# WATER QUALITY IMPACTS OF URBANISATION

## RELATING WATER QUALITY TO URBAN FORM

**Ashantha Goonetilleke & Evan Thomas**
Energy & Resource Management Research Program
Centre for Built Environment and Engineering Research
Queensland University of Technology

April 2004

# EXECUTIVE SUMMARY

## Background

Effective urban resource planning and management entails the mitigation of the impacts of urbanisation on the water environment. The significance stems from the fact that water environments are greatly valued in urban areas as environmental, aesthetic and recreational resources and hence are important community assets. Urbanisation has a profound influence on stormwater runoff quality. This is due to changes to the hydrology of the catchment and the introduction of pollutants resulting from various anthropogenic activities common to urban areas. Though the sources and causes of stormwater pollution are known, its control constitutes an intractable challenge in the drive towards sustainable human settlements.

These difficulties can be ascribed to the fact that the current focus on urban water quality is of relatively recent origin. It is a paradigm shift from the sole focus in the past on quantity issues for flood mitigation. However the techniques and approaches adopted are strongly rooted in quantity research undertaken in the past. This applies not only to modelling philosophies, but also to the conducting of research and data analysis. There is undue reliance on physical processes and the neglect of important chemical processes in describing various stormwater associated phenomena.

Therefore in the absence of appropriate guidance, current approaches to safeguard water quality focus primarily on 'end-of-pipe' solutions. Unfortunately, as a result of entrenched misconceptions, insufficient design knowledge, faulty value judgements or inadequate consideration of life-cycle costs, these approaches tend to be largely ineffective and even counter-productive in the long-term.

## The Project

The research project was located in Gold Coast, Queensland State, Australia. The primary focus of the project was to undertake an in-depth investigation of pollutant wash-off by analysing the hydrological and water quality data from six areas having different land uses in order to correlate urban form to water quality. The project entailed significant field work for water sample collection, extensive laboratory testing and data analysis.

The study areas were selected so as to ensure that there was uniformity in the geological, topographical and climatic variables which could influence the water quality characteristics. The three main catchments selected were already established by Gold Coast City Council and are characterised by differing forms of land development and housing density; ranging from predominantly forested, to rural acreage-residential and forest to mixed urban development. For this project, three smaller subcatchments within the urban catchment were identified for more detailed investigations into effects of increasing urban density on water quality.

Automatic monitoring stations were established at the outlet of each area to record rainfall, stream-flow and a number of water quality parameters. Each station was equipped with an automatic event sampler to augment grab samples taken during low flow conditions. Event samples collected and the grab samples taken during low flow conditions were analysed in the laboratory for a range of water quality parameters.

The data derived were initially analysed using univariate statistical methods to obtain an insight into the trends and patterns of variations in water quality. Subsequently multivariate ‘chemometric’ techniques were applied to identify linkages between various parameters and their correlation with land use. The analytical techniques used included, Principal Component Analysis, Scores Plots and Partial Least Squares Regression.

**The Outcomes**

The primary conclusions derived were:

- The mean values and standard deviations for the primary water quality parameters were generally found to increase with increasing urbanisation. The increase in standard deviations underlies the difficulties in developing predictive water quality models.

- For all three main catchments, the particulate bound component of heavy metals was significantly higher than the dissolved component. This can be attributed to the relatively stable pH values.

- The total concentrations of Al, Mn and Fe, which are sourced from the soil were found to increase with increasing urbanisation probably as a result of erosion.

- There was no appreciable difference in the dissolved components of heavy metal concentrations between study areas which runs counter to the general trend of increasing concentrations with urbanisation. Secondly due to the fact that the dissolved fractions were below detection limit, it could be surmised overall that heavy metals are not a significant issue in these study areas. It is the dissolved component which is readily bioavailable.

- The six study areas behaved quite differently to each other in terms of correlations between different parameters and strong correlationships were not common. This would make it difficult to develop stereotypical strategies for water quality management.

- It was quite common for TOC to be in soluble form (as DOC) in a number of study areas. This parameter can exert a significant influence on urban water quality particularly in relation to the bioavailability of heavy metals and hydrocarbons.

- Primary pollutants such as TN, TP and TOC were commonly in soluble form. This would mean that the effectiveness of structural pollutant abatement measures such as sediment traps is open to question as these measures are dependent on gravity settling.

- The calibration models derived using partial least squares regression for predicting various parameter values were of questionable value generally resulting in large errors of prediction. This in effect means that strong correlationships between parameters were singularly lacking and underlies the reason for the large errors commonly encountered in water quality models.

The outcomes from this study bring into question a number of fundamental concepts routinely accepted in stormwater quality management. The fact that the pollutant characteristics are not consistent across the study areas would mean that the urban form is the overriding factor influencing the water quality. This conclusion would mean that the effectiveness of structural measures would not be universal and stereotypical solutions will not always prove adequate. A significant fraction of the pollution is in dissolved form, it is more bio-available and is therefore more likely to cause pollution in receiving waters. It could well be that this condition is linked to the climatic and rainfall conditions experienced in the study region which significantly influences pollutant composition, build-up and wash-off. Therefore it is important that predictive water quality models developed have the versatility to take these characteristics into consideration.

The above findings underline the need to move beyond the dependency on customary structural measures and end-of-pipe solutions and the key role that urban planning can play in safeguarding urban water environments. The univariate and multivariate statistical data analysis undertaken found that among the different urban forms, stormwater runoff from the area with detached housing in large suburban blocks exhibited the highest concentration and variability of pollutants. This is based on the concentration of various pollutants, their high variability and physico-chemical form. Rural residential on large blocks were only marginally better. It could be concluded that in terms of safeguarding water quality, high density residential development which results in a relatively smaller footprint should be the preferred option.

# ACKNOWLEDGEMENTS

We gratefully acknowledge the financial assistance provided by the Built Environment Research Unit of the Department of Public Works, Queensland, Gold Coast City Council and Queensland University of Technology.

# CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| Al | Aluminium |
| BOD | Biochemical oxygen demand |
| Cd | Cadmium |
| COD | Chemical oxygen demand |
| Cr | Chromium |
| Cu | Copper |
| DOC | Dissolved organic carbon |
| Fe | Iron |
| Hg | Mercury |
| Mn | Manganese |
| $NH_3$ | Ammonia |
| Ni | Nickel |
| $NO_2$ | Nitrite |
| $NO_3$ | Nitrate |
| PAH | Polycyclic aromatic hydrocarbons |
| Pb | Lead |
| SS | Suspended solids |
| TN | Total nitrogen |
| TP | Total phosphorus |
| TSS | Total suspended solids |
| V | Vanadium |
| Zn | Zinc |

# WATER QUALITY IMPACTS OF URBANISATION

# RELATING WATER QUALITY TO URBAN FORM

## 1. INTRODUCTION

### 1.1 Background

Urban expansion transforms local environments and can dramatically alter local conditions. In the context of effective urban resource planning and management, the recognition of the impacts of urbanisation on the water environment is among the most crucial. The significance stems from the fact that water environments are greatly valued in urban areas as environmental, aesthetic and recreational resources and hence are important community assets. Arguably it is the water environment which is most adversely affected by urbanisation. Any type of activity in a catchment that changes the existing land use will have a direct impact on its quantity and quality characteristics.

Land use modifications associated with urbanisation such as the removal of vegetation, replacement of previously pervious areas with impervious surfaces and drainage channel modifications invariably result in changes to the characteristics of the surface runoff hydrograph. Consequently the hydrologic behaviour of a catchment and in turn the streamflow regime undergoes significant changes. The hydrologic changes that urban catchments commonly exhibit are, increased runoff peak, runoff volume and reduced time to peak (ASCE, 1975; Mein & Goyen, 1988). However, urbanisation not only impacts on the hydrologic regime of catchments, but also has a profound influence on the quality of stormwater runoff. These consequences are due to the introduction of pollutants of physical, chemical and biological origin resulting from various anthropogenic activities common to urban areas. As Sartor and Boyd (1972) have identified, urban stormwater runoff constitutes the primary transport mechanism that introduces non-point source pollutants to receptor areas. These contaminants will detrimentally impact on aquatic organisms and alter the characteristics of the ecosystem.

This results in a water body which is fundamentally changed from its natural state (Hall & Ellis, 1985; House et al., 1993).

The pollutant impact and 'shock load' associated with stormwater runoff can be significantly higher than secondary treated domestic sewage effluent (House et al., 1993; Novotny et al., 1985). In summary, the deterioration of water quality, degradation of stream habitats, and flooding, are among the most tangible of the resulting detrimental quality and quantity impacts of urbanisation. Therefore the appropriate management of urban stormwater runoff and streamflow has significant socio-economic and environmental ramifications for urban areas.

## 1.2 The Management Dilemma

The management of quantity impacts of stormwater runoff is relatively straight forward. The common approach is the provision of various structural or physical measures such as detention/retention basins or features such as porous pavements to retain part of the runoff volume and/or attenuate the runoff hydrograph. The primary objective of these measures is to replicate the pre-urbanisation runoff hydrograph. Under appropriate conditions, these structural measures have proven to be effective. However it is important to bear in mind that they are feasible only for relatively low average recurrence interval rainfall/runoff events. The provision of detention facilities for higher order events may not be economically feasible.

Unfortunately, the management of quality impacts due to urbanisation are far more complex. Though the sources and causes of stormwater pollution are widely known (Hall & Ellis, 1985; House et al., 1993), its control constitutes an intractable challenge in the drive towards sustainable human settlements. The current state of knowledge with regards to the process kinetics of pollutant build-up and wash-off is extremely limited. The inter-relationships between various factors and the build-up and wash-off processes of pollutants are complex and little understood. There is no question that the urban environment is adversely affected by a variety of anthropogenic activities which introduces numerous pollutants to the environment. However major uncertainties arise in efforts to articulate the process kinetics of pollutant generation, transmission and dispersion.

These uncertainties and limited knowledge can be ascribed to the fact that the current focus on urban water quality is of relatively recent origin. It is a paradigm shift from the sole focus in the past on quantity issues for flood mitigation and the economic cost it entails. Since the 1970s the move towards urban water quality research is apparent. However the techniques and approaches adopted are strongly rooted in quantity research undertaken in the past. This applies not only to modelling philosophies and water quality models currently available, but also to the conducting of research, data analysis and the reporting of outcomes. There is an undue reliance on physical processes and the neglect of important chemical processes in describing various stormwater associated phenomena.

Therefore in the absence of appropriate guidance and as a direct consequence of the fact that in the past, the major interest of regulatory authorities was quantity impact mitigation, current approaches to safeguard water quality are similarly guided by a primary focus on 'end-of-pipe' solutions. Unfortunately, as a result of entrenched misconceptions, insufficient design knowledge, faulty value judgements or inadequate consideration of life-cycle costs, these approaches tend to be largely ineffective and even counter-productive in the long-term.

The management of water quality impacts do not necessarily lend themselves simple solutions. The provision of appropriate treatment facilities would depend on the targeted pollutants. As an example, the removal of pollutants such as litter is relatively simple. However the removal of other pollutants poses a more challenging task. The provision of gross pollutant traps (GPTs) is a common practice in most urban areas. In addition to litter removal, they may also incorporate sediment removal facilities. These include a screen for litter removal and a sediment trap at the base for sediment removal. However the feasibility of the use of GPTs is open to question due to two significant factors. Firstly, if there is any appreciable time delay in the removal of the collected pollutants, anaerobic conditions could occur in the water collected in the sediment trap due to the decomposition of organic matter present. Therefore in addition to adverse impacts such as odour, the facility could become a pollutant exporter. Also, other than for aesthetic reasons, the contribution to water quality improvement achieved by the removal of gross pollutants or litter is open to question. Secondly, as Allison et al. (1998) have pointed out, the nutrient contribution by leaf litter may not be significant compared to the total nutrient load in stormwater.

The second factor relates to the size range of sediments removed by a sediment trap or for that matter any other treatment measure. Suspended solids act as a mobile substrate for pollutants such as heavy metals and polycyclic aromatic hydrocarbons (PAHs) (Hoffman et al., 1982; Sartor & Boyd, 1972; Shinya et al., 2000; Tai, 1991). As such there is no doubt as to the importance of the removal of suspended solids from urban stormwater runoff. In fact one of the most effective measures for the removal of heavy metals and PAHs from stormwater runoff would be suspended solids separation. However at the same time it is important that the facilities provided are capable of removing the critical size range of sediments which would be carrying a significant pollutant load. Research has shown that due to their physico-chemical characteristics, the finer particulates are more efficient in the adsorption of pollutants and hence will carry a relatively higher pollutant concentration (Andral, 1999; Hoffman et al., 1982; Roger et al., 1998; Sartor & Boyd, 1972). Greb and Bannerman (1997) have raised similar concerns regarding the limited ability of a wet detention pond in removing fine particulates. They found that pollutants are removed at a rate less than the suspended solids removal rate, and the removal rates are influenced by particle size distribution of the suspended solids.

However, outcomes from a number of studies have noted that the fraction of fine particulates in runoff can be small, and as such the total pollutant load would be smaller when compared to the load carried by the coarser particulates (Marsalek et al., 1997; Pitt, 1979). Therefore it has been argued that it is the load rather than the concentration which is of importance and hence the focus should be on the removal of the coarser fraction. Contrary to these findings, other studies have reported a larger fraction of fine particulates being present in stormwater runoff (Andral, 1999; Pechacek, 1994). These contradictory findings clearly point to the fact that it is the catchment characteristics which play the most significant role in urban stormwater runoff quality. Therefore any treatment measures adopted should be designed taking these specific characteristics into consideration. Similarly, street sweeping can be considered as having cosmetic value. The standard street sweeper cannot remove the fine particulates on the road surface that contribute significantly to water pollution (Pitt 1976; Sartor & Boyd 1972).

## 1.3 The Mismatch – Concepts vs. Real World

There have been significant advances in the control of point sources of pollution such as sewage effluent outfalls. However, it is the non point-sources which are the most damaging, the least visible and the most difficult to control. Current approaches to stormwater control, center around conventional concepts of volume and peak flow reduction and primary forms of treatment and reuse. These concepts in themselves are admirable, but their application is open to criticism. Table 1 provides a brief evaluation of the common structural measures adopted in Australia in the implementation of these concepts.

As Table 1 illustrates, commonly adopted measures are based either on, insufficient design knowledge, faulty value judgements or inadequate consideration of life cycle costs. The various structural measures are costly, largely ineffective when dealing with large flows or in dealing with the 'real world' problems and can even be counter productive. Implementation of structural measures can also often be interpreted as being 'seen to be doing something' in response to community pressure.

Modelling is one way where improved design outcomes may be developed. However, based on the current state of knowledge, stormwater pollution does not fit into neat mathematical models which engineers and scientists can use for predictive purposes. Predictive errors of over 100% are common in the use of various models. This is due to the difficulty in mathematical formulation of key anthropogenic activities and the questionable mathematical formulation of key concepts. The quantification of relationships that support models of urban systems is fundamental to the performance of many current models and is crucial for developing improved designs that will work in concert with surrounding natural and constructed systems. The generation and transport of pollution in urban systems during a storm event is multifaceted as it concerns many media, space and time scales (Ahyerre et. al., 1998). These processes are influenced by a range of factors which do not lend themselves to simple mathematical modelling and the simplistic modelling approaches commonly adopted can lead to gross error. The limited data sets available and the large data scatter makes the form of the relationships difficult to determine.

**Table 1 – Issues associated with conventional approaches**

| Treatment device | Primary function/s | Issues |
|---|---|---|
| Retention, detention basins | Volume, peak flow reduction | 1. Can only afford to detain relatively small volumes.<br>2. Sediment build-up, weed infestation entail regular maintenance.<br>3. During dry periods collected water can become anaerobic, breed pests becoming a health hazard and pollutant generator.<br>4. Water feature can attract birds, contributing to pollutant export. |
| Wetlands | Quality improvement | 1. Can only afford to treat relatively small volumes.<br>2. Efficiency in quality improvement not completely proven, particularly removal of very fine sediments, dissolved nutrients.<br>3. Adequate design guidelines for stormwater treatment not available and dependency on wastewater treatment systems.<br>4. Adequate guidelines for weed removal and maintenance not available. |
| Gross pollutant & sediment traps, Vortex devices | Quality improvement | 1. Can only afford to treat relatively small volumes.<br>2. Cannot remove very fine sediments.<br>3. During dry periods collected water can become anaerobic, breed pests becoming a health hazard and pollutant generator.<br>4. Maintenance costs can be very high. |
| Grass swales | Quality improvement | 1. Can be effective in removal of particulate pollutants but not necessarily fine sediment.<br>2. Adequate design guidelines are not available.<br>3. Most paved surfaces such as streets do not have space for grass swales. |
| Rainwater tanks | Volume reduction | Effective in handling only small flows. |

Many factors affect the quality of stormwater runoff with land use being the most important. Though numerous research studies have attempted to relate land use to pollutant loadings, the outcomes reported can be conflicting (Hall & Anderson, 1986; Lopes et al., 1995; Parker et al., 2000; Sartor & Boyd, 1972). This can be attributed to the reliance on physical processes and the neglect of important chemical processes in describing various stormwater associated phenomena.

A fundamental concept in water quality management is the treatment of 'first flush'. However published literature shows that treating the first flush is of doubtful value. There is little conclusive evidence from past research studies to prove the effectiveness of this strategy. As reported by numerous researchers, the 'first flush' has been noted as an important and distinctive phenomenon within pollutant wash-off. The first flush produces higher pollutant concentrations early in the runoff event and a concentration peak preceding the peak flow (Deletic, 1998). It has significant economic implications in relation to the management and treatment of urban stormwater runoff. The economic significance stems from the fact that structural measures for water quality control such as detention/retention basins are often designed for the initial component of urban runoff.

