methods in defining water quality criteria of micro-pollutants in aquatic ecosystems. In proceedings of the third conference of prevention of pollution substances derived from wastes, Rome May 11 –12.

Carpenter S.R, Caraco N.F, Correll D.L., Howarth R.W, Shapley A.N and Smith V.H. (1998): Nonpoint pollution of surface water with phosphorous and nitrogen. Ecol. Appl No. 813 pp 559-568

Chapman D. (1992): Water Quality Assessment. In: Chapman D. on behalf of UNESCO, WHO and UNEP. London. Chapman and Hall: 585p.

Chinonyere O.B. (2003): Waste treatment in the petroleum industry (the KRPC experience). NNPC internal briefing.

Cornwin D.L and Wagenet R.J. (1996): application of GIS to the modeling of non-point source pollutants in the vadose zones: a conference overview. J. Environ. Qual.. vol.25, pp.403-411

Davis T.A, Volesky B. and Mucci A. (2003): A review of the Biochemistry of heavy metals biosorption by brown algae. Water Research, No.37, pp4311 - 4330

Dix H.M. (1981): Environmental pollution. John Wiley and Sons Incorporated, New York.

Fergusson J.F. (1990): The Heavy Elements: Chemistry, Environmental Impacts and Health Effects. Pergamon Publishers, Oxford, England.

Dixon W and Chiswell B (1996): Review of aquatic monitoring programme design. Water Res. 30: 1935 – 1948

Egbo A. (2003): Sustainable Water Resources Development and Environmental Sanitation. Water Front (Oct. – Dec.). NIWASA, Lagos

Ekanem E.J and Irekpita H (2004): Determination of Pollution status of Zaria Rivers by physico-chemical parameter monitoring. *ChemClass Journ*: 127 -137

Eigbeifo C. (1978): Hydrology of Kubanni drainage basin. Unpublished M.Sc thesis, Department of Geology, Ahmadu Bello University, Zaria.

FEPA (1991): Guidelines and standards for Environmental pollution in Nigeria, Federal Environmental Protection Agency, Lagos.

Greenacre M.J (1984): Theory and application of correspondence analysis. Academic Press, London.

Garba M.L and Schoeneick K. (2001): Actual Lead content of Kubanni dam, Samaru, Zaria. *Academy Journal of Science and Engineering.pp*118-123.

Helena B, Pardo R, Vega M, Barrado E, Fernandez J.M and Fernandez L (2003): Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by Principal Component Analysis. Water Res. 34: 807-816.

Hopke P.K (1985): Receptor modelling in environmental chemistry. Wiley Publishers, U.S.A

Iguisi E. and Madaki C.N (2003) : An assessment of the levels of toxicity of industrial effluents from some industrial establishments in Kaduna metropolis. Proceedings of 1st National Conference of Environmental Health Society of Nigeria, Ahmadu Bello University, Zaria.

Ismail A. (1996): Stormwater hydrograph characteristics evaluation for a semi-urban watershed. Unpublished M.Sc thesis, Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria

James A. (1993): An Introduction to water quality modeling. John Willey and sons . New York.

Johnson R.A and Wichern D.W. (1992): Applied Multivariate Statistical Analysis. Prentice and Hall International, England.

Jolainkai P. (1992): Hydrological, Chemical and biological Processes of Pollution Transport and transformation in Rivers and Lakes (a state of the art). UNEP Studies.

Kakulu S.E. and Osibanjo O. (1992): Pollution studies of Nigerian rivers: trace metal levels of surface waters in the Niger Delta area. Int. Jrn. of Env. Studies, 41: 287 – 92.

Keller A.Z and Wilson H. S (1992): Hazards to drinking water supply samples. Sprigerverlag Publishers, London.

KEPA (2000): Policy directives on industrial pollution control in Kaduna state, KDSG Press, Kaduna.

Liu C.W, Lin K.H. and Kuo Y.M. (2003): Application of factor analysis in assessment of ground water quality in a Blackfoot Disease area in Taiwan. Science Tot. Environment. No.313 pp77-89

Mara D (1986): Sewage treatment in hot climate. John Wiley and Sons, England

Mbajiorgu C.C. (2003): A water quality study of Ulasi River at selected locations. Proceedings of the 29th WEDC International conference. Abuja,

Murdoch W.W (ed.) (1975): Environment. Sinauer Associates Inc., Sunderland, England

Nunes M.L, Da Silva E.F. and Almeida S.P (2003): Assessment of water quality in Caiman and Man River (Portugal) using geochemical and biological indices. Water, Air and Soil Pollution, vol.149, no.3, pp. 227 – 250.

Nwaedozie J.M (2000): Environmental Pollution: A case study of wastewater effluent parameters of some industries in Kaduna. African journal of Environmental studies, Vol. 1, No. 1&2 pp84 – 89.

O'Connor D.J. (1967): The temporal and spatial distribution of dissolved Oxygen in streams. Water Resources Research Vol.3 No. 1.

Oguejiofor C. (1994): A pollution study and self-purification of River Galma, Zaria. M.Sc thesis Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria.

Ogunjobi, M. (2000): quantification of aquifer parameters in the Galma basin of the crystalline basement complex of northern Nigeria. Ph.D Dissertation, Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria.

Oke I.A (2000): Tertiary treatment of textile wastewater by means of carbonaceous medium. M.Sc thesis, Department of Water Resources and Environmental Engineering, Ahmadu Bello University, Zaria

Okuofu, C.A (1990): Waste effluent reuse as a means of water conservation. 3rd National Workshop/Exhibition on Water Resources and Environmental Pollution, Durbar Hotel, Kaduna

Okafor C.N, Umoh V.J and Galadima M. (2002): Public health Implications of using unsafe waste water for irrigation of vegetable crops in parts of Kaduna State. Conference proceeding of Environmental Health Society of Nigeria. Zaria. Pp51-58

Oluwande P.A. (1988): Water pollution: causes and effects. 2nd Annual International Workshop on Env. Sanitation and Water Resources, Kaduna.

Osibanjo O. (1988): Hydrogeological and chemical basic data acquisition from NNPC operational areas. Technical report to NNPC Lagos

Otun J.A., Adie D.B. and Oke I.A. (2006): Status of conventional wastewater treatment plants in northern Nigeria. Conference proceeding of SOSEH, Awka, Nigeria

Overton D.E and Meadows M.E (1976): Storm water Modelling. Academy Press, New York.

Reghunat R, Murthy T.R.J and Raghavan B.R (2002): The utility of multivariate statistical analysis technique in hydrogeochemical studies, an example from Karmataka, India. Water Res.No.36 pp 2437-2442

Rosegrant M.W., Cai X and Cline S.A. (2002): World water and food 2005 (dealing with scarcity). International Food Policy Research Institute, Washington D.C, U.S.A.

Rudd T., Sterit R.M and Lester J.N (1984): Formation and conditional stability constants of complexes formed between heavy metals and bacterial extracellular polymers. Water Res.: 18 (3): 379-384.

Salana R., Hatton T. and Dawes W. (1999): Predicting land use inputs in regional scale groundwater recharge and discharge. J. Environ. Qual. No.28, pp 446 – 460.

Shuxia Y.U, Jincheng S., Jinsong Z. and Huaicheng G.U.O (2003): Factor Analysis and Dynamics of water quality of Songhua River, North East China. Soil, Air and Water Pollution, Vol. 144 No1 pp 159-169

Simeonov V., Statis J.A., Samara C., Zachariadas G, Voutsa D., Anthemidis A., Sofomou M., T. Kouimtzis (2003): Assessment of the surface water quality in Northern Greece. Water Research No.37, pp.4119 – 4124.

Singh K.P., Malik A., Mohan D and Sinha S. (2004): Multivariate Statistical technique for evaluation of spatial and temporal variations in water quality of Gomti River (India). Water Research No.38, pp.3980 – 3992.

Streeter H and Phelps E (1925): A study of the purification of Ohio River. U.S Public Health Service Bulletin No. 146. Washington D.C.

Stroud K.A (1995): Engineering mathematics. Macmillan press limited, London.

Tebbut T.H.Y. (1998): The principles of water quality and control. Butterworth Publishers England

United Nations Environment Programme (UNEP) (1992): Earth Audit; The World Environment, 1972 to 1992, where now. UNEP GEMS survey, Nairobi, Kenya.

Vega M.P, Barrado E. and Dehan L. (1998): Assessment of seasonal and polluting effects of water quality of river water by exploratory method analysis. Water Research. No. 32, pp.3581 – 3592.

Vesiling, P.A. (1974); Treatment and Disposal of Wastewater Sludges; Ann Arbor Michigan.

Vesilend R.A., Pierce J.F. and Weiner R.F. (1994): Environmental Engineering. Butterworth Heinemann, Oxford – London.

Wang N., Zhu Y. and Li S. (1999): Analysis of dynamics variations and forming reasons of nourishment material in Songhai Lake. Research Environ. Sci. No.12, pp 26 – 29.