Hall and Ellis (1985) have claimed that the first flush phenomenon is over emphasised and only 60–80% of storms exhibit an early flushing regime. Other researchers too have observed that the first flush is very frequent in urban runoff, but not necessarily always (Angino et al., 1972; Cordery, 1977). As Deletic (1998) has pointed out, in view of the diverse definitions, varying sampling strategies and data collection methods, it is difficult to compare results from different studies. This could possibly explain the differences in reported observations in relation to the occurrence of the first flush.

The qualitative descriptions commonly found in literature cannot be used as an appropriate basis to plan structural pollutant abatement measures. In understanding the first flush, the major difficulty arises with respect to defining this phenomenon in a quantitative manner. As Bertrand-Krajewski et al., (1998) and Saget et al., (1996) have pointed out, the problem stems from the fact that the 'initial component of runoff' which carries the first flush is never precisely defined. This is despite its commonly reported occurrence in qualitative terms. A mere increase in pollutant concentration at the beginning of a storm cannot be interpreted in a quantitative manner. In the context of stormwater pollution management, it is the pollutant load rather than pollutant concentration that is of significance. Due to the corresponding runoff volume being low and despite the increase in pollutant concentration, the pollutant load during the initial phase of runoff could be relatively low when compared to the overall load carried by the runoff event, (Barrett et al., 1998; Cordery, 1977). Therefore under these circumstances whether or not the first flush exists and if so, its characteristics are highly debateable issues. It could be postulated that the first flush is only a convenient expression to describe a concentration peak.  As Delectic (1998) has observed, it is clear that the first flush load cannot be calculated using a universal set of rainfall, runoff and climate characteristics or generic types of regression curves.

The above findings underline the need to move beyond the dependency on customary structural measures and end-of-pipe solutions. The mere provision of standard structural measures is not necessarily effective in removing water quality pollutants per se. Any structural measures to be adopted should depend on targeted pollutants and management strategies adopted should take into consideration the rainfall, runoff and physical characteristics of the area.

# 2. REPORT DETAILS

## 2.1 Background to the Report

This document is the second in a series of two reports focusing on water quality impacts of urbanisation. The research project was initiated at the request of the Gold Coast City Council and the Built Environment Research Unit of the Department of Public Works. The major objective of the research undertaken was to relate stormwater runoff quality to different urban forms. The initial report consisted of a 'state of the art' review of research undertaken in the arena of urban water quality. This report provides a comprehensive outline of the experimental study undertaken.

## 2.2 Report Objectives

The safeguarding of urban water quality is being afforded increasing importance due to the recognition of urban water resources being important environmental assets. It is in this context that the design of the urban form is being subjected to greater scrutiny and innovations adopted in order to minimise its ecological footprint in relation to the water environment. However the relationships between urban form and water quality are not intuitively obvious. This is because the underlying processes which influence pollutant generation, transmission and dispersion are complex and poorly understood. These processes do not lend themselves to simple mathematical modelling.

The key role played by various anthropogenic activities and the difficulty in their mathematical formulation further adds to the complexity of the inherent processes. Consequently, the mere adoption of structural or regulatory measures for urban water pollution mitigation will not suffice. It is important that the mitigative management strategies adopted are appropriately formulated based on a comprehensive awareness of influential factors. This requires a multifaceted strategy that would encompass:

- The continuous improvement and/or development of strategies based on currently available 'state of the art' research outcomes.
- The undertaking of practical research in areas where there is a discernible lack of in-depth knowledge.

The research report provides a detailed discussion of the experimental study. This includes the field work, water sampling and testing, data analysis undertaken and the significant conclusions derived. In the long term it is hoped that this report will implicitly contribute to the development of a comprehensive knowledge base, which will form the basis for the formulation of credible and innovative urban growth management strategies to mitigate the adverse environmental impacts commonly associated with urbanisation.

## 2.3 Scope and Outline of the Report

This research report outlines the experimental study undertaken in order to relate water quality to urban form. The study focused on the primary water pollutants of physical and chemical origin and the microbiological quality of water did not form a part of it.

Chapter 3 outlines the research project including physiographical and land use details and water sampling and testing protocols. The analytical methods adopted including multivariate statistical methods are discussed in Chapter 4. a detailed discussion of the results obtained from the data analysis and conclusions derived are given in Chapter 5. the results obtained a clear understanding of how the urban form can influence water quality.

# 3. RESEARCH PROJECT

## 3.1 Project Description

The water quality research project was located in The City of Gold Coast, Queensland, Australia. The primary focus of the project was to undertake an in-depth investigation of pollutant wash-off by analysing the hydrologic and water quality data from six areas having different land uses in order to correlate urban form to water quality. The project initially commenced as a collaboration between Gold Coast City Council and Queensland University of Technology in July 1999 and encompassed three existing catchment study areas. Subsequently three additional subcatchments were included in December 2001. The project entailed significant field work for water sample collection, extensive laboratory testing and data analysis. The data analysis undertaken includes rainfall event based water quality data collected to end 2003.

## 3.2 Study Areas

The study areas were selected so as to ensure that there was uniformity in the geological, topographical and climatic variables, which could possibly influence the water quality characteristics. The three main catchments were established by the Gold Coast City Council in 1998 and are characterised by the same geology based on the Neranleigh-Fernvale metasediments and similar predominant soil types mainly Kurosols (Isbell 1996). However they have differing forms of land development and housing density; ranging from predominantly forested in the upper Bonogin Valley (or Bonogin), to rural acreage-residential (un-sewered) and forest in the lower Bonogin Valley (or Hardy), to mixed urban development (sewered) in Highland Park (or Hinkler) catchment. The study catchments have been mapped in detail, including detailed landuse, geology, soils, topography and stream morphology (GCCC, 2001).

Three smaller subcatchments within the Highland Park catchment were identified for more detailed investigations into effects of increasing urban density on water quality. These subcatchments are a tenement townhouse development of around 60 properties (Alextown), a duplex housing development with around 20 dual occupancy residences (Gumbeel) and a high-socio-economic single detached dwelling area (Birdlife). The locations of the study areas are shown in Figures 1 and 2. Table 2 provides a summary of the catchment and geological characteristics whilst Table 3 provides a summary of land cover characteristics for each area.

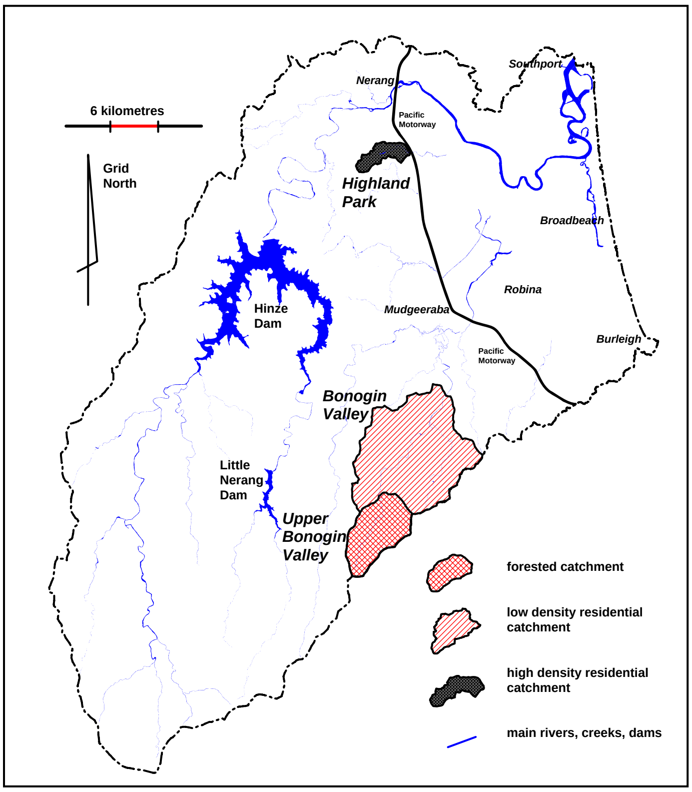



**Figure 1 – Locations of main catchments**

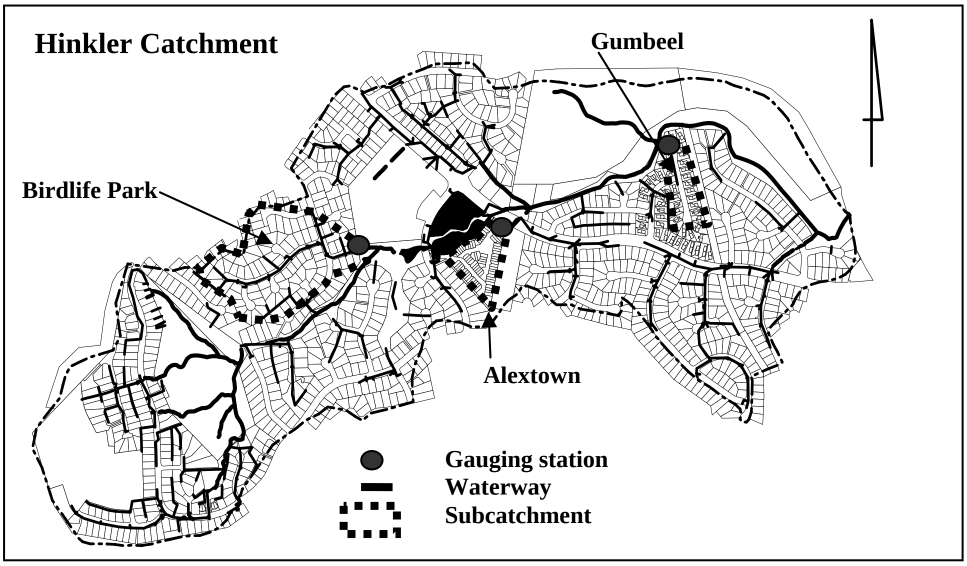



**Figure 2 – Locations of the urban subcatchments**

**Table 2 – Selected study catchments and their geology and soil characteristics**

| Catchment name | Area (ha) | Landuse | Geology | Soil (Isbell 1996) |
|---|---|---|---|---|
| Upper Bonogin Valley Catchment (Bonogin) | 647 | Forest land: 79%<br>Rural residential: 19%<br>Road reserve: 2% | Alluvium: 1.1%<br>Binaburra Rhyolite: 3.5%<br>Neranleigh-Fernvale beds: 95.4% | Grey Dermosols and Rudosols: 1.1%<br>Brown Dermosols and Chromosols: 3.5%<br>Red, Brown, Yellow and Grey Kurosols, Red Ferrosols and Tenosols: 95.4% |
| Lower Bonogin Valley Catchment (Hardy) | 2726 | Forest land: 40.5%<br>Rural residential: 54.9%<br>Road reserve: 4.6% | Alluvium: 2.1%<br>Binaburra Rhyolite: 0.8%<br>Neranleigh-Fernvale beds: 97.1% | Grey Dermosols and Rudosols: 2.1%<br>Brown Dermosols and Chromosols: 0.8%<br>Red, Brown, Yellow and Grey Kurosols, Red Ferrosols and Tenosols: 97.1% |
| Highland Park Catchment (Hinkler) | 161 | Forest land: 9.2%<br>Rural residential: 2.8%<br>Urban residential: 60.4%<br>Road reserve: 16.3%<br>Other (commercial, grazing land etc): 11.3% | Neranleigh-Fernvale beds: 100% | Red, Brown, Yellow and Grey Kurosols, Red Ferrosols and Tenosols: 100% |

**Table 3 – Characteristics of selected study areas**

| Study area | Extent (ha) | Land cover | |
|---|---|---|---|
| | | Impervious area (buildings, roads) | Pervious area (forest, grassland) |
| Forested catchment – Upper Bonogin Valley (Bonogin) | 647 | 2% | 98% |
| Rural acreage residential catchment – Lower Bonogin Valley (Hardy) | 2726 | 9% | 91% |
| Urban Residential Catchment – Highland Park (Hinkler) | 161 | 55% | 45% |
| Town Houses – Alextown subcatchment | 2 | 60% | 40% |
| Duplex Housing – Gumbeel subcatchment | 7.5 | 70% | 30% |
| Detached housing – Birdlife subcatchment | 8.5 | 60% | 40% |

### 3.3 Water Sample Collection

Automatic monitoring stations were established at the outlet of each area to record rainfall, stream-flow and a range of water quality parameters. Each station was equipped with an automatic event sampler to augment grab samples taken during low flow conditions. The automatic monitoring stations recorded rainfall, streamflow, pH, electrical conductivity (EC), temperature and dissolved oxygen concentration (DO). Event samples collected by the automatic sampling devices and the grab samples taken during low flow conditions were analysed for total organic carbon (TOC), dissolved organic carbon (DOC), suspended solids (SS), total dissolved solids (TDS), particle size distribution, heavy metals, polycyclic aromatic hydrocarbons (PAHs), total nitrogen (TN) and total phosphorus (TP).

### 3.4 Water Sample Testing

The primary water quality parameters were evaluated for individual water samples whilst parameters such as heavy metals and polycyclic aromatic hydrocarbons (PAHs) were evaluated on the basis of the event mean concentration (EMC). The sample for determining EMC values was formed from the individual samples collected. The event based water samples collected from each monitoring stations was combined according to the outflow hydrograph to form a single flow weighted event mean sample (Lopes et al. 1995). The parameter values obtained were taken to be the EMC for the specific rainfall event. Additionally, the EMC could also be determined from the individual water samples on a flow weighted basis. The low flow samples were tested individually. The water samples were kept under refrigeration at $4^0$C prior to analysis.

#### 3.4.1 pH-analysis

The pH was measured using a combined pH/EC meter. The TPS WP-81 pH/EC meter was calibrated prior to analysis using standard solutions. The pH measurements were conducted at room temperature.

#### 3.4.2 Electrical Conductivity (EC)

The EC was measured on delivery of the water samples to QUT using a combined pH/EC-meter. The procedures adopted were the same as for the pH measurements.

#### 3.4.3 Total Suspended Solids (SS) and Total Dissolved Solids (TDS)

TSS and TDS were analysed using Standard Method No. 2540D and 2540C respectively (APHA 1999). A well mixed portion of the sample was filtered through a glass fibre filter (pore diameter 0.45µm) which had been pre-cleaned and weighed prior to filtering. The filter paper was dried at $103\text{-}105^0$C and re-weighed. The difference was taken as the weight of TSS. The filtrate was evaporated to dryness in a porcelain dish of known weight at $180^o$C. The increase in dish weight represented the TDS.

#### 3.4.4 Total Organic Carbon (TOC)

TOC was measured using a Shimadzu TOC-5000A analyser. The analysis for TOC was based on Standard Method No. 5310B (APHA 1999).

### 3.4.5 Total Nitrogen (TN)

The nitrogen concentration was analysed using the total sample as well as the dissolved phase. The difference in concentration between the original and dissolved sample was defined as the particulate nitrogen concentration (APHA 1999). The dissolved phase was the filtrate passing through a 0.45μm glass fibre filter. The nitrogen species analysed included nitrate ($NO_3$), nitrite ($NO_2$) and total kjeldahl Nitrogen (TKN) using Standard Method Nos. 4500-$NO_3$-A, 4500-$NO_2$-B and 4500-$N_{org}$-B respectively (APHA 1999). A Hach DR/4000 Spectrophotometer was used for the analysis.

### 3.4.6 Total Phosphorous (TP)

Similar to nitrogen analysis, total and dissolved phosphorous concentrations were measured and the difference was defined as the particulate phosphorous fraction. The analysis was undertaken according to Standard Method 4500-P using a Hach DR/4000 Spectrophotometer.

### 3.4.7 Heavy metals

The sample was initially divided into a dissolved phase and a particulate phase. The dissolved phase was the filtrate passing through a 0.45μm nitrocellulose filter as described in Standard Method No. 3030E (APHA 1999). The dissolved sample was nitric acid preserved for analysis and the particulate sample was digested using nitric acid according to Standard Method No. 3030E (APHA 1999). Following digestion and preservation, the samples were tested for eight metal elements, namely; zinc (Zn), aluminium (Al), cadmium (Cd), iron (Fe), Manganese (Mn), lead (Pb), chromium (Cr) and copper (Cu) using Inductively Coupled Plasma-Mass Spectroscopy (ICP-MS). Reagent blanks and duplicate samples were used for quality control purposes.

### 3.4.8 Polycyclic Aromatic Hydrocarbons (PAH)

Similar to the heavy metal analysis, the samples were divided into a dissolved and a particulate phase prior to testing. The dissolved sample was extracted using liquid-liquid extraction according to Standard Method No. 6440B (APHA 1999) with Dichloromethane (DCM) as a solvent. The dissolved sample was poured into a separatory funnel and mixed with 60mL of DCM. The separatory funnel was then shaken for two minutes with periodic venting to release excess pressure. The organic layer was allowed to separate from the water for ten minutes before the DCM was collected in a 250mL Erlenmeyer flask. Each sample was extracted with DCM three times. The extracted sample was concentrated down to 1 mL using a Kuderna-Danish evaporator and nitrogen gas as described in EPA Method No. 610 (US EPA 1986). Extracted samples were kept under refrigeration until analysis.

The particulate phase was extracted using sonication with DCM-Acetone (Guerin 1999). Similar to the dissolved phase, the extracted sample was concentrated down to 1mL prior to analysis. The particulate extract was dried and cleaned using a silica gel/sodium sulphate column as described in EPA Method No. 610 (US EPA 1986).

Analysis of PAHs was undertaken using a ThermoFinnigan PolarisQ Gas Chromatograph-Mass Spectroscopy (GC-MS). Sixteen species were analysed and similar to heavy metals, reagent blanks and duplicate samples were used for quality control purposes. Both internal and external standards were used to check matrix recoveries. The samples were analysed for Napthalene, Acenapthene, Phenanthrene, Chrysene, Anthracene, Acenapthylene, 2-Bromonaphtalene, Pyrene, Flourene, Flouranthene, Benzo[a]anthracene, Benzo[b]flouranthene, Benzo[a]pyrene, Indeno[1,2,3-cd]pyrene, Dibenzoanthracene, Benzo[g]pyrene. These species are the US EPA specified sixteen priority pollutants (US EPA 2002).

# 4. ANALYTICAL METHODS

## 4.1 Univariate Statistical Analysis

Univariate statistical analysis was undertaken to determine the mean and standard deviation for the primary water quality parameters for the six study areas. It was anticipated that these values would provide an insight into the trends and patterns of variations in water quality with land use. This would provide further information to underpin the outcomes derived from more detailed data analysis. At the initial stages of the research project, using correlation matrices, Rahman et al. (2002) developed a set of preliminary predictive equations relating key pollutant parameters and rainfall characteristics. This was based on the data obtained from July 1999 to July 2001 for the three primary catchments. For Bonogin, an equation was developed to predict TP from TSS. This equation had a high coefficient of determination (95%) and a relatively small standard error of estimate (25%). Unfortunately in the case of Hardy and Hinkler catchments, the various predictive equations developed did not reflect the same degree of statistical accuracy. However most importantly, the study by Rahman et al. (2002) highlighted the importance of developing a deeper understanding of the interactions and linkages between influential parameters.

## 4.2 Multivariate Statistical Analysis

Subsequent to the univariate study, multivariate chemometric techniques were applied to identify linkages between various pollutant parameters and correlations with land use. The analytical techniques used included, Principal Component Analysis (PCA), Scores Plots and Partial Least Squares Regression (PLS).

### 4.2.1 Principal Component Analysis (PCA)

Essentially, PCA is used for pattern recognition. PCA is a multivariate statistical data analysis technique which reduces a set of raw data into a number of principal components which retain the most variance within the original data in order to identify possible patterns or clusters between objects and variables. Detailed descriptions of PCA can be found elsewhere (Adams, 1995; Kokot et al., 1998; Massart et al., 1988). PCA has been used extensively for various applications related to water quality. As examples, Wunderlin et al. (2001) used PCA for the evaluation of spatial and temporal variations in river water quality and Marengo et al. (1995) to characterise water collected from a lagoon as a function of seasonality and sampling site and for the identification of significant discriminatory factors. Hamers et al. (2003) employed PCA to study pesticide composition and toxic potency of the mix of pollutants in rainwater and Librando et al. (1995), for the analysis of micropollutants in marine waters. Similarly Vazquez et al. (2003) used PCA to evaluate factors influencing the ionic composition of rainwater in a region in NW Spain

In order to undertake PCA, the water quality concentration data as mg/L was arranged into a matrix for each study area. The columns defined the variables and the rows, the sample measurement. The raw data was initially subjected to pre-treatment to remove 'noise' which may interfere in the analysis (Adams, 1995, Kokot et al., 1998). Firstly, the data was log transformed to reduce data heterogeneity. Following this, the transformed data was column-centred (column-means subtracted from each element in their respective columns) and standardised (individual column values divided by the column standard deviations). PCA was undertaken on the transformed data for pattern recognition and for the identification of correlations between selected variables.

In undertaking a PCA a fundamental issue to be determined is the number of principal components (PCs) that needs to be considered which will describe most of the variance in the data available. This issue is generally resolved using a Scree Plot.

### Scree Plot

The scree plot is an empirical tests used to determine the number of significant factors in analytical data analysis. The scree plot is so named for its analogous description of the straight line of rubble and boulders which form at the pitch of sliding stability at the foot of a mountain (Adams 1995; Cattell 1966). This describes the levelling off of the residual variance after the significant number of factors has been determined. The scree plot represents the residual variance, V, as a function of the number of eigen vectors that have been extracted. The residual variance of the $r^{th}$ eigen vector is defined by:

$V(r^*) = \sum_{k=r^*+1}^{r} \lambda_k^2$ with $1 \le r^* \le r$ where r is the number of nontrivial eigen values.

Thereafter, it is assumed that the structural eigen vectors explain successively less variance in the data, until the resulting curve levels off at a point r* when the structural information of the data is nearly exhausted. This point determines the number of factors with the remaining information contributing noise within the data set.

### 4.2.2 Scores Plot

The derivation of principle components can be mathematically described as finding a new set of uncorrelated variables (scores) $V_1$ $V_2$…$V_n$, such that the variance decreases from $V_1$ to $V_n$. The scores plot is simply a plot of the scores in the corresponding principal components that have been found to be significant (ie contain most of the data information). The scores plot can provide important information relating to the specific objects (or samples) analysed through PCA. Firstly, specific clusters or groupings between similar object scores can be identified. This for example, can be used to distinguish between water samples that are taken from a polluted section of a river to those from an unpolluted section. Secondly, in using the scores in conjunction with biplots, the correlations of specific clusters with relevant variables can be identified (Adams 1995; Massart et al. 1988). For example, the polluted section of a river may be highly correlated with nitrogen or phosphorus and have a low pH, whereas the unpolluted section may have neutral pH and low nutrients.

### 4.2.3 Partial Least Squares Regression (PLS)

PLS is a generalisation of the more common multiple linear regression (MLR). However, unlike MLR, PLS has the ability to analyse data that is strongly collinear and noisy, and uses numerous X-variables in order to simultaneously model one to several response Y-variables (Wold et al. 2001). This technique has been significantly useful in analytical chemistry in predicting concentrations from spectral data. PLS involves the generation of abstract factors (latent variables) of both the predictor (X) and response (Y) data, which are then rotated towards each other in order to optimise the regression between the two data sets. There are two common forms of PLS that are typically employed and the decision as to which one to adopt depends on the number of response data being regressed. PLS1 is used to model singular Y-variables which are fairly independent and measure different factors. However on the other hand if the Y-variables measure similar objects and are highly correlated, the PLS2 model is utilised, which uses the X-variables to predict several Y-variables at once.

PLS has been used for a diverse array of analysis relating to water quality. As examples, Marhaba et al. (2003) employed PLS to predict the dissolved organic carbon concentration in river water, Librando et al. (1995) for the analysis of micropollutants in marine waters and Dabakk et al. (1999) for developing predictive water quality models.

In the analysis undertaken, the predictor variables X are the concentrations measured from the catchment which were used to predict the selected response variable (Y), such as total phosphorus utilising the PLS1 routine. Firstly, the data was log transformed to reduce data heterogeneity. Following this, the transformed data was column-centred (column-means subtracted from each element in their respective columns) and standardised (individual column values divided by the column standard deviations). These are standard data pre-treatment options and are utilised in order to remove or reduce irrelevant sources of variation or 'noise' which may interfere in the analysis. Secondly, a decision whether to use just a calibration data set to develop the model or include a validation set as well needed to be made. Although using just a calibration set will provide a good indication of whether the model is useful for prediction, a second, independent validation data set is useful as it will allow a more beneficial test of the model's overall efficiency.

The PLS1 model estimates new X-scores, which are estimates of the latent variables. These X-scores are both predictors of Y and models of X (both modelled by same latent variables). The X-scores are generally optimised and represented by the least amount of orthogonal factors which provide minimal error in the model. The PLS1 model uses these scores and develops error predictions based on calculated reduced eigen values (or error of calibration), cross-validation (simulates independent validation by a leaving-one-out routine where each sample is left out to allow an independent estimate of error for that sample) or predicted residual error sum of squares or PRESS. PRESS is used to evaluate the ability of the model to predict the X variable, with the model developed using the calibration data and subsequently as a check against a validation data set. Basically, the number of factors or components required in the PLS1 model is determined when the predicted errors are minimised. Generally, if only a calibration data set is used, the cross-validation provides a more suitable estimate. However, if a validations set is available; PRESS would provide a better error prediction.

Once the number of components are determined, the X-scores are recalculated to include the necessary information retained in those components (or latent variables). These components generally contain the most variance in the data, with the other components usually made up of ‘noise’. These newly acquired data is then used to predict the Y-variable. General statistics showing model performance are also provided.

The difference between the recalculated X-scores and the measured X, and the predicted Y-variables and measured Y, called the residuals, are used to establish how well the model fits the predictions. To express the model’s performance, standard measures of fit, $R^2X$ (amount of variation of X explained in terms of sum of squares) and $R^2Y$ (amount of variation of Y explained in terms of sum of squares), $Q^2$ (explained variation of Y using cross-validated data) and RMSEP (Root Mean Squared Error of Prediction) are determined. Generally, the $R^2$ values provide an explanation of the amount of variance explained in the model and the predicted data. The $Q^2$ value provides and indication of the amount of the predicted ‘Y’ explained by the cross-validation method. As these are presented as percentages, typically the higher the percentage the better the model and predictions. However, this entails that the errors in the model are fairly small and do not influence the predictive ability. As such the RMSEP is used as an assessment of the predictive errors, hence the smaller the RMSEP, the better the model.

# 5. RESULTS AND DISCUSSION

## 5.1 Univariate Statistical Analysis

Table 4 gives the mean and standard deviation for the measured parameters for all the study areas for the measured rainfall events from July 1999 to December 2003. Based on the data given in Table 4 the following conclusions can be derived.

**The primary catchment areas; Bonogin, Hardy, Hinkler**

1. There was no significant change in pH values obtained for individual catchments and values between catchments. This could be related back to the soil conditions prevalent in the study areas which are of similar characteristics.

2. Considering the other primary water quality parameters measured, EC, TN, TP, SS and TOC, the mean values and the standard deviations obtained increase with increasing urbanisation. The increase in mean values can be attributed to the increase in the pollutant load in stormwater runoff due to urbanisation. The increase in standard deviations is even more significant. It indicates a high variability in stormwater runoff quality. This underlies the difficulties in predicting the quality of urban runoff and the large margins error usually associated with predictive modelling.

3. The trend of increasing standard deviation with urbanisation does not apply only in the case of SS for Hardy and Hinkler catchments. It is postulated that this is due to the fact that the Hardy catchment does not have kerb and channelling and that this leads to relatively higher erosion along its roadways.

4. It is also significant that despite the high canopy cover in the Bonogin catchment, the Hinkler catchment comparatively still exhibits the highest TOC concentration.

**Table 4 – Mean and Standard Deviations of the measured parameters**

| Study area | Bonogin | | Hardy | | Hinkler | | Alextown | | Gumbeel | | Birdlife | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Parameter** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| pH | 6.49 | 0.20 | 6.67 | 1.28 | 6.78 | 0.33 | 6.73 | 0.25 | 6.69 | 0.38 | 7.04 | 0.39 |
| EC (uS/cm) | 91.44 | 56.74 | 151.99 | 73.77 | 222.45 | 138.77 | 74.78 | 28.48 | 103.11 | 46.26 | 161.74 | 84.48 |
| SS (mg/L) | 80.71 | 112.27 | 149.41 | 209.52 | 171.52 | 111.43 | 130.91 | 253.95 | 58.49 | 59.48 | 181.70 | 238.16 |
| TN (mg/L) | 2.34 | 0.47 | 2.66 | 3.18 | 6.77 | 14.08 | 2.06 | 1.11 | 3.31 | 3.79 | 2.01 | 1.96 |
| TP (mg/L) | 0.27 | 0.15 | 0.23 | 0.39 | 0.48 | 0.80 | 0.45 | 0.27 | 0.75 | 0.72 | 0.73 | 0.96 |
| TOC (mg/L) | 13.94 | 6.28 | 21.50 | 27.51 | 102.75 | 466.28 | 11.35 | 4.16 | 10.37 | 5.76 | 11.52 | 6.23 |
| Zn (mg/L) P | 0.17 | 0.19 | 0.22 | 0.22 | 0.17 | 0.20 | 0.18 | 0.17 | 0.05 | 0.03 | 0.14 | 0.14 |
| Cu (mg/L) P | 0.01 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.04 | 0.03 | 0.01 | 0.01 | 0.04 | 0.03 |
| Cd (mg/L) P | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 |
| Cr (mg/L) P | 0.02 | 0.00 | 0.02 | 0.01 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 |
| Pb (mg/L) P | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.02 | 0.02 | 0.01 | 0.02 | 0.03 | 0.02 |
| Al (mg/L) P | 1.67 | 1.26 | 2.43 | 1.71 | 4.12 | 2.62 | 2.52 | 1.90 | 1.08 | 0.78 | 4.31 | 3.49 |
| Mn (mg/L) P | 0.05 | 0.03 | 0.10 | 0.07 | 0.13 | 0.09 | 0.06 | 0.04 | 0.03 | 0.03 | 0.19 | 0.16 |
| Fe (mg/L) P | 1.66 | 1.27 | 2.56 | 1.53 | 4.42 | 3.37 | 2.77 | 1.79 | 1.09 | 0.94 | 16.03 | 16.09 |
| Zn (mg/L) D | 0.04 | 0.01 | 0.05 | 0.02 | 0.04 | 0.01 | 0.06 | 0.02 | 0.05 | 0.02 | 0.04 | 0.01 |
| Cu (mg/L) D | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| Cd (mg/L) D | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 |
| Cr (mg/L) D | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Pb (mg/L) D | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Al (mg/L) D | 0.12 | 0.08 | 0.11 | 0.09 | 0.04 | 0.05 | 0.06 | 0.09 | 0.03 | 0.01 | 0.03 | 0.02 |
| Mn (mg/L) D | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Fe (mg/L) D | 0.11 | 0.05 | 0.17 | 0.06 | 0.08 | 0.05 | 0.04 | 0.03 | 0.05 | 0.03 | 0.03 | 0.03 |
| Total PAH (ppb) | 14.06 | 8.82 | 16.78 | 4.89 | 28.88 | 4.35 | 20.80 | 11.00 | 9.55 | 1.43 | 18.51 | 13.08 |

P – Particulate fraction D – Dissolved fraction SD – Standard Deviation

5. In the case of heavy metals, for all three catchments the particulate bound components was significantly higher than the dissolved component. In fact the dissolved components were below the detection limit for most of the heavy metal species tested. It is the dissolved component which is readily bioavailable. As Tai (1991) has pointed out the pH value has a significant impact on the desorption of pollutants adsorbed on particulates. Furthermore, Tai (1991) noted that the ratio of trace metals released at pH 6 against pH 8.1 for similar suspension concentrations was about 180 for Zn, 45 for Pb and 25 for iron (Fe).The values obtained for pH indicated that it is quite stable at the point of measurement. However if there are subsequent changes to the pH values the dissolved heavy metal concentrations would change accordingly.

6. The high concentration values of Al, Mn and Fe can be attributed to the fact that these are being sourced from the soil. This would be the result of erosion as runoff flows over the soil surface. Once again the values obtained increases with increasing urbanisation. It is also significant that due to the neutral nature of the surface runoff pH, the dissolved concentrations of these metals are in the same range despite the wide variation in the suspended solids concentrations.

7. In the case of total polycyclic aromatic hydrocarbons (PAHs), Hinkler exhibits the highest concentration. This is to be expected due to the urban land use in this catchment and the primary source of PAHs being road surface runoff. It is also important to note that the standard deviation for Hinkler is relatively low. This would mean that the concentration vlues are consistently high for the catchment.

**The subcatchment areas; Alextown, Gumbeel, Birdlife**

1. Once again the pH values are relatively stable and there is no significant difference between the three study areas.

2. In terms of the other primary pollutants, the trend in changes in data values is not very clear as in the case of the primary catchment areas. However other than for TN and SS, Birdlife subcatchment shows the highest variability in pollutant concentrations. In terms of the other two study areas, it is difficult to distinguish between Alextown and Gumbeel.

3. It was not possible to determine the reasons for the higher standard deviation for SS for Alextown which is only marginally higher than for Birdlife. However it is important to note that the average SS concentration for Birdlife is higher when compared to Alextown.

4. In the case of heavy metals the conclusions noted in items 4 and 5 above for the three primary catchments are also applicable for the three subcatchments.

5. The total PAH concentration and standard deviation, the values for Alextown and Birdlife are in the same range whereas for Gumbeel the values are significantly lower. The primary source of PAHs is from motor vehicles and the resulting runoff from roadways. For all three subcatchments, the percentage impervious area is in the same range. Therefore it could postulated that this difference in concentrations would be due to lesser vehicle usage in the case of Gumbeel. In the case of Alextown even though the total length of roadways is relatively small there is a greater concentration of houses.

**General Comments**

It is important to note that in the case of all the study areas there was no appreciable difference in heavy metal concentrations other than Fe in Birdlife subcatchment. This runs counter to the general trend in increase in heavy metal concentrations with urbanisation as noted in numerous research studies. The type of urbanisation present in the study areas was residential and they are located a distance away from the industrial areas in the region. Therefore the primary source of heavy metals in these areas would be from roadways. As evident from the data given in Table 3, there is an appreciable difference in percentage of road surfaces in the different study areas. Therefore due to the fact that the dissolved heavy metal concentrations were below detection limit, it could be surmised overall that the heavy metals are not a significant issue in these study areas.

### 5.2 Multivariate Statistical Analysis

For PCA, the EMC values obtained for individual rainfall events were used as the objective was to investigate correlationships between different parameters values. However PLS regression was undertaken using parameter values obtained for individual samples as the objective was to develop predictive relationships for deriving other parameter values.

#### 5.2.1 Bonogin catchment

**PCA Analysis**

The PCA undertaken consisted of six variables and six objects. In the case of Birdlife, the overall length of roadways is significantly higher. This in turn would translate to relatively higher concentrations. From the scree plot as shown in Figure 3, it was found that two initial PCs were retaining 79% of the overall variance. The Scores plot and the Biplot developed accordingly are shown in Figure 4 and 5 respectively. In the Biplot, the eigen vectors representing parameters which are close together can be considered to be correlated whilst an angle of $90^0$ or over would mean that the parameters are not correlated.