Wang J., Huang C.P. and Allen H.E. (2003): Modelling of heavy metals uptake by sludge particulates in the presence of dissolved organic matter. Water Research No.37, pp 4835 – 4842.

Wengliang D. (2001): A study of several specific secure two party computation problems. Ph.D thesis, Purdue University, West Lafayette, Indiana, U.S.A.

Wood A. (1995): Constructed wetland in water pollution control fundamental to their understanding. Water Science and Technology. Vol. 32, No2. pp. 21-29

Wulder M. (2005): Practical guide to Multivariate Statistics. Pacific Forestry Center, National Resources, Canada.

Wunderlin D.A, Diaz M.P, Ame M.V, Pesce S.F, Hued A.C. and Bistoni M.A. (2001): Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality: A case study of Tuquia river basin (Cordoba – Argentina). Water Research. No.35, pp. 2881 –2894.

APPENDIX

Appendix A: Average Laboratory Results for pH

STN	MONTHS													
	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec	**MN**	**SD**
1	6.85	7.02	6.93	6.63	7.35	7.25	7.3	7.1	7.19	6.92	6.36	6.58	**6.96**	**0.31**
2	6.54	6.93	6.33	6.43	7.42	6.72	7.19	7.06	6.98	7.06	6.42	6.38	**6.97**	**0.37**
3	7.03	7.25	6.82	6.6	7.48	6.99	7.13	7.01	7.05	7.31	6.84	6.36	**6.99**	**0.31**
4	6.66	7.22	6.93	6.65	7.29	6.94	7.55	6.95	6.82	6.66	6.92	6.80	**6.95**	**0.28**
5	6.99	7.25	7.12	6.72	7.32	7.2	6.97	6.90	6.84	6.70	7.14	6.85	**7.00**	**0.20**
6	6.92	7.24	6.70	6.83	7.16	7.37	7.65	7.25	7.39	7.11	7.69	7.12	**7.23**	**0.27**
7	6.84	7.23	6.66	6.74	7.75	6.89	7.23	7.12	7.17	6.85	7.16	7.14	**7.07**	**0.29**
8	6.83	7.02	6.62	6.67	7.21	6.96	7.4	7.05	7.06	6.97	7.04	7.22	**7.00**	**0.22**
9	6.8	6.89	6.68	6.74	7.79	6.86	7.5	7.27	7.53	7.04	7.13	6.71	**7.08**	**0.37**
10	6.57	6.89	6.7	6.51	7.09	7.03	7.93	6.97	6.97	6.94	6.59	6.33	**6.88**	**0.41**
11	6.32	6.86	6.9	6.96	7.27	7.2	7.52	7.06	6.87	6.73	7.19	6.67	**6.96**	**0.32**
12	6.58	6.87	6.63	6.71	7.04	7.04	7.38	6.9	6.88	6.74	6.89	6.69	**6.93**	**0.35**
MN	**6.74**	**7.06**	**6.78**	**6.68**	**7.41**	**7.04**	**7.40**	**7.05**	**7.07**	**6.92**	**6.95**	**6.74**	6.99	
SD	**0.21**	**0.17**	**0.21**	**0.14**	**0.25**	**0.19**	**0.26**	**0.12**	**0.22**	**0.19**	**0.37**	**0.31**		**0.24**

Appendix A1: 2005 Laboratory Results for pH

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	7.1	7	6.9	6.6	7.32	7.28	7.3	7.1	7.17	6.6	5.38	6.46	6.8508	0.54761
S2	6.7	7.06	6.4	6.44	7.44	6.82	7.1	7.22	6.85	6.98	5.52	6.56	6.7575	0.50026
S3	7.26	7.35	6.74	6.8	7.5	6.9	7.06	7.23	7.06	7.85	6.33	6.34	7.035	0.44824
S4	6.76	7.24	7.08	6.7	7.26	7.02	6.68	7.14	6.54	6.71	6.58	7.01	6.8933	0.25889
S5	7.16	7.42	7.34	6.8	7.28	7.03	6.93	6.9	6.78	6.72	7	7.04	7.0333	0.22701
S6	6.94	7.22	6.72	6.9	7.22	7.38	7.66	7.28	7.39	7.3	8.18	7.07	7.2717	0.3815
S7	6.74	7.35	6.74	6.77	7.9	7.01	7.26	7.14	7.15	6.48	7.67	7.21	7.1183	0.4074
S8	6.84	7.01	6.63	6.72	7.27	6.91	7.62	7.2	7.1	6.51	6.73	7.18	6.9767	0.31555
S9	6.82	6.7	6.66	6.74	7.81	6.8	7.61	7.56	7.65	7.79	6.81	6.69	7.1367	0.49009
S10	6.6	6.69	6.81	6.62	7.1	6.91	8.2	7.01	7.01	7.08	6.05	6.23	6.8592	0.53587
S11	6.54	6.88	7.01	7.03	7.3	6.76	7.41	6.8	6.84	6.55	7.29	6.65	6.9217	0.29291
S12	6.74	6.92	6.42	6.9	7.43	6.9	7.5	7.01	6.86	6.29	6.97	6.69	6.8858	0.34768
MN	6.85	7.07	6.7875	6.752	7.4025	6.977	7.361	7.133	7.0333	6.905	6.7092	6.7608		
SD	0.2234	0.2485	0.2678	0.156	0.2375	0.187	0.401	0.196	0.2945	0.5112	0.8181	0.3337		

Appendix A.2: 2006 Laboratory Results for pH

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	6.6	7.02	6.96	6.63	7.38	7.22	7.3	7.2	7.2	7.24	7.34	6.7	7.0658	0.2817
S2	6.38	6.8	6.26	6.42	7.4	6.62	7.28	6.9	7.11	7.14	7.37	6.4	6.84	0.41807
S3	6.8	7.15	6.9	6.4	7.46	7.08	7.2	6.79	7.04	6.77	7.35	6.38	6.9433	0.33811
S4	6.56	7.2	6.78	6.6	7.32	6.86	7.2	6.76	7.1	6.61	7.26	6.59	6.9033	0.29345
S5	6.82	7.08	6.63	6.63	7.36	7.01	7.01	6.9	6.9	6.68	7.28	6.66	6.9133	0.24655
S6	6.9	7.26	6.76	6.76	7.1	7.32	7.58	7.22	7.19	6.92	7.2	7.21	7.1183	0.24282
S7	6.94	7.13	6.71	6.71	7.6	6.77	7.2	7.1	7.13	7.22	6.65	7.07	7.0192	0.27632
S8	7.82	7.03	6.62	6.62	7.18	7.01	7.18	6.9	7.02	7.43	7.35	7.26	7.1183	0.33634
S9	6.78	7.08	6.74	6.74	7.77	6.9	7.39	6.98	7.41	7.27	7.45	6.9	7.1175	0.33483
S10	6.54	7.09	6.4	6.4	7.08	7.15	7.66	7.05	6.93	6.8	7.13	6.43	6.8883	0.38657
S11	6.1	6.84	6.89	6.89	7.24	7.32	7.4	7.01	6.9	6.91	7.09	6.69	6.94	0.33831
S12	6.52	6.82	6.9	6.9	6.51	7.18	7.46	7.2	6.9	7.19	6.81	6.67	6.9217	0.28989
MN	6.73	7.0417	6.7125	6.642	7.2833	7.037	7.322	7.001	7.0692	7.015	7.19	6.7467		
SD	0.418	0.1497	0.2107	0.172	0.3139	0.219	0.185	0.158	0.1545	0.2661	0.2405	0.3022		

Appendix B: Laboratory Results for Electrical Conductivity (µs/cm^2)

STN	MONTHS													
	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec	**MN**	**SD**
1	50	0	65	45	40	50	65	100	105	110	10	42	**61.8**	**29.8**
2	72	80	55	50	50	80	120	170	200	200	20	55	**96**	**61.9**
3	177	105	170	80	150	130	160	220	225	300	170	135	**171**	**62**
4	142	100	90	60	165	70	70	70	160	320	135	100	**123.5**	**72**
5	220	200	105	120	140	145	155	165	185	205	255	175	**172.5**	**43.2**
6	410	430	490	460	550	900	700	700	470	1100	1050	560	**651.7**	**242.7**
7	110	100	75	50	290	300	370	400	390	480	435	410	**284**	**157.3**
8	402	370	490	400	600	590	610	690	750	1000	725	540	**597.3**	**179.4**
9	173	190	210	220	200	365	800	900	1000	700	450	100	**442.3**	**321.4**
10	43	55	70	65	60	95	70	50	70	70	47	60	**62.9**	**13.9**
11	20	75	40	50	60	145	175	130	105	95	50	60	**83.8**	**47**
12	165	200	245	290	130	180	300	145	105	100	460	150	**205.8**	**103.7**
MN	**165.3**	**163.8**	**175.4**	**157.5**	**327.9**	**254.2**	**299.6**	**311.7**	**316.3**	**390**	**317.3**	**199**	**255.3**	
SD	**128.2**	**122.8**	**160.2**	**149.1**	**150.6**	**255.9**	**263**	**290.8**	**292**	**358.7**	**322.1**	**191.1**		**206**