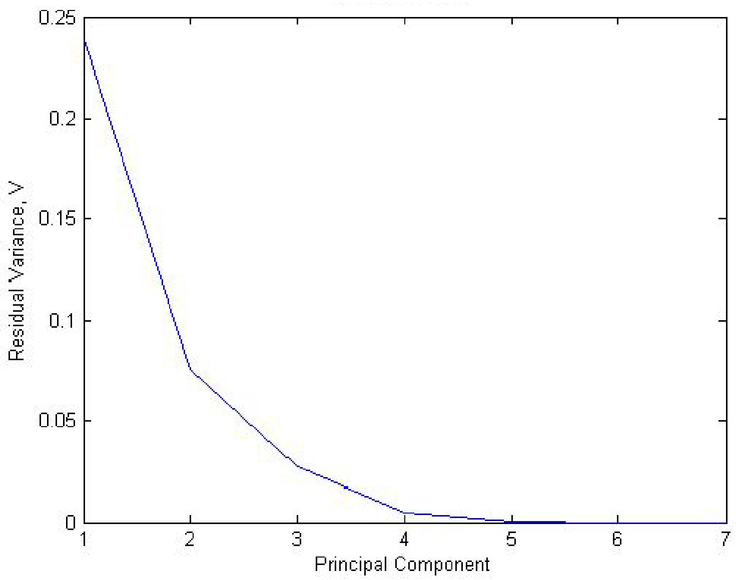



**Figure 3 – Scree Plot for Bonogin Catchment**

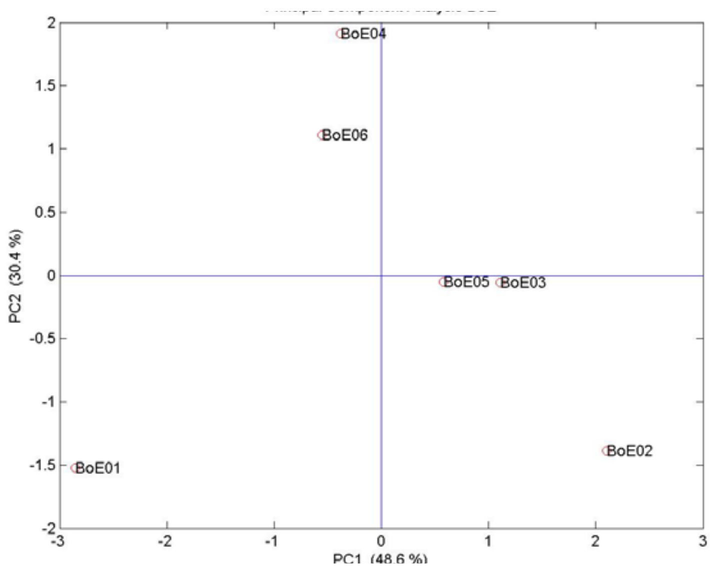



**Figure 4 – Scores Plot for Bonogin Catchment**

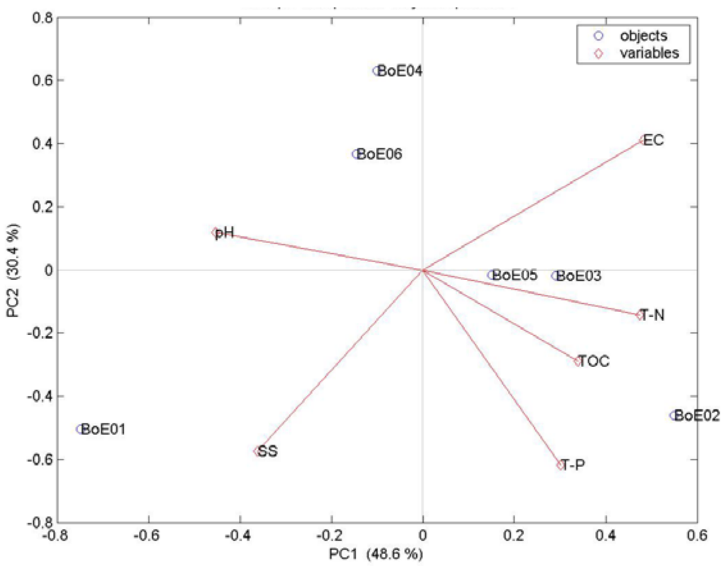



**Figure 5 – Biplot for Bonogin Catchment**

Based on the Scores and the Biplot the following conclusions can be derived:

- Both pH and EC are not correlated with the other variables.

- TN, TP and TOC are correlated as the vectors representing these parameters are close to each other.

- As the vectors representing TN, TP and TOC are perpendicular or at a greater angle with the SS vector it can be surmised that these parameters are primarily in soluble form.

- The fact that TOC is not correlated with SS would mean that the organic carbon is primarily in dissolved form or as dissolved organic carbon (DOC). DOC is an important water quality parameter. DOC absorbs and reacts with sunlight energy, complexes metals, provides an energy source for microorganisms and associates with hydrophobic substances. Additionally, organic carbon adsorbed on suspended solid particles enhances their sorption capacity for combining with hydrocarbons and some heavy metals. Though some of these characteristics can be considered to be beneficial, the organic matter is liable to microbial decomposition, thereby returning the pollutants back into the dissolved phase (Parks & Baker 1997; Roger et al. 1998; Warren et al. 2003; Westerhoff & Anning 2000).

- The fact that TN, TP and TOC are primarily in soluble form would mean that conventional structural pollutant abatement measures such as sediment traps will not be particularly effective other than for the removal of SS.

**PLS Regression**

*PLS – TOC regressed* (excludes pH) No change as evenly correlated

Calibration = 21 objects

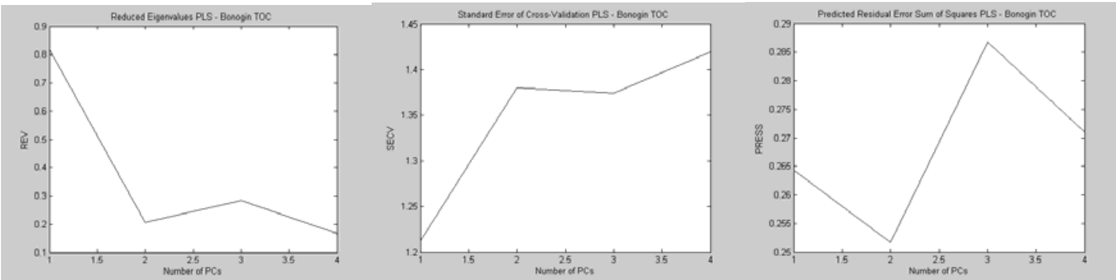
Validation = 21 objects

PLS was performed on the data which was separated as:

- Y = TOC, X = EC, TP, TN and SS
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

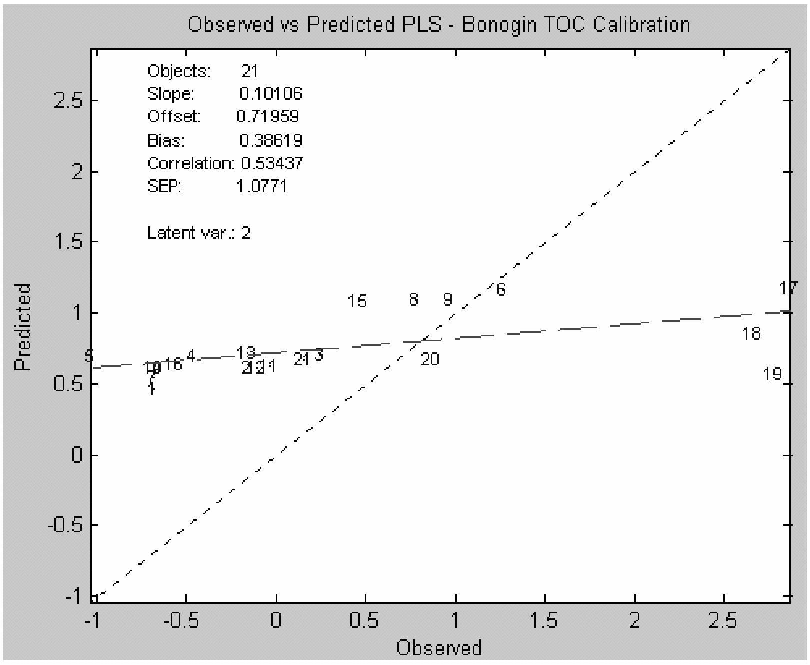
The resulting PLS analysis for the Bonogin data indicated that two significant factors out of four variables used were required to model the predictions for TOC, as indicated by the predicted error plots shown in Figure 6. Due to the number of variables being used to predict TOC, the model was expected to perform well. However the resulting Observed vs Predicted Calibration plot shown in Figure 7 performed a little poorer than expected, as indicated by an $r^2$ value of 0.53. This is most likely due to the fact that the selected variables are not strongly correlated. The calibration model was highly biased towards over predicting the values, and this is indicated in the resulting errors of fit.

These indicate that most of the X data variance was extracted ($R^2X$ = 81.2%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$ = 99%) also indicated good predictability using cross validation. The $R^2Y$ value (amount of Y variance explained), although still poor (44%), provided much more data variance extraction than the previous model. Unfortunately, this is not adequate to suggest a suitable model. Similarly, the resulting error of prediction (RMSEP = 5.1), is very high indicating that substantial errors exist in the prediction model. Due to the calibration only providing a mediocre performance, a validation was not undertaken.




n = 2

**Figure 6 – PLS analysis error plots for TOC for Bonogin Catchment**




$R^2X$=81.242%
$R^2Y$=44.232%
$Q^2$=99.06%
RMSEP=5.1317

**Figure 7 – TOC calibration plot for Bonogin Catchment**

*PLS – TP regressed* (excluding pH and EC)

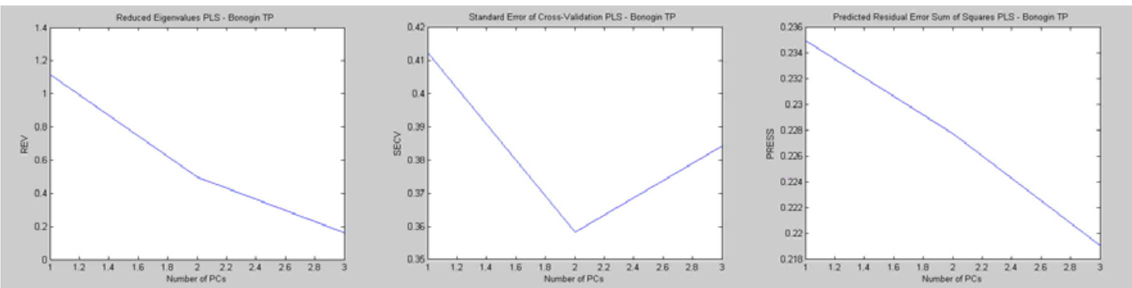
Calibration = 21 objects

Validation = 21 objects

PLS was performed on the data which was separated as:

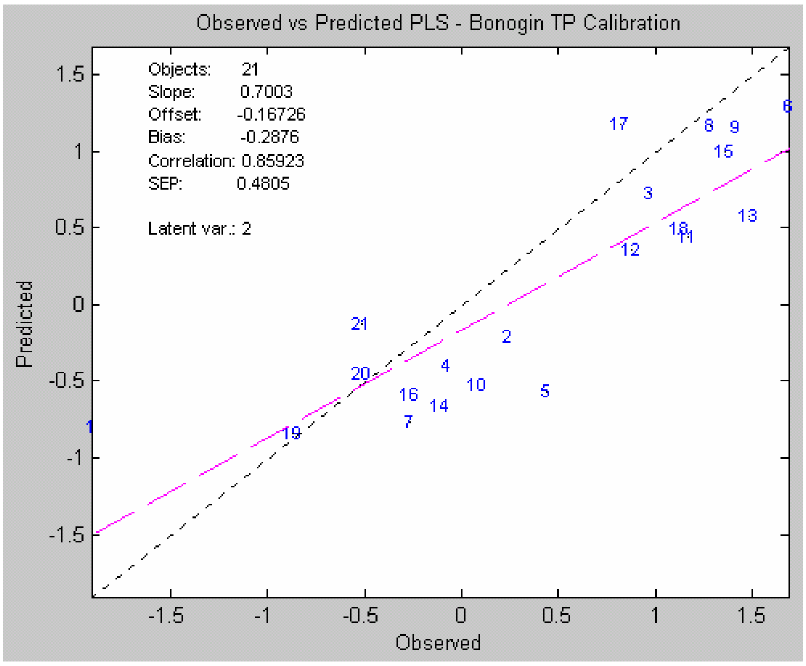
- Y = TP, X = TOC, TN and SS
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the resulting analysis for the TP, the PLS model indicated that two significant factors out of three variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 8. The resulting Observed vs Predicted Calibration and Validation plots shown in Figures 9 and 10 performed well, as indicated by $r^2$ values of 0.86 and 0.92 respectively. This is a suitable model. The reason for this is due to the variables being more strongly correlated. Neither the calibration or validation model were very biased, although a some bias is evident in the validation model, although this is quite low. The errors of fit also indicate that the model is predicting very well. Most of the X data variance was extracted ($R^2X$ = 94%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$ = 99%) also indicated good predictability using cross validation. The $R^2Y$ value (amount of Y variance explained), was very good (71%), indicating that the prediction of TP values is good. The resulting error of prediction (RMSEP = 1.8), indicates that the errors in the prediction model are reduced compared to the previous model.




n = 2

**Figure 8 – PLS analysis error plots for TP for Bonogin Catchment**




$R^2X$=93.933%
$R^2Y$=71.005%
$Q^2$=99.258%
RMSEP=1.7825

**Figure 9 – TP calibration plot for Bonogin Catchment**

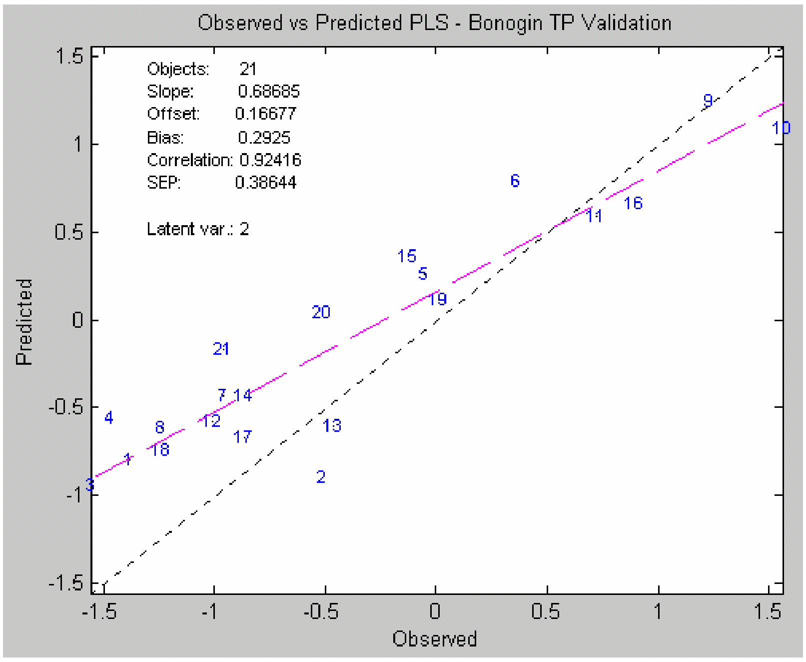



**Figure 10 – TP validation plot for Bonogin Catchment**

*PLS – TN regressed* (excluding pH and EC)

Calibration = 21 objects

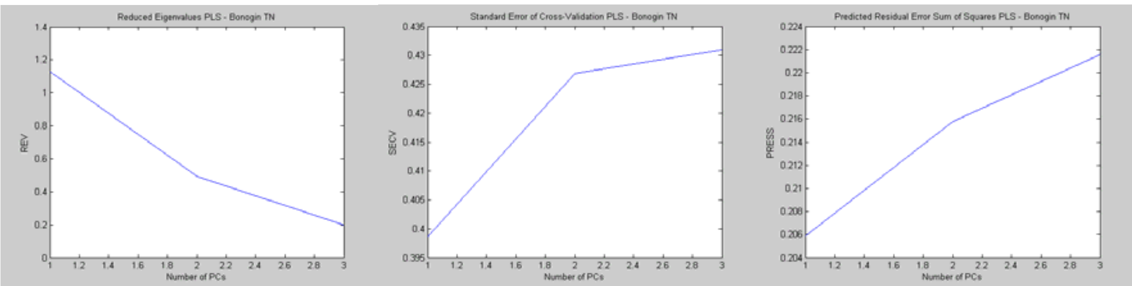
Validation = 21 objects

PLS was performed on the data which was separated as:

- Y = TN, X = TOC, TP and SS

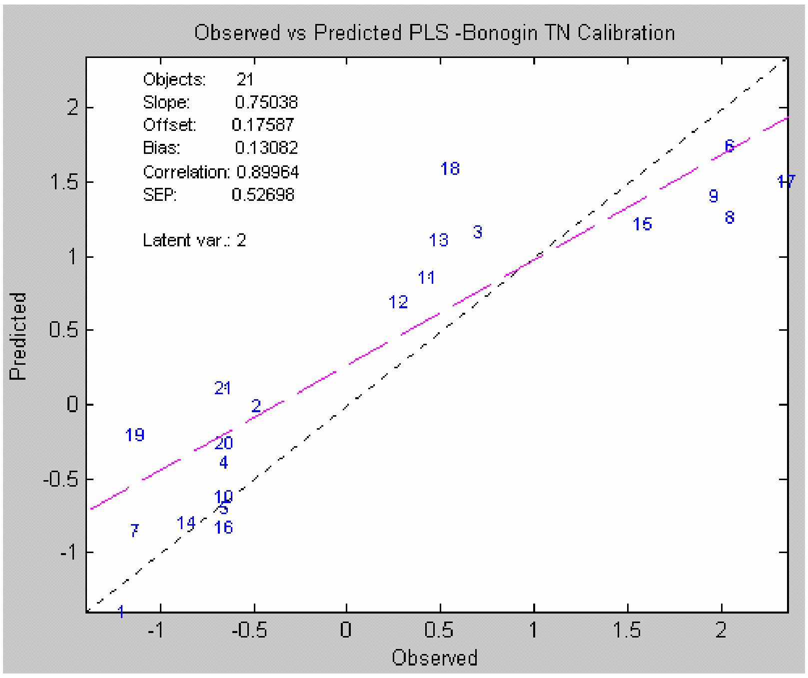
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the resulting analysis for TN, the PLS model indicated that three significant factors out of three variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 11. The resulting Observed vs Predicted Calibration and Validation plots shown in Figure 12 and 13 indicated that the PLS model also performed well, as indicated by $r^2$ values of only 0.89 and 0.90 respectively. Similar to the previous model, the main reason for the good model performance is due to the variables used for prediction being more closely correlated. Neither the calibration or validation model retained much bias towards either the predicted or observed values. The errors of fit also indicated that the model is predicting very well. Most of the X data variance was extracted ($R^2X$ = 98%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation Q2 = 99%) also indicated good predictability using cross validation. The $R^2Y$ value (amount of Y variance explained), was very good (81%), even higher than for the TP model. The resulting error of prediction (RMSEP = 1.4), was also lower, indicating that the errors in the prediction model are minimal.




n = 3

**Figure 11 – PLS analysis error plots for TN for Bonogin Catchment**




$R^2X$=97.8%
$R^2Y$=80.538%
$Q^2$=99.527%
RMSEP=1.404

**Figure 12 – TN calibration plot for Bonogin Catchment**

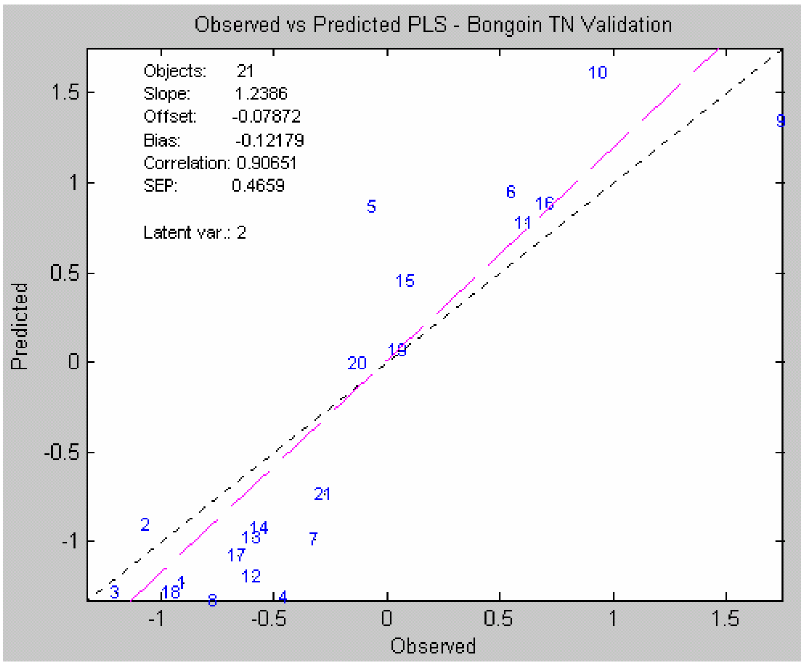



**Figure 8 – TN validation plot for Bonogin Catchment**

### 5.2.2 Hardy Catchment

**PCA Analysis**

PCA analysis for Hardy EMC data involved six variables and twenty four objects, BiE01 - BiE24 (ie 24x6 data matrix). One outlier was found and removed, providing a 23x6 data matrix for analysis. From the analysis and the corresponding Scree plot shown in Figure 14, it was determined that the first three PCs were significant as they account for 81.5% of the data variance. Consequently, PC1, PC2 and PC3 were investigated in the PCA analysis resulting in 3D plots. The 3D Scores plot and the Biplot developed are shown in Figures 15 and 16 respectively.

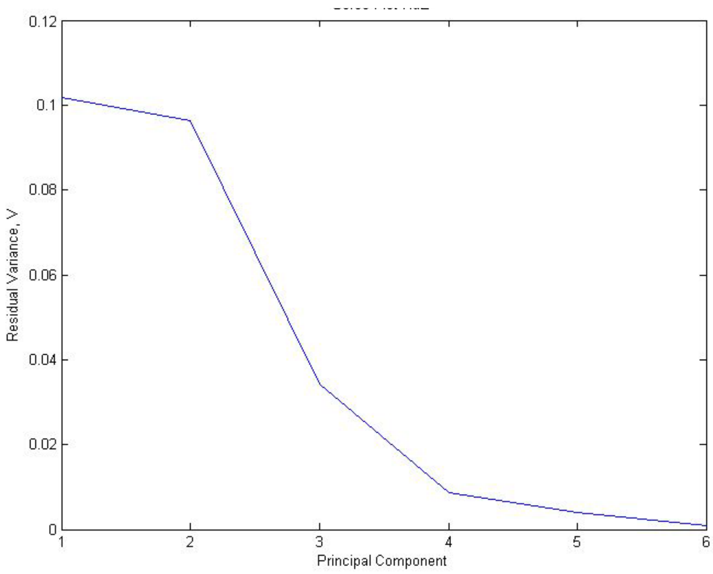



**Figure 14 – Scree Plot for Hardy Catchment**

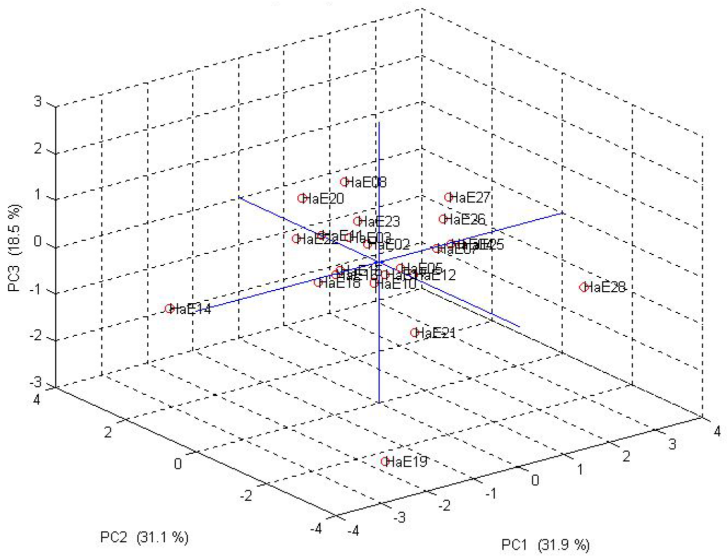



**Figure 15 – 3D Scores plot for Hardy Catchment**

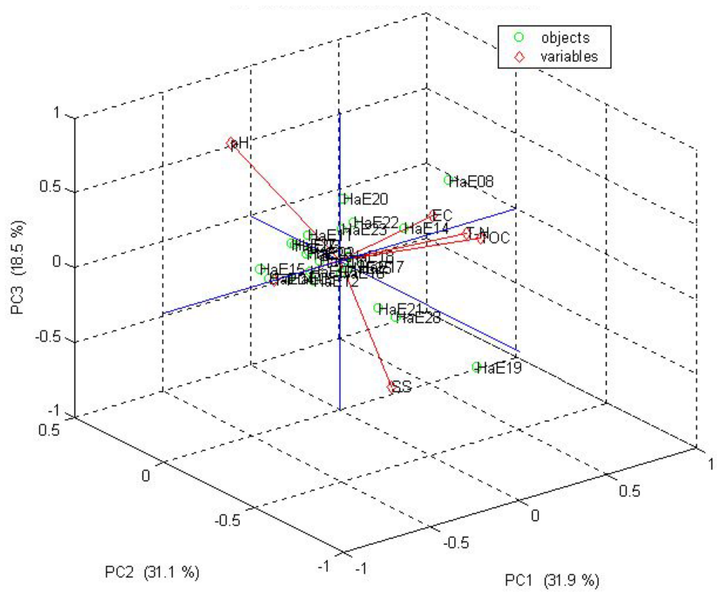



**Figure 16 – 3D Biplot for Hardy Catchment**

The interpretation of the 3D Biplot can be visually confusing. Hence to facilitate interpretation, the Biplot has been re-plotted in the three planes individually as shown in Figures 17 – 19.

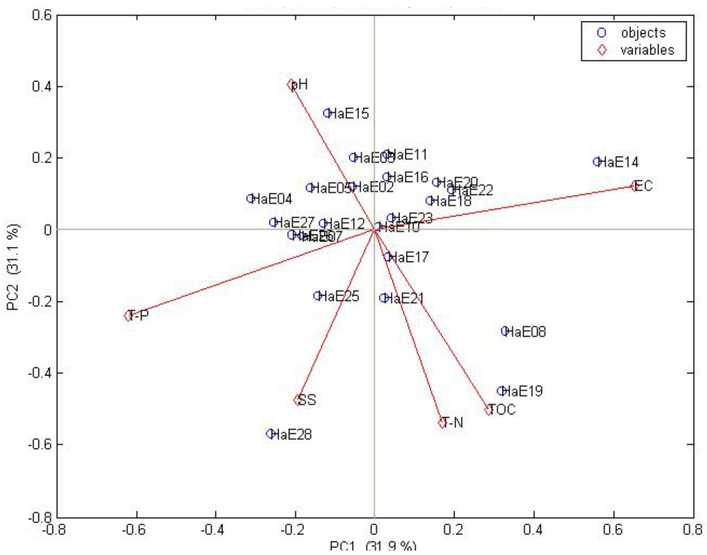



**Figure 17 – Biplot for Hardy Catchment (axes 1 and 2)**

Based on the Scores and the Biplot the following conclusions can be derived:

- Both pH and EC are not correlated with the other variables.

- TN and TOC are strongly correlated with each other, but are not correlated with TP. This would mean that TN is in organic form whilst TP is primarily in inorganic form.

- TN and TOC and also TP show some correlation with SS, thus indicating that a fraction of these parameters are in solid form. However, as SS and TP both fall negatively on PC1 and PC2, it would indicate that TP is primarily in solid form, whilst TOC and TN are primarily in dissolved form. However, due to the minor correlation with SS, TN and TOC are most probably in equal fractions of dissolved and solid forms.

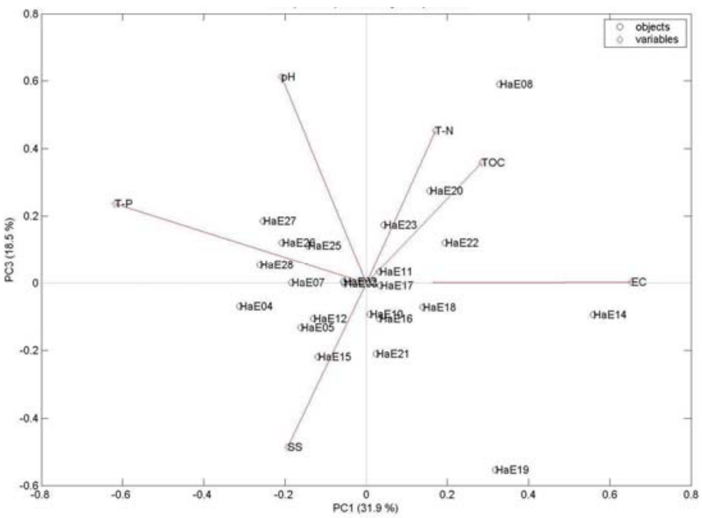



**Figure 18 – Biplot for Hardy Catchment (axes 1 and 3)**

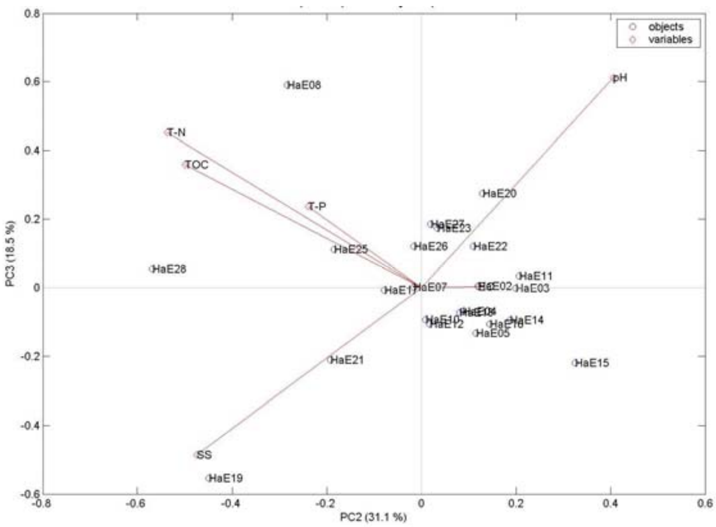



**Figure 19 – Biplot for Hardy Catchment (axes 2 and 3)**

- The fact that TOC is partially in dissolved form, the comments made in relation to DOC for the Bonogin catchment are also relevant here.

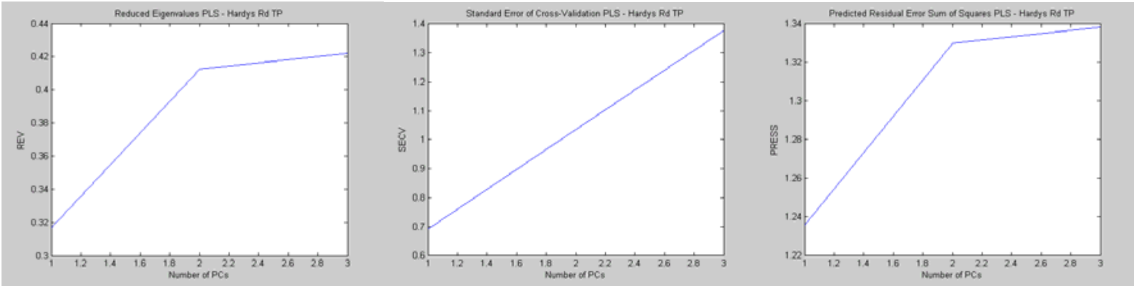
- The fact that a significant fraction TN and TOC are in soluble form and TP is in solid form would mean that conventional structural pollutant abatement measures such as and sediment traps will be effective only in the removal of SS and TP. Structural measures would only be partially effective in the case of TN and TOC.

*PLS – TP regressed* (excludes pH TOC and SS)
Calibration = 54 objects
Validation = 54 objects

PLS was performed on the data which was separated as:

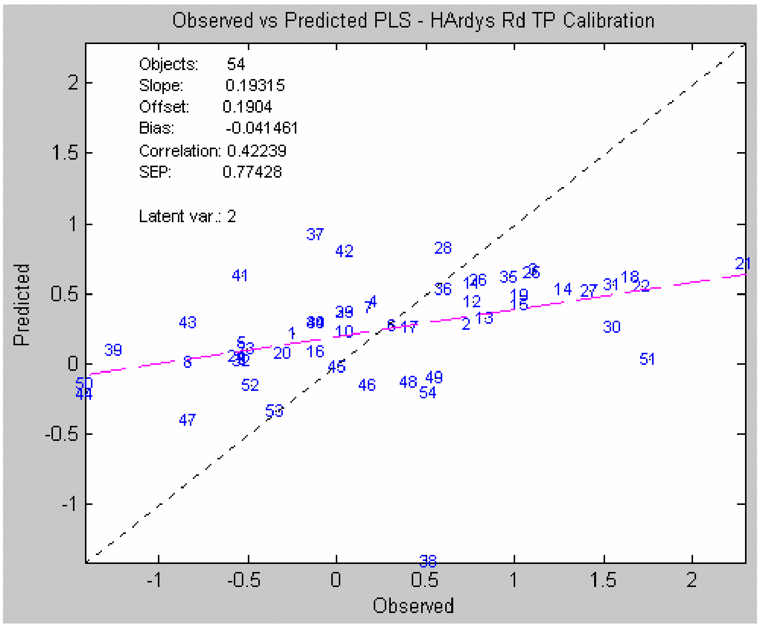
- Y = TP, X = TN and EC
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TP, the PLS model indicated that two significant factors out of two variables used were required to model the predictions as indicated by error plots shown in Figure 20. However, the error prediction plots also suggested that the error will continually increase, indicating that the model will perform rather poorly. The corresponding Observed vs Predicted Calibration and Validation plots shown in Figures 21 and 22 confirmed that the model is weak, with $r^2$ values of 0.42 and 0.17 respectively. The main reason for this, is related to the correlation between the variables selected, and the fact that only two variables can be used. Only minor bias was observed in the validation model, mostly towards under prediction of the values, primarily due to the cluster of points located towards the observed axis. The errors of fit also indicated that the model is predicting poorly. Most of the X data variance when extracted ($R^2X$ = 97%) in the models, and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=88%) indicated good predictability using cross validation. However, this reduced slightly compared to other models, mostly probably due to the arrangement of values for X-Y data sets. The $R^2Y$ value (amount of Y variance explained), was poor (27%), indicating that the prediction of TP values is not suitable. Similarly, the resulting error of prediction (RMSEP = 3.2) also indicated that errors in the prediction model evident were becoming more prominent.