Appendix B.1: 2005 Laboratory Results for Electrical Conductivity (µS/cm^2)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	51	30	67	46	38	41	65	40	50	30	13	44	42.917	14.9512
S2	74	77	52	46	48	76	135	168	202	199	27	52	96.333	62.6902
S3	185	107	180	81	150	140	160	240	250	310	179	130	176	64.5009
S4	144	100	78	65	170	85	75	162	162	370	209	148	147.33	83.8194
S5	240	200	110	130	165	165	170	200	220	310	449	270	219.08	91.9392
S6	380	400	410	420	510	510	530	550	640	780	800	400	527.5	145.047
S7	100	80	80	50	340	350	400	420	380	760	370	390	310	203.961
S8	444	370	410	400	600	570	600	700	800	700	299	110	500.25	196.685
S9	146	180	200	220	200	500	700	800	500	500	740	110	399.67	253.164
S10	46	60	70	70	60	100	70	50	65	99	54	55	66.583	17.2756
S11	20	100	40	55	65	95	160	110	60	139	79	59	81.833	40.7405
S12	190	200	250	300	60	60	300	150	110	100	519	200	203.25	129.346
MN	168.33	158.67	162.25	156.9	200.5	224.3	280.4	299.2	286.58	358.08	311.5	164		
SD	132.14	118.89	132.62	142	187.59	199.5	223.7	256.8	244.73	269.01	270.73	126.63		

Appendix B.2: 2006 Laboratory Results for Electrical Conductivity ($\mu S/cm^2$)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	49	33	63	44	42	59	65	60	55	800	27	40	111.42	217.18
S2	70	83	58	54	52	84	105	172	198	201	13	58	95.667	61.61
S3	169	103	150	79	150	120	160	200	200	290	16	140	148.08	68.24
S4	140	100	92	55	160	55	65	88	158	270	60.9	52	107.99	64.68
S5	200	200	100	110	115	125	140	130	150	100	60.8	80	125.9	42.47
S6	40	460	570	500	590	1290	870	850	300	1320	1300	720	734.17	410.38
S7	120	120	70	240	240	250	340	380	300	200	500	430	265.83	130.83
S8	360	370	570	600	600	610	1000	1100	1200	700	601	90	650.08	319.30
S9	200	200	220	200	200	230	900	1200	900	900	160	90	450	397.15
S10	40	50	70	60	60	90	70	50	65	41	40	65	58.417	15.02
S11	20	50	40	45	55	195	280	250	150	51	21	61	101.5	92.49
S12	40	200	240	280	200	300	300	140	100	250	401	100	212.58	103.66
MN	120.67	164.08	186.92	188.9	205.33	284	357.9	385	314.67	426.92	266.73	160.5		
SD	95.715	126.95	181.71	180.2	185.33	337.1	339.5	400.1	342.57	388.32	370.19	196.21		

Appendix C: Field data for Temperature (°C)

	MONTHS													
STN	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec	**MN**	**SD**
1	24	20	27	20	17	19	19	22	25	26	27	24	**22.6**	**3.5**
2	27	26	26	20	17	19	16	23	25	24	25	24	**22.7**	**3.6**
3	25	25	25	20	16	19	20	23	25	26	29	25	**23.2**	**3.7**
4	25	25	25	20	18	19	20	22	25	25	27	25	**23**	**3.0**
5	26	27	26	20	17	19	20	21	26	25	27	25	**23.2**	**3.7**
6	29	30	26	20	17	20	20	22	26	26	29	27	**24.2**	**4.4**
7	24	26	25	22	19	20	20	26	28	28	32	28	**25**	**3.6**
8	26	27	26	22	20	22	20	25	28	30	33	28	**25.8**	**3.6**
9	25	26	27	22	19	21	22	23	26	28	29	29	**24.5**	**3.3**
10	28	29	27	23	19	19	20	23	28	27	27	25	**24.4**	**3.7**
11	28	29	27	25	20	22	24	26	28	27	29	28	**25.9**	**2.8**
12	27	28	26	22	20	23	25	25	28	28	31	25	**25.5**	**3.0**
MN	**26.2**	**26.3**	**26**	**21.5**	**18.2**	**20**	**20.7**	**23**	**26.5**	**26.5**	**28.8**	**25.8**	**24.1**	
SD	**1.6**	**2.56**	**0.7**	**1.5**	**1.4**	**1.5**	**2.2**	**1.6**	**1.6**	**1.7**	**2.1**	**1.7**		**3.3**

Appendix C.1: 2005 Field Results for Temperature (°C)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	23.5	19	24.5	20.5	16	19	19	23	26	20	34	24	22.375	4.68132
S2	27	25.5	25	20	16	19	15	23	26	16	30	24	22.208	4.90574
S3	25.5	25	25	20	15	18	21	24	25	24	30	25	23.125	3.98933
S4	26	25	24	18	18	19	20	21	26	20	25	24	22.167	3.12856
S5	27	27	26	19	18	19	21	21	26	19	26	25	22.833	3.61395
S6	31	31	28	20	18	20	20	21	27	21	35	29	25.083	5.69622
S7	24	26	26	22	21	21	20	26	27	37	39	30	26.583	6.09707
S8	27	27	26	22	21	21	21	25	28	30	37	29	26.167	4.70654
S9	25	28	27	22	19	21	23	23	26	29	37	29	25.75	4.78872
S10	26	27	27	23	18	19	20	24	27	25	36	24	24.667	4.75458
S11	29	29	27	26	20	22	24	25	30	25	37	30	27	4.45176
S12	27	28	26	22	20	22	23	24	28	26	34	25	25.417	3.70401
MN	26.5	26.458	25.958	21.21	18.333	20	20.58	23.33	26.833	24.333	33.333	26.5		
SD	2.056	2.919	1.1766	2.105	1.9695	1.348	2.314	1.67	1.3371	5.7577	4.5594	2.6112		

Appendix C.2: 2006 Field Results for Temperature (^{o}C)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	24.5	21	29.5	19.5	18	19	19	21	24	32	20	24	22.625	4.39072
S2	27	26.5	27	20	19	19	17	23	24	32	20	24	23.208	4.39761
S3	24.5	25	25	20	17	20	19	22	25	28	28	25	23.208	3.5513
S4	24	25	26	22	18	19	20	23	24	30	29	26	23.833	3.7132
S5	25	27	26	21	16	19	19	21	26	31	23	25	23.25	4.20227
S6	27	29	24	20	16	20	20	23	25	30	20	25	23.25	4.20227
S7	24	26	24	22	17	19	20	26	27	29	19	26	23.25	3.79294
S8	25	27	26	22	19	19	19	25	28	30	19	26	23.75	3.98006
S9	25	24	27	22	19	21	21	23	26	28	19	29	23.667	3.3934
S10	28	28	27	23	19	20	20	22	28	29	18	26	24	4.0899
S11	27	29	27	24	20	22	24	26	26	29	21	26	25.083	2.93748
S12	27	28	26	23	20	21	24	24	28	30	22	25	24.833	3.06989
MN	25.667	26.292	26.208	21.54	18.167	19.83	20.17	23.25	25.917	29.833	21.5	25.583		
SD	1.4196	2.3008	1.4994	1.437	1.4035	1.03	2.038	1.712	1.5643	1.3371	3.5548	1.3114		

Appendix D: Laboratory results for Total Dissolved Solids (mg/L)

STN	MONTHS													
	Jan	Feb	Mar	April	May	June	July	Aug	Sep	Oct	Nov	Dec	**MN**	**SD**
1	380	156	128	66	114	176	148	176	168	153	1050	796	**292.6**	**308.5**
2	402	198	82	86	132	145	200	218	398	469	928	140	**333**	**269**
3	226	198	177	90	126	179	186	274	230	295	958	480	**285**	**233.8**
4	240	126	126	90	134	176	145	228	214	391	606	394	**239**	**152**
5	142	320	126	114	170	245	564	366	170	175	488	378	**271.5**	**150.3**
6	336	204	304	266	392	593	690	630	392	1030	1004	918	**563.3**	**294.5**
7	163	225	225	188	212	262	354	340	266	666	578	742	**351.8**	**198.1**
8	444	308	244	266	304	388	363	404	438	655	822	906	**461.8**	**217**
9	164	226	226	190	234	322	438	570	910	644	732	548	**433.7**	**245.4**
10	138	92	232	108	242	303	521	366	152	178	996	668	**333**	**271.5**
11	102	102	124	94	118	166	174	188	134	148	640	526	**209.7**	**178.6**
12	194	204	228	186	154	163	178	190	142	84	614	604	**245.1**	**173.8**
MN	**244.3**	**196.6**	**185.2**	**145.3**	**194.3**	**260**	**330**	**329.2**	**301.2**	**407.3**	**784.7**	**641.7**	**335**	
SD	**116.6**	**71.3**	**67**	**71.1**	**86.3**	**130**	**187.5**	**149.3**	**220**	**291**	**197.9**	**183.7**		**104.3**