n = 2

**Figure 20 – PLS analysis error plots for TP for Hardy Catchment**




$R^2X$=96.761%
$R^2Y$=27.035%
$Q^2$=88.19%
RMSEP=3.2592

**Figure 21 – TP calibration plot for Hardy Catchment**

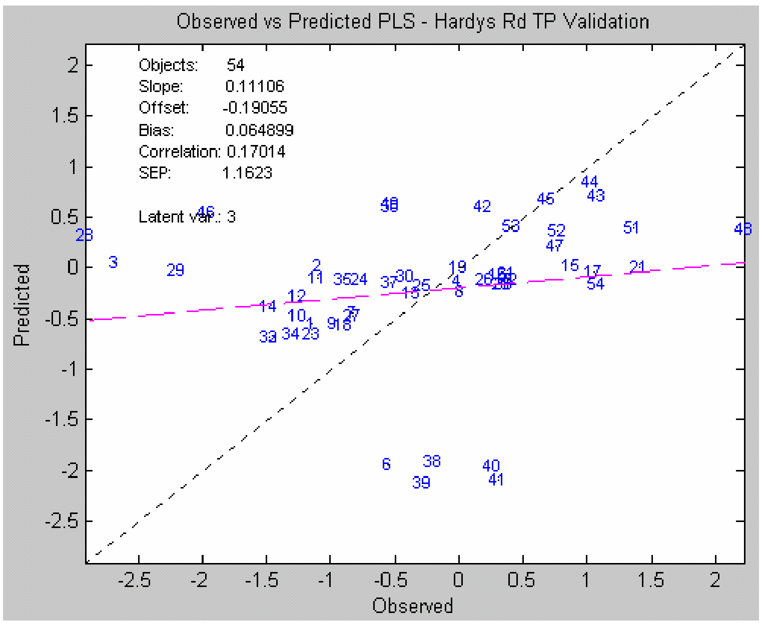



**Figure 22 – TP validation plot for Hardy Catchment**

*PLS – TN regressed* (excludes pH TOC and SS)

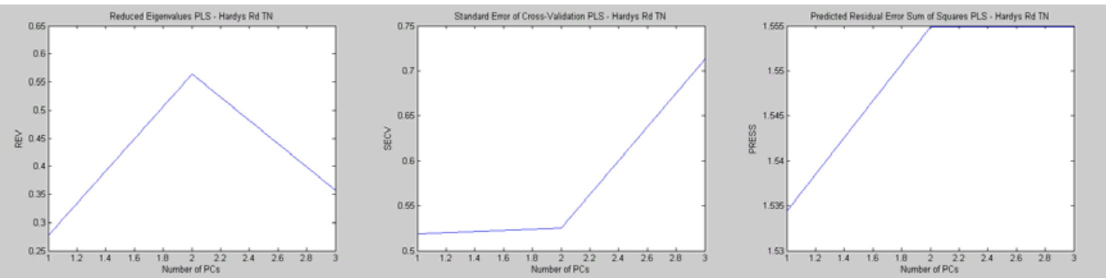
Calibration = 54 objects

Validation = 54 objects

PLS was performed on the data which was separated as:

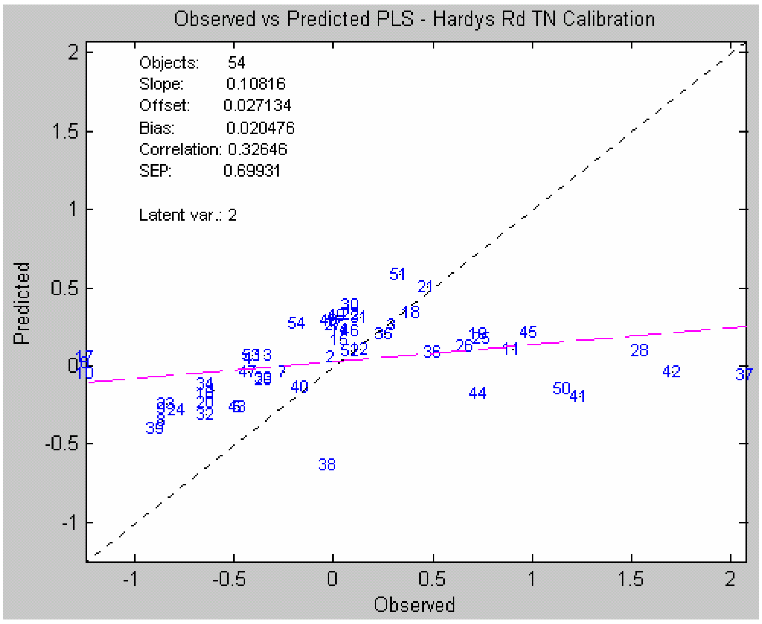
- Y = TN, X = TP and EC
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TN, the PLS model indicated that two significant factors out of two variables used were required to model the predictions for TP. However, error prediction plots shown in Figure 23 also suggested that the errors will mostly increase, indicating that the model will perform poorly. Corresponding Observed vs Predicted Calibration and Validation plots shown in Figures 24 and 25 confirmed that the model is weak, with $r^2$ values of 0.32 and -0.1 respectively. The main reason for this is related to correlation between the variables selected, and the fact that only two variables can be used. Both the calibration and validation models retained bias towards the observed values (under predicting) with the validation model retaining the most bias. Calibration model only retained minor bias. Errors of fit also indicated that the model is predicting poorly. Most of the X data variance when extracted ($R^2X$ = 84%) in the models, and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=75%) indicated suitable predictability using cross validation. However, this reduced slightly compared to other models, probably due to the arrangement of values for X-Y data sets. The $R^2Y$ value (amount of Y variance explained), was very poor (11%), indicating that the prediction of TN values is not suitable. Also the resulting error of prediction (RMSEP = 3.6), indicated that errors in the prediction model evident were becoming more prominent.




n = 2

**Figure 23 – PLS analysis error plots for TN for Hardy Catchment**




$R^2X$=84.533%
$R^2Y$=11.113%
$Q^2$=75.219%
RMSEP=3.6015

**Figure 24 – TN calibration plot for Hardy Catchment**

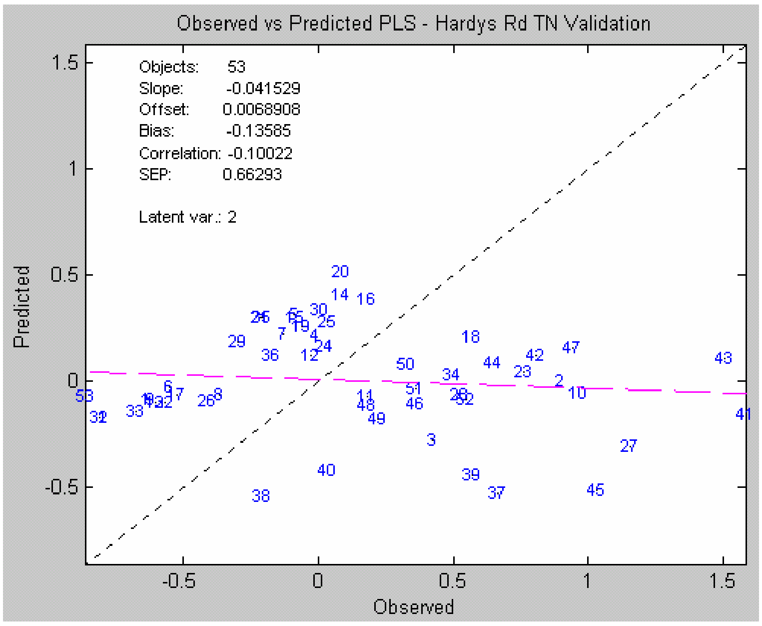



**Figure 25 – TN validation plot for Hardy Catchment**

### 5.2.3 Hinkler Catchment

**PCA Analysis**

The analysis of Hinkler EMC data involved six variables; pH, EC, SS, TOC, TP and TN and thirty three objects, BiE01 - BiE33 (ie 33x6 data matrix). One outlier was discovered and removed, providing a 32x6 data matrix for analysis. From the analysis and the corresponding Scree plot as shown in Figure 26, it was determined that the first three PCs were significant, contributing 65.5% of the data variance. Consequently, PC1, PC2 and PC3 were investigated in the PCA analysis. The resulting 3D Scores plot and Biplot are shown in Figures 27 and 28 respectively.

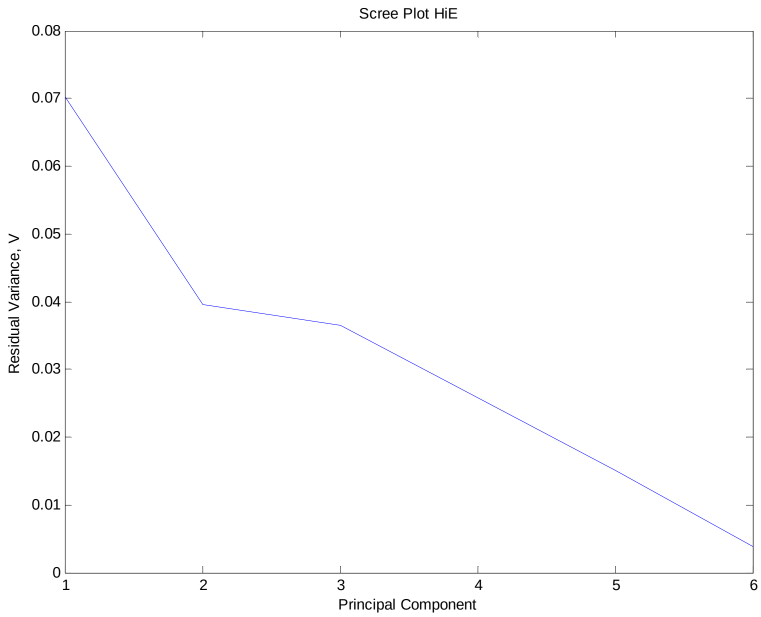



**Figure 26 – Scree plot for Hinkler Catchment**

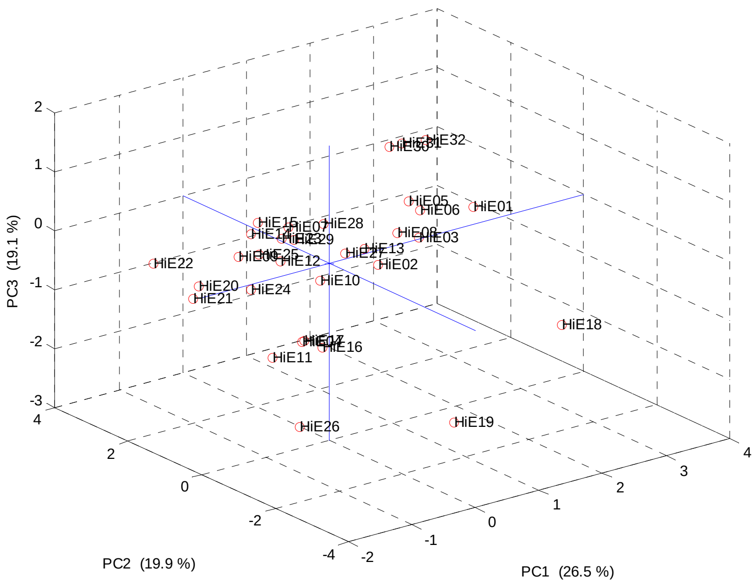



**Figure 27 – 3D Scores plot for Hinkler Catchment**

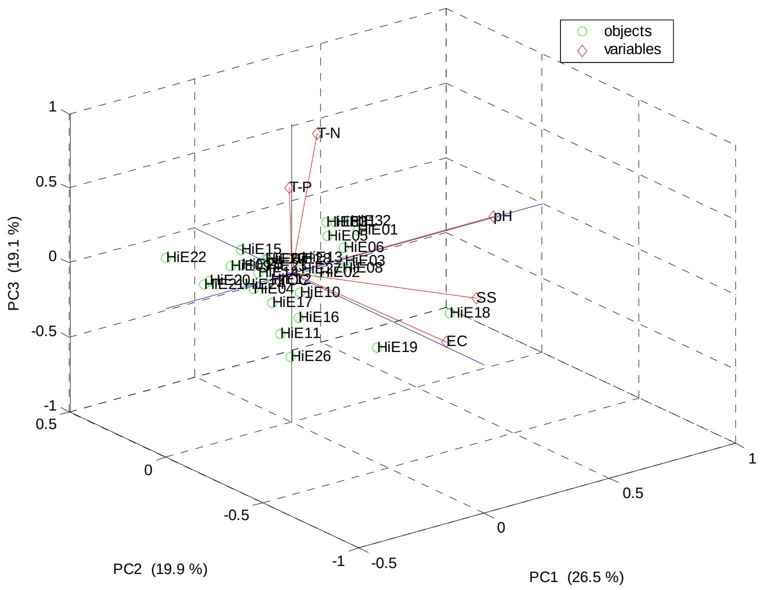



**Figure 28 – 3D Biplot plot for Hinkler Catchment**

To facilitate the interpretation of the Biplot, the three planes have been re-plotted individually in Figures 29 – 31.

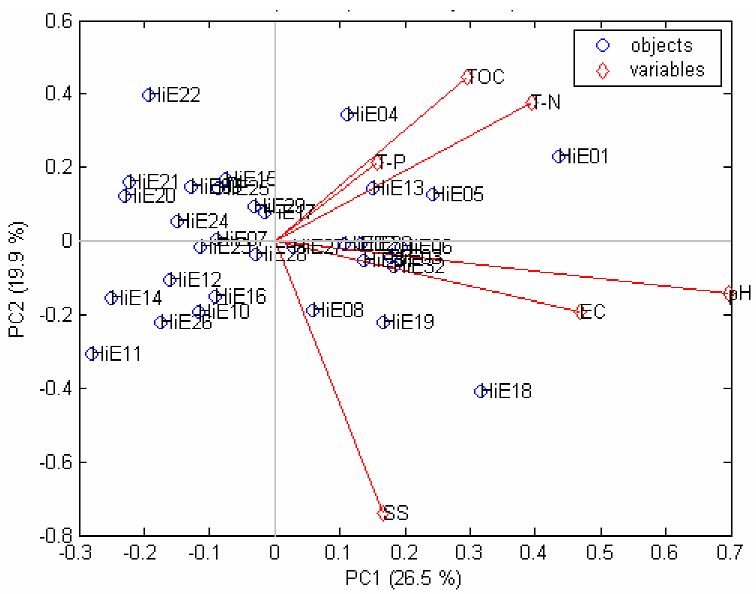



**Figure 29 – Biplot for Hinkler Catchment (axes 1 and 2)**

Based on the Scores and Biplot, the following conclusions can be derived:

- pH and EC are correlated with each other but only weakly correlated with the other variables.

- TP, TN and TOC are closely correlated with each other but uncorrelated with SS. Hence these parameters would be primarily in dissolved form.

- The fact that TOC is primarily in dissolved form, or as DOC and the comments made in relation to the Bonogin Catchment are also relevant here.

- The fact that TP, TN and TOC are in dissolved form would mean that structural pollutant abatement measures such as sediment traps will only be effective in the removal of SS.

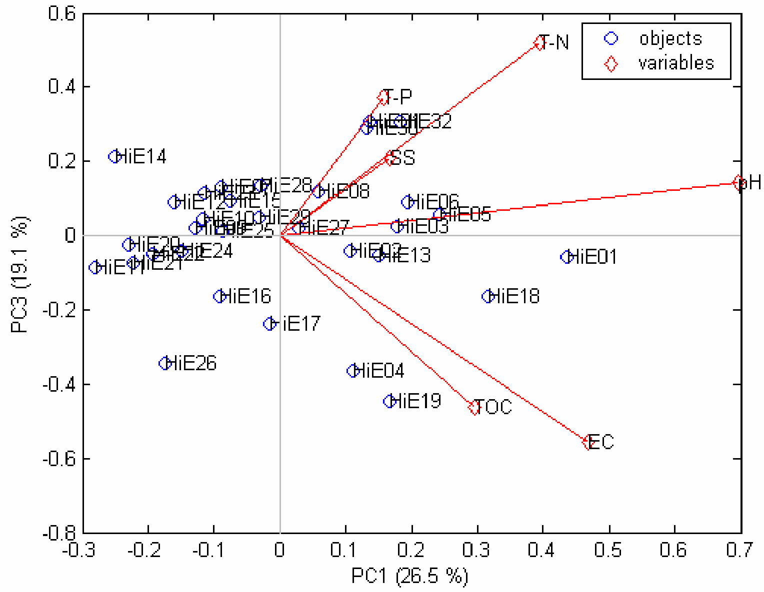



**Figure 30 – Biplot for Hinkler Catchment (axes 1 and 3)**

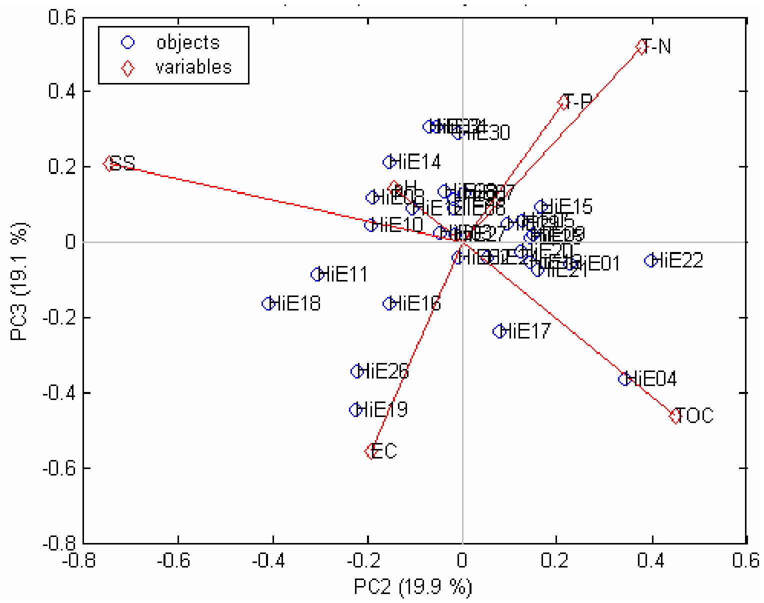



**Figure 31 – Biplot for Hinkler Catchment (axes 2 and 3)**

**PLS Regression**

*PLS – TP regressed* (excludes pH and EC)

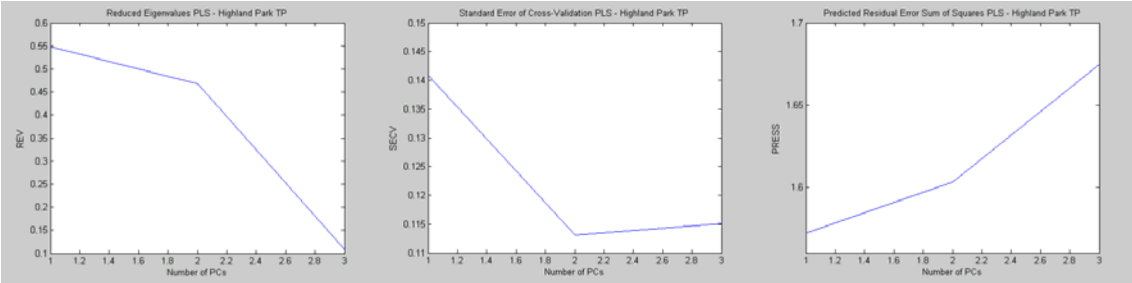
Calibration = 49 objects

Validation = 49 objects

PLS was performed on the data which was separated as:

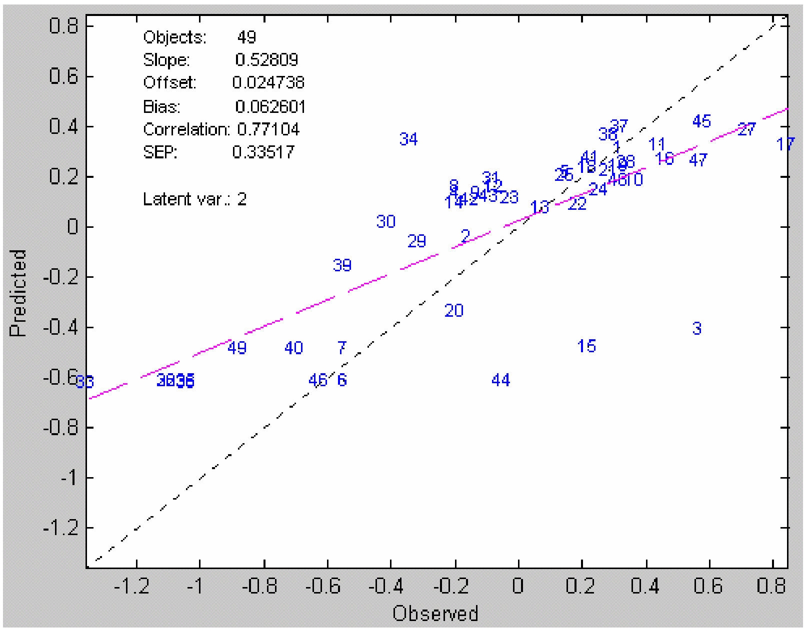
- Y = TP, X = TN , TOC and SS
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TP, the PLS model indicated that two significant factors out of three variables used were required to model the predictions as indicated by the predicted error plots shown in Figure 32. The corresponding Observed vs Predicted Calibration plot shown in Figure 33 indicated that the model provides a good prediction, with $r^2$ value of 0.77. However, the validation model as shown in Figure 34 did not perform as well ($r^2$ =0.1). The main reason for this is most probably due to the splitting of the data into X and Y data, with some inconsistencies (ie higher or lower values) in one of the data sets. Minimal bias was observed for the calibration model. However the validation model was found to be highly biased towards under predicting values (biased to the observed values). The errors of fit also indicated that the model is predicting poorly. Most of the X data variance was again extracted ($R^2X$ = 90%) in the models, however the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$ = 62%) did not perform as well with a much lower % variance obtained using cross validation. The $R^2Y$ value (amount of Y variance explained), provided an improved variance (62%), although not appreciably strong. This suggests that the prediction of TP values although not very precise, may still be acceptable using the calibration model. The resulting error of prediction (RMSEP = 1.7), also indicated that errors in the prediction model are evident, although they are still minimal.




n = 2

**Figure 32 – PLS analysis error plots for TP for Hinkler Catchment**




$R^2X$=89.99%
$R^2Y$=62.378%
$Q^2$=41.333%
RMSEP=1.671

**Figure 33 – TP calibration plot for Hinkler Catchment**

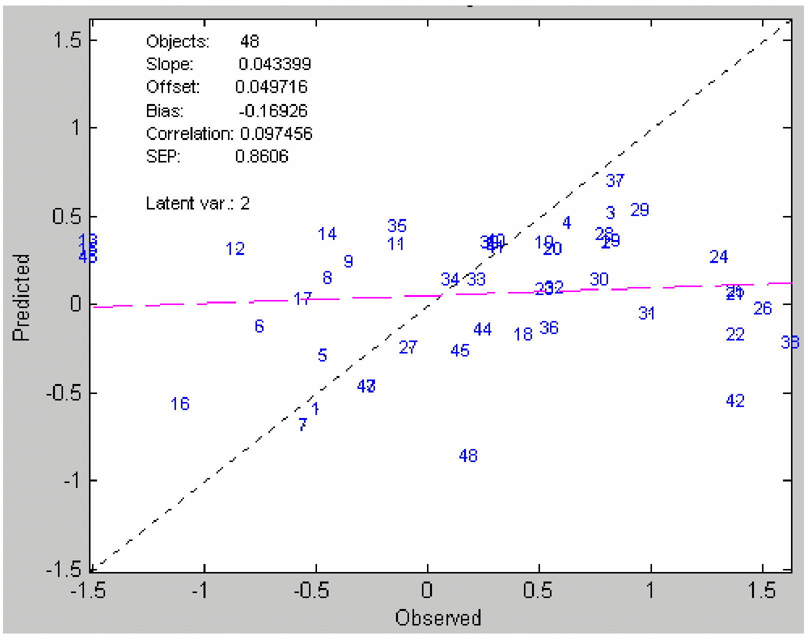



**Figure 34 – TP validation plot for Hinkler Catchment**

*PLS – TOC regressed* (excludes pH) No change equally correlated

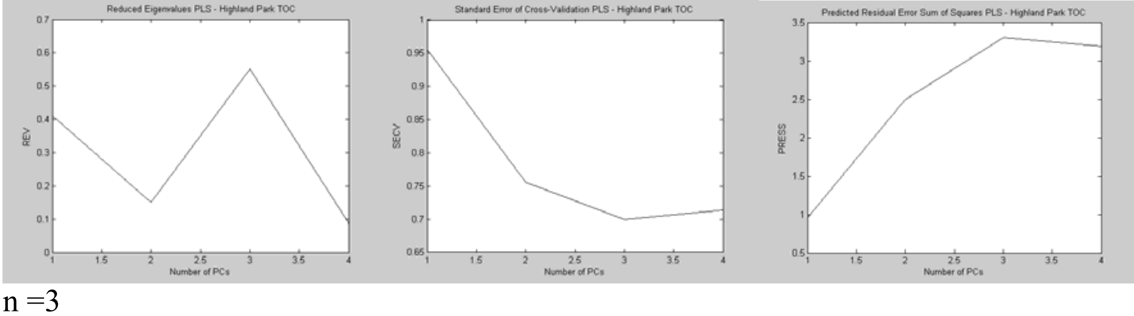
Calibration = 49objects

Validation = 48 objects

PLS was performed on the data which was separated as:

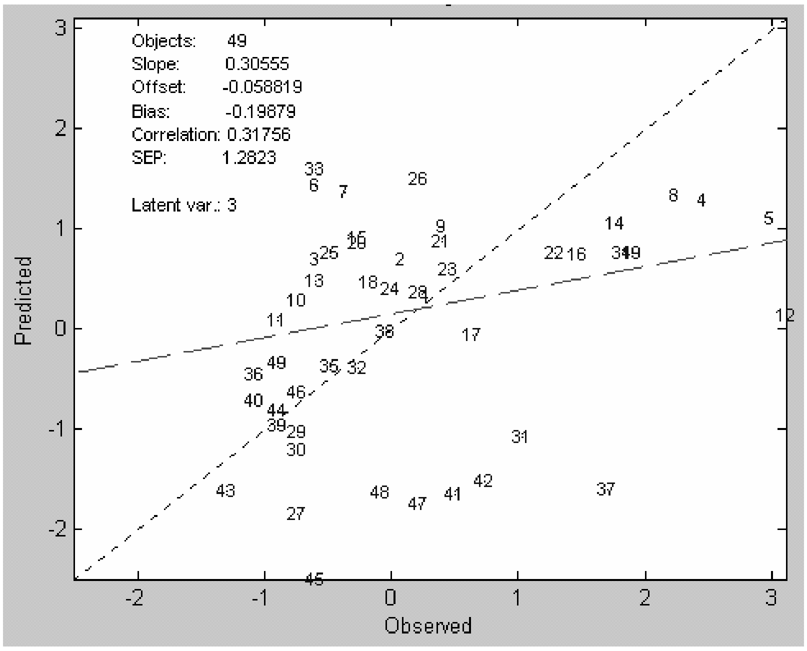
- Y = TOC, X = TN, TP, SS and EC
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TOC, the PLS model indicated that three significant factors out of four variables used were required to model the predictions as indicated by the predicted error plots shown in Figure 35. The corresponding Observed vs Predicted Calibration and Validation plots shown in Figures 36 ad 37 confirmed that the model performance was weak with $r^2$ values of only 0.31 and 0.1 respectively. The main reason for this is related to the correlation between the variables selected, in particular, TOC not being strongly correlated with the other variables. Both the calibration and validation models are highly biased towards the observed data, thereby under predicting the TOC values. The errors of fit also indicate that the model is predicting poorly. The amount of X data variance extracted ($R^2X$ = 74%) in the models was slightly less than usual. However, the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$ = 54%) provided much less data variance for using the cross validation method. Consequently, the model is obviously performing poorly as a result of these low values. The $R^2Y$ value (amount of Y variance explained), was poor (54%), indicating that the prediction of TOC values, although higher than previous, is only a slight improvement from predicting TP values. The resulting error of prediction (RMSEP = 5.2), also indicated that errors in the prediction model are very high indicating poor model performance.




n =3

**Figure 35 – PLS analysis error plots for TOC for Hinkler Catchment**




R2X=74.191
R2Y=50.888
Q2=54.509
RMSEP=5.1916

**Figure 36 – TOC calibration plot for Hinkler Catchment**

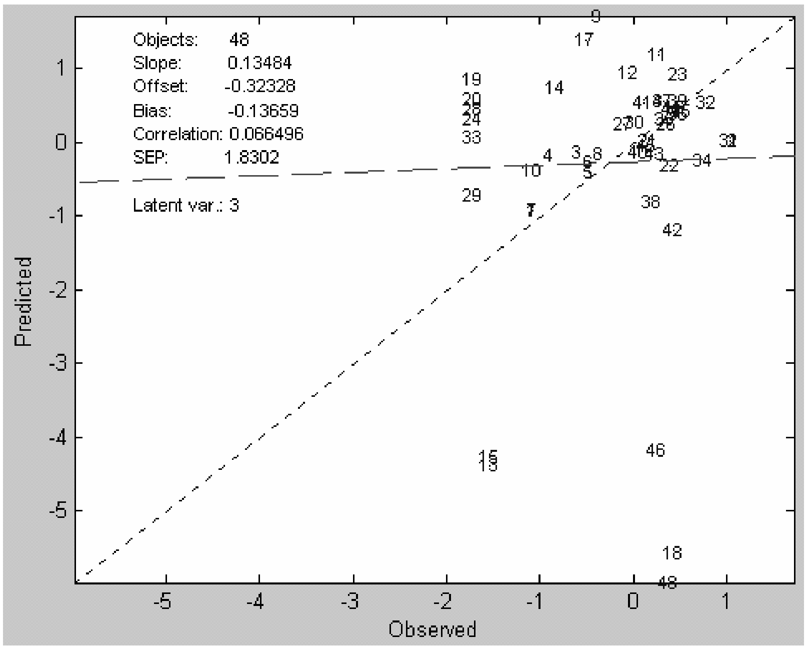



**Figure 37 – TP validation plot for Hinkler Catchment**

*PLS – TN regressed* (excludes pH and EC)

Calibration = 49 objects

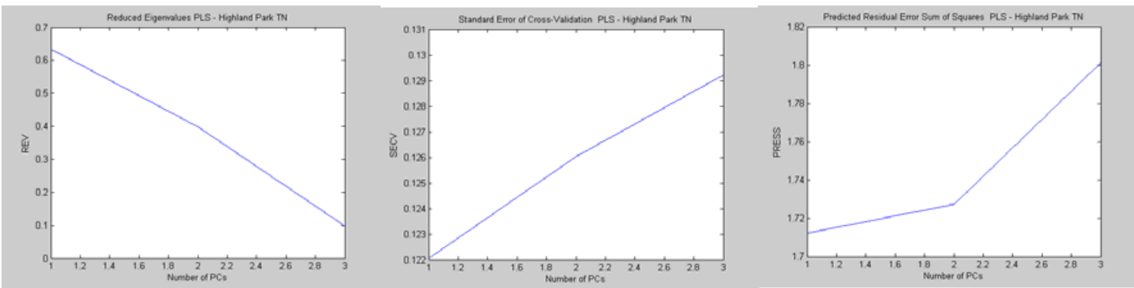
Validation = 48 objects

PLS was performed on the data which was separated as:

- Y = TN, X = TN, TP, SS

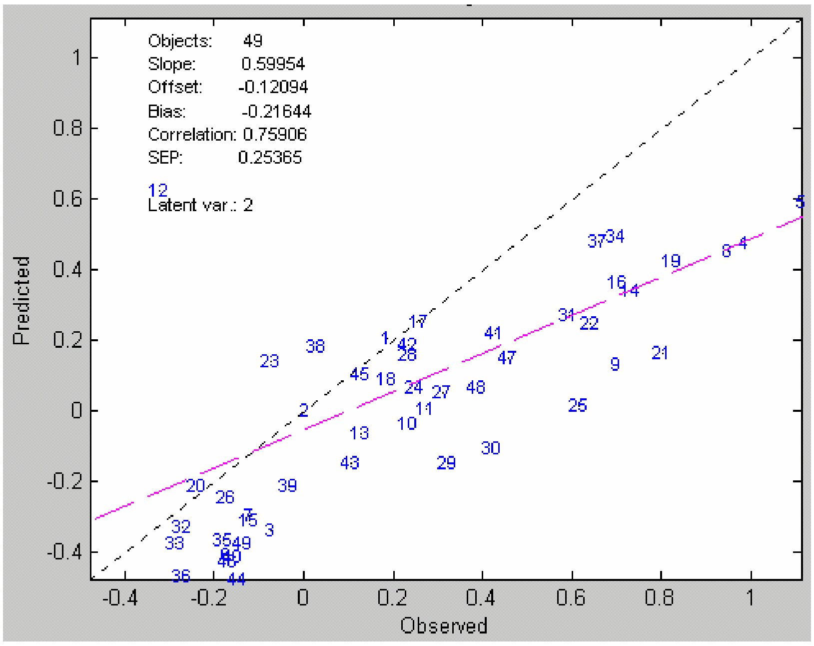
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TN, the PLS model indicated that three significant factors out of four variables used were required to model the predictions. However, the predicted errors as indicated by the predicted error plots shown in Figure 38 indicated that the errors in the models may continuously increase, suggesting that the model may perform poorly. The corresponding Observed vs Predicted Calibration plot shown in Figure 39 shows good model performance with a $r^2$ value of 0.76. However, the Validation plot given in Figure 40 indicated much poorer model performance ($r^2$ =0.03). As the variables used in this PLS model are highly correlated together, this is most likely due to distinct values in the X or Y validation data causing a poorer model performance. Both the calibration and validation models are biased towards the observed data, causing under prediction of the predicted TN values. The errors of fit also indicate that the model is predicting poorly. Most of the X data variance was again extracted ($R^2X$=97.7%) in the models. However the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=8.8%) did not perform as well with a very low % variance obtained using cross validation. The model is obviously performing more poorly as a result of these low values. The $R^2Y$ value (amount of Y variance explained), was on average 47%, indicating that the prediction of TN values, is only a slight improvement on the other models derived previously. The resulting error of prediction (RMSEP=1.3), also indicates that errors in the prediction model are fairly small, suggesting that the model performed well. However, due to the low $Q^2$ and mediocre $R^2Y$, and because of the bias in the model predictions, the model for predicting TN is not particularly suitable.




n=3

**Figure 38 – PLS analysis error plots for TN for Hinkler Catchment**




$R^2X$=97.713%
$R^2Y$=47.03%
$Q^2$=8.8893%
RMSEP=1.3396

**Figure 39 – TN calibration plot for Hinkler Catchment**

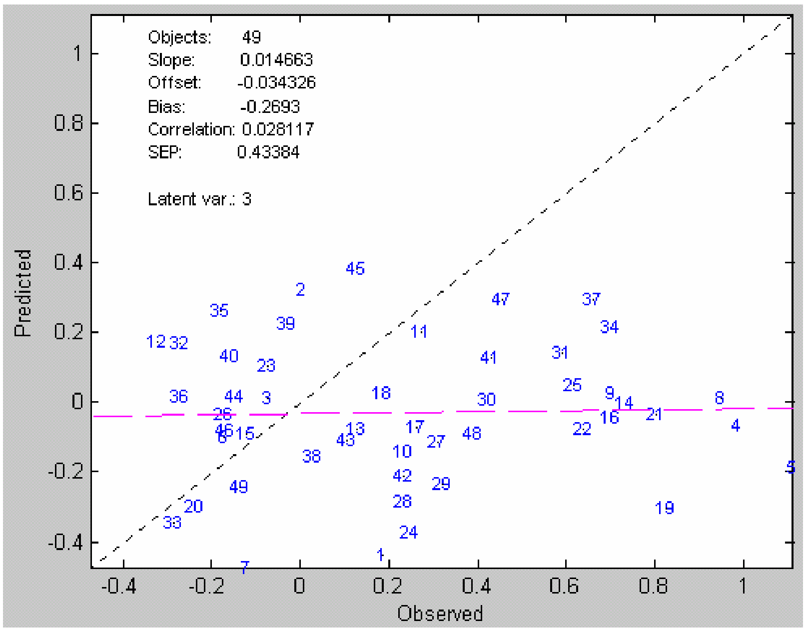



**Figure 22 – TN validation plot for Hinkler Catchment**

### 5.2.4 Alextown Catchment

The analysis of the Alextown EMC data involved six variables; pH, EC, SS, TOC, TP and TN and twenty one objects, AlE01 - AlE21 (ie 21x6 data matrix). One outlier was found, and removed providing a 20x6 data matrix. From this analysis and the corresponding Scree plot as shown in Figure 41, it was determined that the first three PCs were significant, contributing 77.8% of the data variance.