Appendix D.1: 2005 Laboratory Results for Total Dissolved Solids (mg/L)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	420	246	156	62	110	156	128	150	164	106	100	192	165.83	93.1293
S2	424	230	59	78	131	109	210	170	270	530	796	82	257.42	222.233
S3	252	176	200	80	102	148	175	290	196	290	616	260	232.08	138.542
S4	240	170	162	90	130	152	110	200	190	492	412	300	220.67	123.041
S5	204	480	100	110	180	218	528	400	140	190	326	356	269.33	144.93
S6	400	288	408	323	444	678	680	700	374	940	1000	830	588.75	247.866
S7	216	260	240	126	164	214	238	400	230	932	699	501	351.67	243.863
S8	387	316	248	216	268	476	386	408	400	580	1004	1110	483.25	286.958
S9	218	246	226	200	248	340	530	722	928	358	814	496	443.83	254.506
S10	155	94	204	100	270	252	270	322	104	116	862	804	296.08	262.715
S11	94	98	146	98	128	162	168	174	186	136	570	452	201	150.143
S12	190	200	220	172	158	168	170	190	144	138	588	498	236.33	146.337
MN	266.67	233.67	197.42	137.9	194.42	256.1	299.4	343.8	277.17	400.67	648.92	490.08		
SD	112.2	103.51	87.572	76.47	98.626	167.2	187.1	196.5	224.05	300.87	272.84	295.8		

Appendix D.2: 2006 Laboratory Results for Total Dissolved Solids (mg/L)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	340	66	100	70	118	196	168	202	172	200	2000	1400	419.33	616.126
S2	380	166	105	94	133	181	190	266	526	408	1060	198	308.92	271.346
S3	200	220	254	100	150	210	197	218	264	300	1300	700	342.75	336.535
S4	240	82	90	90	138	200	180	256	238	290	800	496	258.33	205.256
S5	80	160	150	108	158	272	600	332	207	160	650	400	273.08	188.629
S6	172	120	200	220	340	508	700	560	410	1120	1008	1006	530.33	354.863
S7	110	190	210	250	260	310	470	280	302	400	457	655	324.5	148.714
S8	501	300	240	316	340	300	340	400	476	730	640	702	440.42	168.936
S9	110	206	226	180	220	304	346	418	892	930	650	600	423.5	280.277
S10	121	90	260	116	214	354	772	400	200	240	1130	532	369.08	309.508
S11	110	106	102	90	108	170	180	174	190	160	750	600	228.33	213.977
S12	198	208	236	200	150	158	180	190	140	30	640	710	253.33	204.093
MN	213.5	159.5	181.08	152.8	194.08	263.6	360.3	308	334.75	414	923.75	666.58		
SD	130.64	69.34	66.826	78.58	81.609	100.6	222.6	116.6	214.9	336.52	420.21	302.15		

Appendix E: Laboratory results for Nitrates (mg/L)

STN	MONTHS													
	Jan	Feb	Mar	April	May	June	Jul	Aug	Sept	Oct	Nov	Dec	**MN**	**SD**
1	14.6	19.7	6.2	9.3	3.3	2.3	1.8	2.6	7.6	3.3	7.2	1	**6.6**	**5.7**
2	3.6	10.4	9.2	5.2	1.2	0.8	2.2	2.2	12	6	6.2	2.3	**5.1**	**3.8**
3	3.3	7.6	11.6	3.6	1.4	1.2	2.4	3.6	12.8	4.5	3.4	4.2	**5.0**	**3.8**
4	2.5	10.2	11.3	6.2	6.8	4.2	13.4	3.4	10.8	5.6	3	5.2	**6.9**	**3.7**
5	2.9	6.6	2.8	2.1	0.8	1	4.4	3.6	12.5	14.2	16.6	3	**5.9**	**5.4**
6	18.4	4.8	3.6	5.8	1.2	1.8	4	3.8	13.6	8.4	2.6	4	**6.0**	**5.1**
7	2.4	4.7	6.0	7	0.6	1.6	2.4	3	7.6	6.1	1.4	1.6	**4.0**	**2.2**
8	23.6	11	8.2	2.4	0.8	1.4	4.6	5	12	4.4	2.2	4.8	**6.7**	**6.4**
9	7.6	3	10.8	4.6	2.8	1	2.1	2.8	18	5.3	2.4	4	**5.4**	**4.8**
10	6.2	2.6	3	3.6	1.6	1.8	5.1	4.6	8.8	3.6	7.6	5	**4.5**	**2.2**
11	3	3.6	4.2	2.8	1	2	2.8	1.8	7.4	2.9	2.8	2.6	3.1	1.6
12	4.5	2.8	6.2	7.9	1.8	4	6.8	2.8	8	2.2	2	5	4.5	2.3
MN	**7.7**	**7.3**	**6.9**	**5.0**	**2.2**	**1.9**	**4.3**	**3.3**	**10.9**	**5.5**	**4.8**	**3.6**	**5.3**	
SD	**7.2**	**5.0**	**3.2**	**2.3**	**1.7**	**1.1**	**3.2**	**0.9**	**3.2**	**3.2**	**4.3**	**1.4**		**1.2**

Appendix E.1: 2005 Laboratory Results for Nitrates (mg/L)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	12.6	18.9	5	8.8	2.2	2.1	0.8	2	7	3	2.4	1.2	5.5	5.51444
S2	3	9.8	6	4.5	1	0.8	1.6	2.1	10	6	5.4	2.2	4.3667	3.17786
S3	3	5.2	7.2	7.7	1	1.2	2	2.3	10.6	5	4.2	4.6	4.5	2.89294
S4	2.5	12.4	6.6	5.8	9	4	17.4	3.7	5.5	5.5	3.3	4.8	6.7083	4.31771
S5	2	8	2.6	2	0.6	0.7	4.2	3	7.5	12.4	16.2	2.2	5.1167	4.94861
S6	1.7	5	3	6	1.4	1.6	3.8	4.4	7.2	12	3.5	4.4	4.5	2.94896
S7	2.8	3	5.6	7.2	0.4	1	1.2	3	7.4	5.3	0.9	1.1	3.2417	2.52999
S8	20.9	15.2	8.3	2	0.5	1	4	4.8	10	5.4	2.9	5	6.6667	6.10712
S9	8.2	3.2	13.5	5	3	1.2	2	3.2	20	5.8	3	0.5	5.7167	5.72186
S10	5.9	2.5	2.9	3.4	1.5	1.9	4.8	3.8	8.4	4.3	6.1	5.5	4.25	2.00295
S11	2.5	3.4	3.8	2.8	0.9	2.1	2.8	1.7	7.3	3.2	3.1	2.8	3.0333	1.55525
S12	4.3	2.4	5.9	8.4	1.7	4.4	6.5	2.9	8.7	3	2.4	5.8	4.7	2.36528
MN	5.7833	7.4167	5.8667	5.3	1.9333	1.833	4.258	3.075	9.1333	5.9083	4.45	3.3417		
SD	5.7206	5.5335	3.0046	2.418	2.3484	1.207	4.462	0.968	3.7293	3.1295	3.9463	1.8798		

Appendix E2: 2006 Laboratory Results for Nitrates (mg/L)

	MONTHS													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	16.6	20.5	7.4	9.8	4.6	2.5	2.8	3.2	8.2	3.6	12	0.8	7.6182	6.41324
S2	4.2	11	12.4	5.9	1.4	0.8	2.8	2.4	14	6	7	2.4	5.8583	4.45512
S3	3.6	10	16	10.2	1.8	1.6	2.8	4.9	15	4	2.6	3.8	6.3583	5.11423
S4	2.5	8	16	6.6	4.6	4.4	9.4	3.1	16.1	6	2.7	5.6	7.0833	4.67174
S5	2.8	5.2	3	2.2	1	1.3	4.6	4.2	17.5	16	17	3.8	6.55	6.33224
S6	19.8	4.6	4.2	5.6	1	2	4.2	3.2	20	1.8	1.7	3.6	5.975	6.64532
S7	20	6.4	6.4	6.8	0.8	1.2	3.6	3.6	7.8	6.9	1.7	2.1	5.6083	5.17397
S8	26	6.8	8.1	2.8	1.1	1.8	5.2	5.2	14	3.4	1.5	4.6	6.7083	7.03995
S9	7	3.8	8.5	4.2	2.6	0.8	2.2	2.4	16	4.8	1.8	0.3	4.5333	4.3389
S10	6.5	2.7	3.2	3.8	1.7	1.7	5.4	5.4	9.2	2.9	9.1	4.5	4.675	2.56129
S11	3.5	3.8	4.6	2.8	1.1	1.9	2.8	1.9	7.5	2.6	2.6	2.4	3.125	1.65975
S12	4.7	3.2	6.5	7.4	1.9	3.7	7.1	2.7	7.3	1.4	1.6	4.2	4.3083	2.27734
MN	9.7667	7.1667	8.025	5.675	1.9667	1.975	4.408	3.517	13.127	4.95	5.1083	3.175		
SD	8.3624	4.961	4.541	2.645	1.3289	1.096	2.126	1.165	4.4486	3.879	5.1018	1.5938		

Appendix F: Laboratory results for Phosphates (mg/L)

STN	MONTHS													
	Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec	**MN**	**SD**
1	2.4	1.1	4.5	1.0	3.8	1.0	2.1	2.7	2.5	1.8	2.1	1.8	**2.2**	**1.1**
2	3.4	0.9	3.1	1.6	4.6	2.7	3.1	1.6	6.3	4.4	1.8	1.6	**2.9**	**1.6**
3	2.6	1.0	1.8	9.8	5	1.6	1.6	3.2	3	3.3	2.1	1.8	**3.1**	**2.4**
4	1.6	1.0	1.6	1.6	3.8	1.9	1.9	2.6	3.1	4.6	2.2	1.7	**2.4**	**1.0**
5	1.1	1.0	1.5	1.8	3.7	2.2	2.2	2.5	2.10	1.90	1.8	1.7	**2.0**	**0.7**
6	1.6	1.2	1.8	3.7	8.7	8.7	8.7	6.9	5.3	10.6	6.5	2.4	**5.5**	**3.3**
7	1.6	1.0	1.7	1.5	3.9	2.1	2.1	2.4	2.1	6.9	2.2	2.1	**2.5**	**1.6**
8	1.7	1	1.9	2.8	7.3	4.5	4.5	5.6	9.7	11	5.4	2.1	**4.7**	**3.2**
9	1.7	1.2	2	1.9	3.7	4.2	4.2	4.9	18.6	5.4	3	2	**4.3**	**4.7**
10	1.5	1.3	1.6	2	4.7	2.8	2.8	2.1	2.5	3	2.2	1.8	**2.5**	**1.2**
11	2	1.4	3.1	2.2	6.3	1.8	1.8	1.1	0.9	0.7	1.2	1.8	**2.0**	**1.5**
12	1.4	1	1.5	1.8	3.6	1.5	1.5	1.1	2.7	1	2.1	1.9	**1.7**	**0.8**
MN	**1.9**	**1.1**	**2.2**	**2.6**	**4.9**	**3.0**	**3.0**	**3.1**	**4.9**	**4.6**	**2.7**	**1.9**	**3.0**	
SD	**0.6**	**0.2**	**0.9**	**2.4**	**1.7**	**2.1**	**2.0**	**1.8**	**4.9**	**3.4**	**1.6**	**0.2**		**1.2**

Appendix F.1: 2005 Laboratory Results for Phosphates (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	2.5	1.2	4.3	1.1	4.2	1	2.1	3.3	2.8	3	3.7	2	2.6	1.15758
S2	3.6	1	3.5	1.5	4.1	2.4	2.5	2.2	8.9	3.6	2.5	1.8	3.1333	2.04421
S3	2.9	1	1.2	11.4	4.9	1.4	1.3	2.9	5.2	4.1	3	1.9	3.4333	2.89021
S4	2.1	1.1	1.4	1.4	4.3	2.8	1.5	2.7	3.1	5.5	1.7	1.8	2.45	1.32699
S5	1	1	1.4	1.9	4.2	2.1	2.2	2.4	2.3	2.1	2	1.6	2.0167	0.83539
S6	1.8	1.2	1.8	4.1	13.4	13.6	12.2	9.6	6	15.1	8.8	2.6	7.5167	5.2263
S7	1.7	1.1	1.7	1.6	4.6	2	2.1	2.5	2.3	10.4	8.7	2.5	3.4333	3.00464
S8	1.5	0.9	1.8	3.3	10.8	4.1	4.3	4.5	13	15.8	5.8	3	5.7333	4.82613
S9	1.4	1.2	2.2	2.2	3.2	3.9	4.1	4.7	21.2	6	3.9	2.2	4.6833	5.39121
S10	1.7	1.2	1.4	1.6	5.6	2.7	2.8	2.2	2.3	4.1	2.3	1.8	2.475	1.25417
S11	2.1	1.5	2.9	1.2	6.7	1.7	1.5	1.1	0.9	1.4	1.2	1.7	1.9917	1.57391
S12	1.5	1.1	1.4	2.1	3	1.2	1.1	1	2.4	1.6	2.7	1.9	1.75	0.67082
MN	1.9833	1.125	2.0833	2.783	5.75	3.242	3.142	3.258	5.8667	6.0583	3.8583	2.0667		
SD	0.7209	0.1545	0.9796	2.853	3.1722	3.404	3.028	2.288	5.9358	5.0103	2.5794	0.425		

Appendix F.2: 2006 Laboratory Results for Phosphates (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	2.3	1.2	4.7	0.9	3.4	1	2.1	3.3	2.8	3	3.7	2	2.5333	1.16801
S2	3.2	0.9	2.7	1.7	5.1	3	3.7	1	3.7	3.2	1.1	1.4	2.5583	1.33107
S3	2.3	0.9	2	8.2	5.1	1.8	1.9	3.5	3.8	2.6	1.6	1.6	2.9417	2.02101
S4	1.1	1.2	1.8	1.8	3.2	2	2.3	2.5	3.1	3.7	2.7	1.6	2.25	0.8152
S5	1.3	1.1	1.6	1.6	3.2	2.3	2.2	2.6	1.9	1.7	1.6	1.8	1.9083	0.58381
S6	1.4	1.2	1.6	1.7	4	3.8	5.2	4.2	4.6	6.1	4.2	2.2	3.35	1.65557
S7	1.5	1	1.7	1.5	3.2	2.2	2.1	2.3	1.9	3.4	4.3	2.3	2.2833	0.93209
S8	1.9	1.1	2.1	2.3	3.8	5	4.8	6.7	6.4	6.2	5.1	1.2	3.8833	2.08407
S9	2	1.2	1.8	1.6	4.2	4.5	4.3	5.1	11	4.3	2.1	1.8	3.6583	2.69965
S10	14	1.4	1.8	2.4	3.8	2.9	2.8	2	2.8	2	2.1	1.8	3.3167	3.42632
S11	1.9	1.3	3.3	3.2	5.9	1.9	2.1	1.2	0.9	0.5	1.3	1.9	2.1167	1.45529
S12	1.3	0.9	1.6	1.9	4.2	1.9	1.9	1.2	3	0.4	1.5	1.4	1.7667	0.99484
MN	2.85	1.1167	2.225	2.4	4.0917	2.692	2.95	2.967	3.825	3.0917	2.6083	1.75		
SD	3.5595	0.1642	0.9324	1.912	0.8723	1.197	1.215	1.712	2.6629	1.8584	1.3675	0.3289		

Appendix G: Laboratory results for Sulphate (mg/L)

STN	MONTHS													
	JAN	FEB	MAR	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC	**MN**	**SD**
1	62	130	112	31	72	117	100	56	75	101	28	60	**78.7**	**33.3**
2	150	185	79	100	87	114	112	36	398	810	20	45	**178**	**222.4**
3	14	80	160	47	24	50	75	75	58	112	53	50	**66.5**	**39.2**
4	35	40	100	51	55	125	95	180	110	148	140	130	**100.8**	**47**
5	20	40	60	52	17	350	110	110	90	80	26	30	**82**	**90.7**
6	50	40	40	25	35	165	150	95	75	180	44	40	**78.3**	**55.9**
7	120	118	115	112	99	86	170	40	72	436	18	80	**122**	**106.5**
8	25	50	40	20	20	120	125	105	125	168	70	50	**76.5**	**50.1**
9	20	80	80	50	15	152	64	77	152	114	74	70	**79**	**43.3**
10	24	40	40	128	106	326	144	445	124	72	70	70	**132.4**	**126.6**
11	22	70	90	70	18	44	554	33	24	22	22	20	**82.4**	**150.5**
12	20	20	100	19	17	49	62	38	59	4	35	3.6	**35.6**	**28.1**
MN	**46.8**	**74.4**	**84.7**	**58.8**	**47.1**	**141.5**	**146.8**	**107.5**	**113.5**	**187.3**	**50**	**54.1**	**92.7**	
SD	**41.9**	**48.2**	**36.3**	**36.6**	**35.1**	**100**	**132.6**	**114.4**	**96.3**	**224.8**	**35.1**	**32.4**		**36.7**

Appendix G1: 2005 Laboratory Results for Sulphates (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	59	117	112	28	67	94	119	39	90	100	20	50	74.583	35.3591
S2	120	170	40	95	78	104	110	128	119	730	20	50	147	188.285
S3	16	76	116	54	34	47	68	71	48	90	50	45	59.583	26.562
S4	37	39	98	53	59	131	102	170	110	151	132	137	101.58	45.2397
S5	20	40	63	60	14	410	180	165	84	92	22	30	98.333	112.153
S6	58	40	38	20	39	205	170	120	81	240	41	50	91.833	74.343
S7	131	139	142	144	108	101	196	42	74	496	21	78	139.33	122.281
S8	34	53	51	20	20	162	169	141	100	166	70	40	85.5	59.0724
S9	20	76	84	50	15	205	61	104	120	176	84	78	89.417	56.4808
S10	30	42	47	70	111	440	187	36	81	72	86	74	106.33	113.175
S11	23	107	100	70	18	54	838	588	29	22	22	24	157.92	266.499
S12	20	21	115	18	18	40	65	65	52	8	33	34	40.75	29.9245
MN	47.333	76.667	83.833	56.83	48.417	166.1	188.8	139.1	82.333	195.25	50.083	57.5		
SD	39.232	46.819	35.079	36.35	35.745	133.1	210.5	149.3	28.474	211.63	35.872	30.888		

Appendix G.2: 2006 Laboratory Results for Sulphates (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	65	143	114	34	76	140	115	72.8	60	102	36	70	85.65	36.7989
S2	180	200	118	105	96	126	118	96	5.7	390	20	40	124.56	101.595
S3	13	84	200	40	25	53	83	79	70	124	56	55	73.5	49.5425
S4	33	41	102	49	51	119	68	189	111	145	148	124	98.333	49.7947
S5	20	40	57	44	20	290	40	65	77	68	30	30	65.083	73.2573
S6	42	40	42	30	31	125	130	70	72	120	47	30	64.917	38.7731
S7	109	97	88	80	90	97	145	38	70	410	15	72	109.25	100.27
S8	16	48	29	20	21	78	81	69	150	170	70	60	67.667	49.1775
S9	20	84	76	50	15	99	69	50	128	74	64	62	65.917	31.1578
S10	18	38	33	186	110	212	101	30	167	76	54	67	91	65.5716
S11	21	69	80	71	19	38	270	303	19	22	22	18	79.333	99.6743
S12	20	19	85	20	17	58	59	59	66	37	37	38	42.917	22.2321
MN	46.417	75.25	85.333	60.75	47.583	119.6	106.6	93.4	82.975	144.83	49.917	55.5		
SD	50.358	52.023	46.74	46.78	35.58	71.1	60.13	77.35	48.372	126.6	35.592	27.998		

Appendix H: Laboratory results for Dissolved Oxygen (mg/L)

STN	MONTHS													
	JAN	FEB	MAR	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC	**MN**	**SD**
1	5.0	6.8	4.8	8.0	8.0	6.4	6.0	4.8	4.4	2.8	6.8	4.4	**5.7**	**1.6**
2	5.0	6.0	6.0	6.0	7.2	6.4	4.0	4.0	4.4	5.4	4.0	4.4	**5.2**	**1.1**
3	7.0	6.4	6.0	5.6	7.2	3.6	2.0	3.2	2.8	4.0	4.0	4.4	**4.6**	**1.8**
4	5.4	6.8	6.8	7.2	7.6	6.0	6.4	5.2	2.8	3.8	4.0	4.8	**5.6**	**1.5**
5	4.0	6.4	0.2	4.4	5.6	6.0	2.4	2.6	2.6	3.0	3.2	3.6	**3.6**	**1.8**
6	1.2	5.6	4.8	2.8	4.4	4.4	4.2	4.2	3.8	4.0	3.0	2.0	**3.7**	**1.2**
7	4.0	4.6	4.8	5.6	8.0	3.6	6.4	4.4	5.2	4.0	6.0	4.4	**5.0**	**1.3**
8	1.8	2.8	2.0	1.8	1.8	1.8	1.6	1.6	1.6	1.2	1.0	0.8	**1.7**	**0.5**
9	4.0	5.2	4.8	4.8	4.8	4.8	4,2	4.4	4.4	4.6	4.8	5.6	**4.7**	**0.5**
10	3.2	6.0	6.0	6.0	7.2	4,4	4.8	5.6	4.0	5.4	4.0	3.6	**5.1**	**1.2**
11	4.6	5.6	6.0	6.8	8.4	6.8	6.4	5.2	3.6	2.8	5.6	5.6	**5.6**	**1.5**
12	3.4	6.8	4.0	5.2	7.2	5.6	5.2	3.2	3.2	3.6	6.0	4.0	**4.8**	**1.4**
MN	**4.1**	**5.8**	**5.2**	**5.4**	**6.5**	**5.0**	**4.5**	**4.0**	**3.6**	**3.6**	**4.4**	**3.9**	**4.6**	
SD	**1.6**	**1.2**	**1.3**	**1.8**	**1.9**	**1.5**	**1.7**	**1.2**	**1.0**	**1.3**	**1.6**	**1.4**		**0.9**

Appendix H.1: 2005 Laboratory Results for Dissolved Oxygen (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	4.8	7.2	4.5	10.8	10.4	6.7	6.3	5.1	4.6	3.2	8.2	5.3	6.425	2.37185
S2	5.2	4.8	4.9	4.6	8.6	7.7	4.8	4.6	5.1	5.1	4.4	4.6	5.3667	1.33576
S3	7.4	6.1	6.3	4.4	9.1	3.5	2.1	3	3.1	5.4	5.6	4.2	5.0167	2.02836
4	5.1	7.1	7.5	8.2	8.8	6.3	6.7	5.4	2.5	5.1	5.7	7	6.2833	1.68082
S5	3.8	6.7	5.4	4.2	5.9	6.3	2.3	2.5	2.5	3.1	3.8	3.8	4.1917	1.53946
S6	1.3	5.9	5.1	2.5	3.9	4.2	4.4	3.6	4	6.7	2.8	2.1	3.875	1.56329
S7	4.4	4.8	3.8	5.3	9.6	3.6	6.1	4.7	4.7	5.7	5.6	4	5.1917	1.59115
S8	1.9	2.9	2.1	1.9	1.7	1.7	1.8	1.7	1.7	1	0.9	1.8	1.7583	0.50355
S9	4.2	4.7	4.9	5.8	5.8	5	4.7	4.6	4.6	4.3	5.1	5.9	4.9667	0.58205
S10	4.4	6.9	5.4	5.7	5.8	5.3	5.8	5.9	4.8	6.8	5.5	4	5.525	0.854
S11	4.8	6.4	7.1	7.1	5.8	7.3	7.9	5.5	4.2	3.2	8.4	6.7	6.2	1.55095
S12	3.6	7.1	4.8	4.9	8.6	6	6.1	3.9	3.6	4.8	6.5	4.6	5.375	1.52323
MN	4.2417	5.8833	5.15	5.45	7	5.3	4.917	4.208	3.7833	4.5333	5.2083	4.5		
SD	1.5681	1.3265	1.4324	2.414	2.5993	1.774	1.963	1.294	1.0994	1.6735	2.0987	1.5909		

Appendix H.2: 2006 Laboratory Results for Dissolved Oxygen (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	5.2	6.4	5.1	5.2	5.6	6.1	5.7	4.6	4.2	2.4	5.4	3.5	4.95	1.13016
S2	4.8	7.2	7.1	7.4	5.8	5.1	3.2	3.4	3.7	5.7	4	4.2	5.1333	1.51077
S3	6.7	6.7	5.7	6.8	6.1	3.7	1.9	3.4	2.5	2.6	2.4	4.6	4.425	1.89599
S4	5.7	6.5	6.1	6.2	6.4	5.7	6.1	4.9	2.7	2.5	2.3	4.7	4.9833	1.59478
S5	4.2	6.1	5.6	4.8	5.3	5.7	2.5	2.7	2.5	2.9	2.6	3.4	4.025	1.41236
S6	1.1	5.3	4.6	3.1	4.9	4.6	4.4	4.8	3.6	1.7	1.2	1.9	3.4333	1.56921
S7	3.6	4.8	5.8	5.9	6.4	3.6	6.7	4.1	5.7	2.3	2.4	4.8	4.675	1.48882
S8	1.7	2.7	1.9	1.7	1.9	1.9	2	2.1	2.1	1	1.1	1.6	1.8083	0.45218
S9	3.8	5.7	4.7	3.8	3.8	4.6	3.7	5.3	4.2	4.9	4.5	5.3	4.525	0.68241
S10	4	5.1	6.6	6.3	8.6	3.5	3.8	4.9	3.2	4	2.5	3.2	4.6417	1.76092
S11	4.4	4.8	4.9	6.5	6.7	6.5	4.9	4.9	3	2.4	2.8	4.5	4.6917	1.43112
S12	3.2	6.5	3.2	5.5	5.8	5.2	4.4	2.5	2.8	2.4	2.3	3.2	3.9167	1.48681
MN	4.0333	5.65	5.1083	5.267	5.6083	4.683	4.108	3.967	3.35	2.9	2.7917	3.7417		
SD	1.5663	1.2169	1.4349	1.668	1.6329	1.318	1.567	1.103	1.0014	1.3301	1.2609	1.1572		

Appendix I: Laboratory Results for Chemical Oxygen Demand (mg/L)

STN	MONTHS													
	JAN	FEB	MAR	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC	**MN**	**SD**
1	172	248	260	408	360	304	380	580	416	680	296	196	**358.3**	**150.3**
2	192	272	380	344	320	240	400	720	328	840	156	168	**363.3**	**211.8**
3	256	260	408	448	748	296	392	776	340	300	368	184	**398**	**185.2**
4	240	280	440	428	428	180	368	604	416	704	244	88	**368.3**	**174.9**
5	296	260	480	560	340	292	536	380	390	4200	276	424	**396.2**	**96**
6	416	320	48	348	488	304	488	800	644	768	440	308	**485.3**	**172.7**
7	400	350	348	344	420	280	344	700	420	552	384	184	**393.8**	**130.1**
8	456	424	660	420	488	304	476	736	420	816	336	152	**470.7**	**186.1**
9	460	320	620	444	308	300	400	992	784	1000	280	320	**519**	**267.6**
10	208	428	612	572	360	340	400	624	336	656	376	376	**440.7**	**140.9**
11	252	416	528	356	400	244	400	624	452	620	260	424	**414.7**	**129.2**
12	268	408	380	392	324	196	296	760	480	660	42	156	**395**	**175.7**
MN	**301.3**	**332.2**	**430.3**	**422**	**412**	**273.3**	**406.7**	**691.3**	**419.7**	**668**	**319.7**	**248.3**	**416.9**	
SD	**104**	**70.6**	**170.9**	**77.4**	**119.9**	**48.2**	**65.3**	**148.6**	**188.5**	**187.3**	**82.3**	**115.9**		**51.4**

Appendix I.1: 2005 Laboratory Results for Chemical Oxygen Demand (mg/L)

Station	Months													
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	163	198	286	489	378	319	361	696	441	860	242	200	386.08	211.06
S2	253	313	456	415	384	254	380	1008	367	1120	112	202	438.67	308.64
S3	205	219	428	493	678	310	372	931	306	260	176	182	380	227.57
S4	252	224	352	538	342	162	386	726	426	818	258	106	382.5	217.22
S5	311	320	520	412	272	306	643	437	407	340	282	367	384.75	108.86
S6	478	322	510	277	510	274	512	1000	721	916	320	372	517.67	243.13
S7	480	368	365	361	357	294	386	805	601	584	208	162	414.25	179.09
S8	308	509	429	357	390	258	452	824	504	1112	352	170	472.08	258.74
S9	400	340	496	355	310	354	380	1141	1098	1490	290	352	583.83	411.06
S10	198	336	524	458	374	374	372	749	330	852	292	250	425.75	196.10
S11	239	500	556	409	480	209	480	700	474	650	230	318	437.08	161.45
S12	281	316	300	431	389	210	310	798	384	630	600	148	399.75	188.67
MN	297.33	330.42	435.17	416.3	405.33	277	419.5	817.9	504.92	802.67	280.17	235.75		
SD	104.87	97.304	91.39	72.45	107.36	62.2	89.7	184.4	219.99	343.39	119.88	93.36		

Appendix I.2: 2006 Laboratory Results for Chemical Oxygen Demand (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	181	298	234	327	342	289	399	464	391	500	350	192	330.58	99.57
S2	230	231	304	276	256	226	320	432	289	560	200	134	288.17	112.45
S3	307	301	388	403	890	281	412	621	374	340	560	186	421.92	188.72
S4	228	336	538	343	514	198	350	482	406	590	230	70	357.08	156.91
S5	281	200	440	708	408	278	429	323	373	500	270	276	373.83	137.22
S6	354	318	451	419	466	267	464	600	567	600	560	575	470.08	114.06
S7	320	333	331	335	483	266	302	595	239	520	560	206	374.17	130.69
S8	604	339	891	483	586	350	500	648	336	520	320	134	475.92	197.35
S9	520	300	744	533	306	246	420	843	470	510	270	288	454.17	190.90
S10	218	335	700	686	346	306	428	499	342	430	460	502	437.67	145.45
S11	265	332	501	303	320	279	320	548	429	590	290	530	392.25	119.48
S12	255	500	460	353	259	182	281	722	576	690	240	164	390.17	194.77
MN	313.58	318.58	498.5	430.8	431.33	264	385.4	564.8	399.33	529.17	359.17	271.42		
SD	126.69	72.049	193.46	144.6	178.21	46.08	69.27	138.8	100.83	88.98	138.07	170.63		

Appendix J: Laboratory results for Biochemical Oxygen Demand (mg/L)

STN	MONTHS													
	JAN	FEB	MAR	APRIL	MAY	JUNE	JULY	AUG	SEPT	OCT	NOV	DEC	**MN**	**SD**
1	9.6	4.8	9.6	9.6	14	9.6	19.2	9.6	9.6	4.8	9.6	14.4	**10.4**	**4.0**
2	14.4	9.6	14.4	4.8	4.8	9.6	19.2	4.8	14.4	9.6	14.4	9.6	**10.8**	**4.6**
3	16.8	14.4	14.4	14.4	4.8	9.6	96	14.4	9.2	14.4	19.2	19.2	**20.5**	**24.1**
4	17.4	9.6	9.6	9.6	4.8	9.6	19.2	9.6	14.4	33.6	9.6	14.4	**13.5**	**7.5**
5	14.4	9.6	14.4	24	4.8	14.4	38.4	8.8	8.9	9.2	9.6	9.6	**13.8**	**9.1**
6	28.8	29.6	38.4	38	34	43.2	249.6	62.4	38.4	76.8	81.6	14.4	**61.2**	**62.5**
7	20	22	22	19.2	4.8	14.4	96	14.4	14.4	57.6	14.4	9.6	**25.7**	**25.7**
8	29.6	24	28.4	48	24	48	153.6	43.2	43.2	81.6	76.8	24	**52.1**	**37.4**
9	28.8	28.8	24	24	4.8	28	19.2	38.4	14.4	84	28.8	9.6	**27.7**	**20**
10	17.2	14.4	19.2	14.4	4.8	14.4	19.2	9.6	9.6	9.6	9.6	9.6	**12.6**	**4.5**
11	17.2	21.6	14.4	14.4	4.8	24	38.4	4.8	14.4	14.4	9.6	14.4	**16**	**9.1**
12	28.8	21.6	14.4	19.2	4.8	14.2	96	9.6	76.8	9.6	14.4	14.4	**27**	**28.7**
MN	**20.3**	**17.5**	**18.6**	**20**	**9.6**	**20**	**72**	**19.1**	**22.3**	**33.8**	**25.2**	**13.6**	**24.3**	
SD	**7.0**	**8.2**	**8.5**	**12.4**	**9.8**	**13.4**	**71**	**18.5**	**20.6**	**31.9**	**25.8**	**4.5**		**16.5**

Appendix J.1: 2005 Laboratory Results for Biochemical Oxygen Demand (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	10.1	5.8	11.04	10.6	16.1	10.6	22.1	11.5	9.1	5.5	10.1	15.1	11.47	4.53743
S2	16.6	10.1	12	5.4	5.3	8.4	21.1	3.8	20.2	15.4	23	15.4	13.058	6.5654
S3	17.64	13	11.9	15.7	5.8	8.6	174.4	16.1	12.9	22.5	30.1	30.7	29.945	46.1334
S4	18	8.6	10.6	10.8	5.5	7.9	21.5	9.1	17.3	55.1	6.1	21.6	16.008	13.5799
S5	16.2	11.5	15.6	19.7	5.9	13.2	50.7	7	10.7	11.6	14	15.4	15.958	11.6028
S6	38.9	33.7	45.3	41	40.8	54	269.9	99.8	65.3	132.4	142.6	21.3	82.083	71.0307
S7	24	26.4	26.4	20.9	5.7	11.5	152.9	17.6	21.6	32.6	22.1	16.3	31.5	38.8892
S8	34	20.4	31.24	50.4	19.2	63.4	273.3	48.4	67.9	91.8	122.9	38.4	71.778	70.2612
S9	33.1	32	24.8	26.9	6.2	25.6	20.4	41.9	17.3	137.8	46.1	15.4	35.625	34.0383
S10	18.9	17.3	22.5	16.6	5.3	16.7	15.4	7.7	12	15.4	15.5	11.5	14.567	4.75917
S11	20.6	22.7	16.6	17.3	5	32.6	60.7	5.8	18.7	23.4	132.44	18.7	31.212	34.9255
S12	31.4	25.9	15.9	22.8	5.4	15.6	164.4	10.8	134.2	13.1	22.5	17.3	39.942	51.9516
MN	23.287	18.95	20.323	21.51	10.517	22.34	103.9	23.29	33.933	46.383	48.953	19.758		
SD	8.9255	9.3412	10.363	12.85	10.64	18.62	98.68	27.97	37.413	47.761	51.657	7.6218		

Appendix J.2: 2006 Laboratory Results for Biochemical Oxygen Demand (mg/L)

	Months													
STN	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	MN	SD
S1	7.9	3.8	8.2	8.6	11.9	8.6	16.3	7.7	10.1	4.1	8.6	13.7	9.125	3.56603
S2	12.2	9.1	16.8	4.3	3.3	8.4	17.3	5.8	8.6	3.8	5.8	3.8	8.2667	4.89143
S3	16	15.8	12.9	13.1	3.8	10.8	17.6	12.7	5.5	6.3	8.3	7.9	10.892	4.50746
S4	16.8	10.6	8.6	8.4	4	10.6	16.9	10.1	11.5	12.1	13.1	7.2	10.825	3.70555
S5	8.6	7.7	13.2	28.3	3.7	11.2	26.1	10.6	7.1	6.8	20.6	3.8	12.308	8.30438
S6	18.7	25.5	31.5	35	27.2	15.6	29.3	25	11.5	21.2	6.3	7.5	21.192	9.40246
S7	16	17.6	17.6	17.5	3.9	32.4	39.1	12.2	7.2	23.4	30.7	2.9	18.375	11.3738
S8	25.2	27.6	25.6	45.6	28.8	17.3	33.9	38	18.5	20.6	11.5	9.6	25.183	10.5919
S9	24.5	25.6	19.2	21.1	3.4	32.6	18	34.9	11.5	30.2	3.7	5.8	19.208	11.0778
S10	15.6	11.5	15.9	12.2	4.3	22.4	23	11.5	7.2	3.8	5.8	7.7	11.742	6.49174
S11	13.8	20.5	12.2	11.5	4.6	12.1	16.1	3.8	10.1	5.4	3.3	10.1	10.292	5.27472
S12	26.2	17.3	13	15.6	4.2	12.8	27.8	8.4	19.4	6.1	6.3	11.5	14.05	7.62561
MN	16.792	16.05	16.225	18.43	8.5917	16.23	23.45	15.06	10.683	11.983	10.333	7.625		
SD	6.0389	7.7238	6.7537	12.2	9.3623	8.549	7.771	11.28	4.3493	9.3157	7.997	3.2633		

APPENDIX K: FORTRAN PROGRAMME FOR PRINCIPAL COMPONENT ANALYSIS

The programme calculates the covariance matrix or the correlation matrix. from the set of data. The characteristic roots and vectors are then calculated from them. The principal component is extracted from the calculations of the eigenvalues and eigenvectors.

The number of stations or months and variables are entered and the required steps listed below are followed for the identification of principal components. The principal components are identified in a range of selected values. The components are not sorted in order of magnitude by the programme. The first and subsequent principal components are determined by visual inspection.

```
C       MEAN, COVARIANCE AND CHAR NEEDED
        DIMENSION X (10,100),XBAR (10), S(10,10), R(10,10)
   1    VEC(10,10), OOT(10),Y(100)
        READ (5, 5) M, N, MOD
   5    FORMAT (3I3)
        DO 10 I=1, M
  10    READ (5,15) (X(X,J), J=1,N)
  15    FORMAT (8F10.0)
        CALL MEAN (M, N, X, XBAR)
        IF (MOD) 25, 25, 20
  20     DO 22 I=1, M
        DO 22 J=1, N
  22    X(I,J)=9X9I,J)=XBAR(I)/SQRT (S(I,1)/FLOAT(N=1))
        CALL CORR (M, S, R)
        WRITE (6, 24)
  24    FORMAT ('COMPONENTS EXTRACTED FROM CORRELATION
        MATRIX')
        GO TO 40
  25    DO 35 I-1, M
        DO 35 J=1,N
        R(I,J) = S(I,J)/FLOAT (N=1)
  35    R(I,J) = R(I,J)
        WRITE (6, 37)
  37    FORMAT ('COMPONENTS EXTRACTED FROM COVARIANCE
        MATRIX')
  40    CALL CHAR (R,M,VEC,ROOT)
        DO 70 K=1, M
        DO 50 I=1, N
        Y(I) = 0.0
        DO 50 J=1,M
  50    Y(I) = Y(I)+X(J,I)*VEC(J,K)
```

```
      WRITE (6, 55) ROOT (K)
55    FORMAT (/,/,'CHARACTERISTIC ROOT=',
      F10.4,/,/'CHARACTERISTIC VECTOR'))
      WRITE (6, 65)(VEC(1,K),J=1,M)
65    FORMAT(10F8.4)
      WRITE (6, 68)
68    FORMAT ('VALUE OF COMPONENT')
70    WRITE(6,75) (Y(I),I=1,N)
75    FORMAT (5E15,7)
      CALL EXIT
      END
```

APPENDIX L: FORTRAN POGRAMME FOR DISCRIMINANT ANALYSIS

The programme calculates the coefficients of the discrimination function. It evaluates the value of function which enables variables to be classified into groups. An item is in group one (1) if the value function is greater than or equal to zero. The value of function belongs to group two (2) if the value of function is less than or equal zero.

```
C       MEAN, COVARIANCE AND INVERSE NEEDED
        DIMENSION AND DECIMAL PLACES SELECTED
        READ DATA FROM DATA FILE
  5     FORMAT (313, F.10.0)
        DO 10 I= 1, m
 10     READ (5,15)(X 9I,J), J=1,N1
 15     Format (8F10.0)
        DO 20 I=1,M
 20     READ (5,15)(X1(I,J), J-1, N2)
        CALL MEAN (M, N1,X , XBAR)
        CALL COVAR (M, N1, X , S, XBAR)
        DO 30 I=1, M
        XBD (I) = XBAR (I)
        XBP (I) = XBAR (I)
        DO 30 J-1,M
 30     A (I,J) = 59I,J)
        CALL MEAN (M, N2, X1, XBAR)
        CALL COVAR (M, N2, X1,S, XBAR)
        DO 40 I=1,M
        XBD (I) = XBD (I) = XBAR (I)
        XBP (I) =XBP (I) = XBAR (I)
        DO 40 J=I,M
 40     A(I,J) = (A(I,J) = S(I,J))/FLOAT (N1+N2=2)
        CALL INVS (A,M)
        DO 60 I=1,M
        XBAR (I) = 0.0
        DO 60 J=1,M
 60     XBAR (I) = XBAR 9I)+A (I,J)*.XBD(J)
        CON =0.0
        D=0.0
        DO 70 I=1,M
        D=D*XBD (I)*XBAR(I)
 70     CON=CON+XBP (I)*XBAR(I)
        CON=0.5*CON
        I1=1
        WRITE (6,75) I1
 75     FORMAT ('GROUP',I2,/,' ITEM VALUE OF CLASIFICATION, /,'
        NUMBER IR FUNCTION')
        S1=0.0
        DO 90 I=1,N
        CL=0.0
```

```
      DO 78 J=1, M
78    CL=CL+XBAR (J)*X(J,I)
      CL=CL-CON
      IF (CL) 80, 85, 85
79    NC=2
80    GO TO 90
      85    NC=1
      S1=S1+1
90    WRITE (6, 95) I, CL, NC
95    FOMAT (I5, E15.7, 7X, I1)
      I1-2
      Write (6,75)I1
      S2 – 0.0
      DO I10 !=1,N2
      CL=0.0
      Do 98 J-1,M
98     CL = CL=XBAR (J)*XI (J,I)
      CL = CL-CON
      IF (CL) 100, 105, 105
100    NC=1
110    WTITE (6,95)1, CL,NC
      SS1 = N1-N2
      SS2 = N2-S2
      WRITE (6,I115) S1,SS1, S2,SS2
115   FORMAT ('SUMAY OF ERORS', /, 9X' COECT ONCORRECT',/,'
      GROUP 1', 3X, F4.0, 6X,F4.0 /,'GROUP 2' 3X,F4.0, 6X, F4,0)
      WRITE (6,12O)
120    FORMAT (COEFICIENTS",/, VARIABLE',/, 4X, 'SLOPE' )
      DO 125 I=1,M
125    WRITE (6,130) I, XBAR (I)
130   FORMAT (4X, I2, 3X, E15, 7)
      WRITE (, I35) CON
135   FORMAT (3X, 'Y=STAR, E15, 7)
      XBAR (10 = D8FLOAT (N1*N20 / FLAOT (N1*N2)
      XBAR (2) = FLOAT (N1+N2=2) *M/ FLAOT 9N1+N2=M1=10
      WRITE (6, 140) XBAR (1), XBAR (2)
140    FORMAT ('T=SQR, ACTUAL = ', E15, 7, 3X, 'T=SQR, CRITICAL
      = ', E15, 7)
      CALL EXIT
      END
```