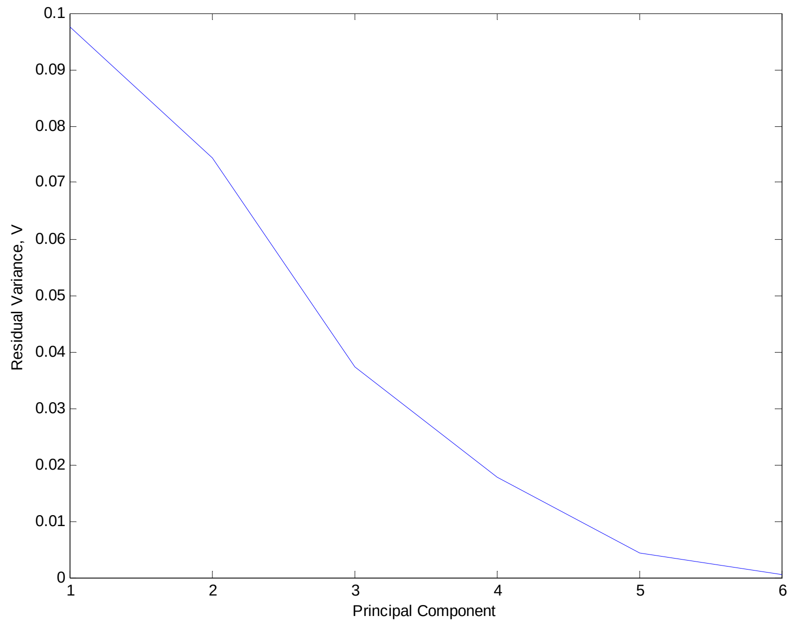



**Figure 41 – Scree Plot for Alextown Catchment**

The resulting Scores plot and Biplot are shown in Figures 42 and 43. To facilitate the interpretation of the 3D Biplot, the three planes have been re-plotted individually in Figures 44 – 46.

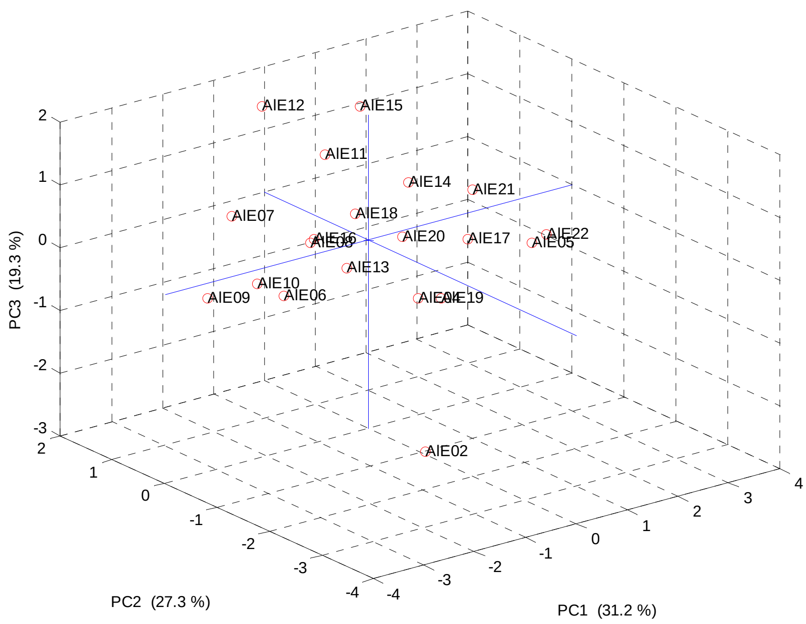



**Figure 42 – 3D Scores plot for Alextown Catchment**

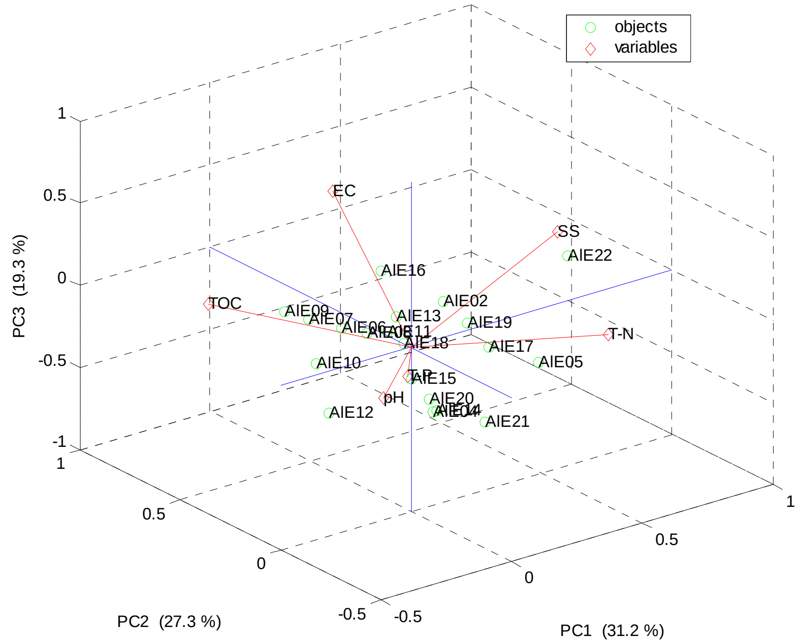



**Figure 43 – 3D Biplot for Alextown Catchment**

Based on the Scores plot and Biplot, the following conclusions can be derived:

- There are no appreciable correlation among any of the parameters other than between TN and SS.

- There is some correlation between TP and EC but there is no correlation with SS and hence TP would be in soluble form. This could mean that TP is primarily in its reactive form ($PO_4^+$) or one of the many possible salts it could form.

- The fact that TOC is not correlated with SS would mean that TOC is primarily in dissolved form and the comments made in relation to the Bonogin catchment are also relevant here.

- The fact that TOC and TP are not correlated with SS and is primarily in soluble form would mean that structural pollutant abatement measures such as sediment traps will only be effective in the removal of TN and SS.

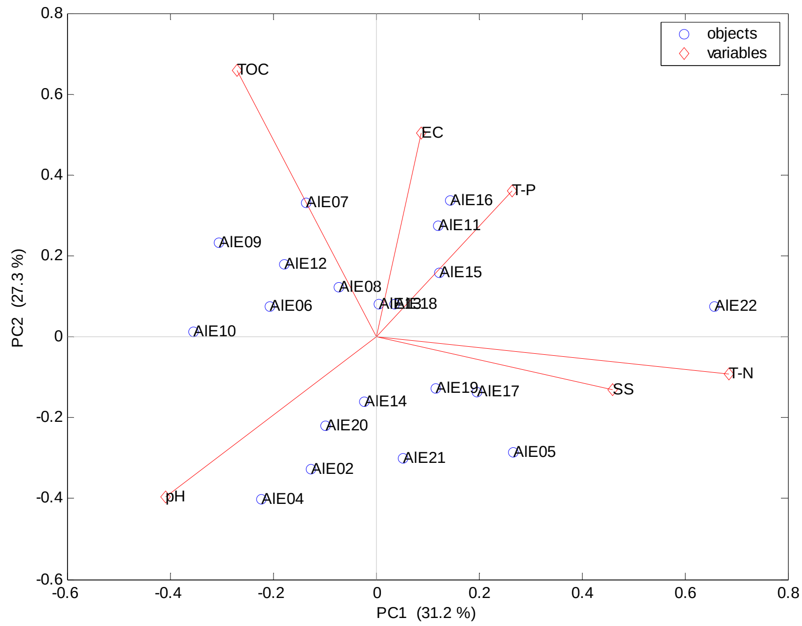



**Figure 44 – Biplot for Alextown Catchment (axes 1 and 2)**

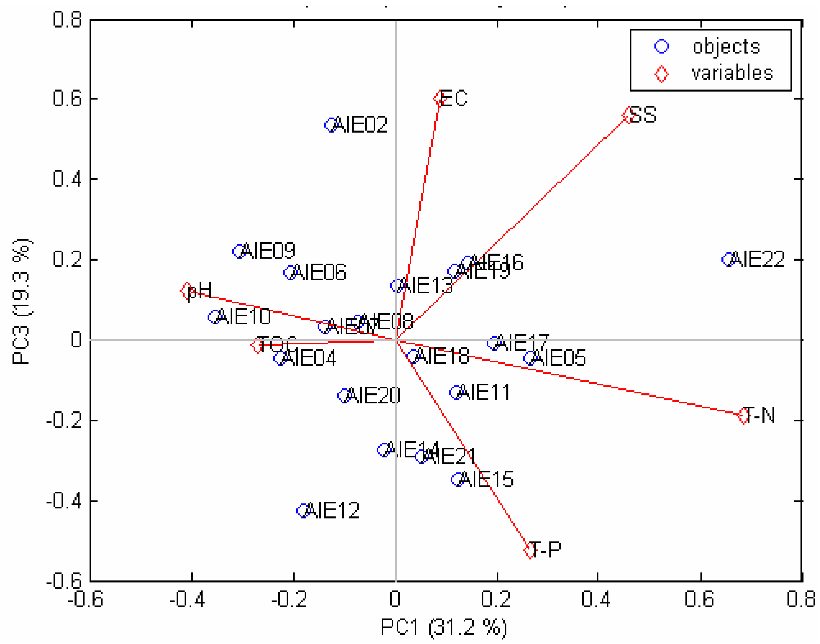



**Figure 45 – Biplot for Alextown Catchment (axes 1 and 3)**

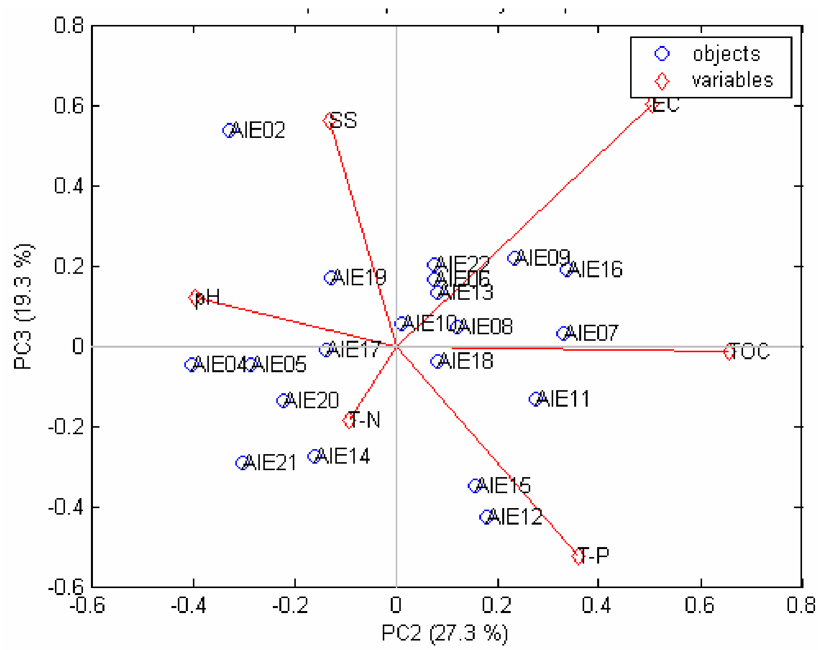



**Figure 46 – Biplot for Alextown Catchment (axes 2 and 3)**

**PLS Regression**

*Alextown PLS – TP regressed* (excluding pH and SS)

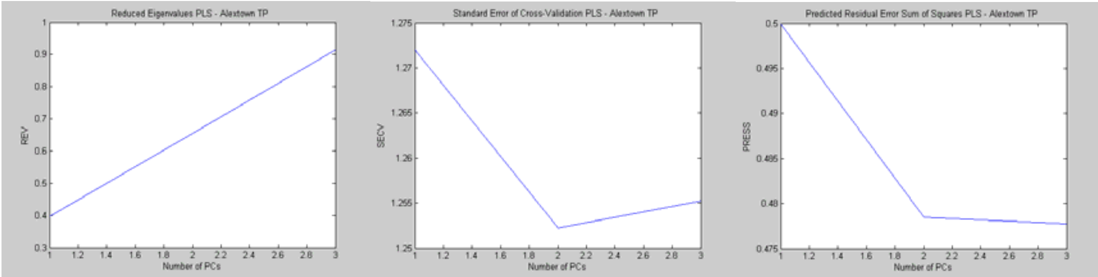
Calibration = 20 objects

Validation = 17 objects

PLS was performed on the data which was separated as:

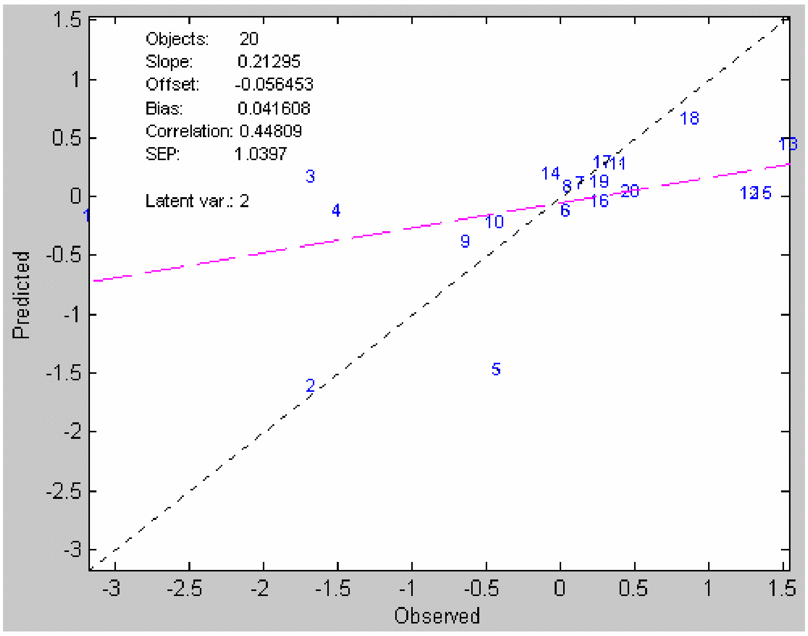
- Y = TP, X = EC, TN and TOC
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TP, the PLS model developed indicated that ttwo significant factors out of three variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 47. The corresponding Observed vs Predicted Calibration plot as shown in Figure 48 indicates that the model provided a poor fit for the data, with a $r^2$ value of 0.45. This is due to the calibration over fitting on the predicted results. The validation data as shown in Figure 49 was found to under predict the predicted values, with slightly more bias towards the observed. This is to be expected although EC and TOC are correlated, only a minor relationship with TN exists. The resulting errors of fit indicate that most of the X data variance was utilised ($R^2X$=88%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation) also indicated good predictability using cross validation. However, the $R^2Y$ value (amount of Y variance explained) was low (23%), indicating that the prediction was not acceptable. The resulting error of prediction was also high, indicating that errors in the prediction model were having an impact.




n = 2

**Figure 47 – PLS analysis error plots for TP for Alextown Catchment**




$R^2X$=87.71%
$R^2Y$=23.022%
$Q^2$=97.279%
RMSEP=3.2073

**Figure 48 – TP calibration plot for Alextown Catchment**

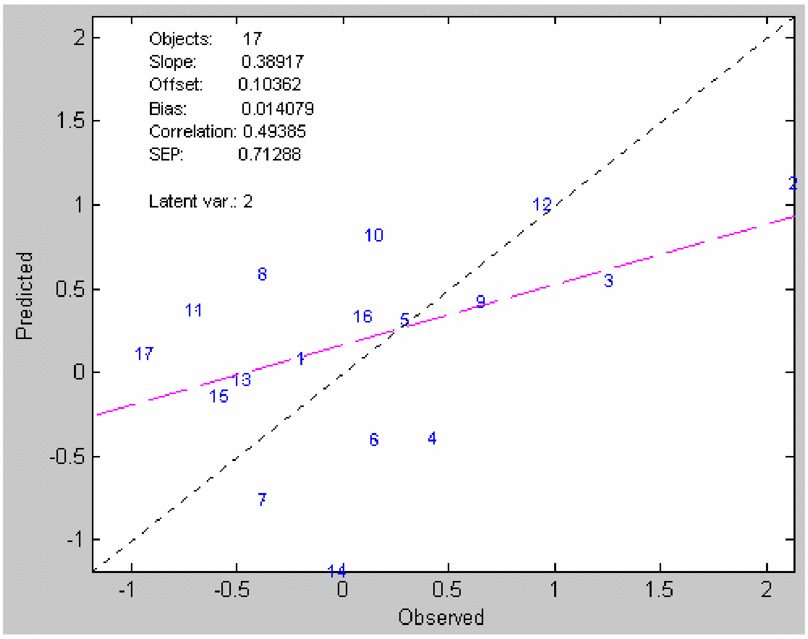



**Figure 49 – TP validation plot for Alextown Catchment**

*PLS – TOC regressed* (excluding pH, SS, TN)

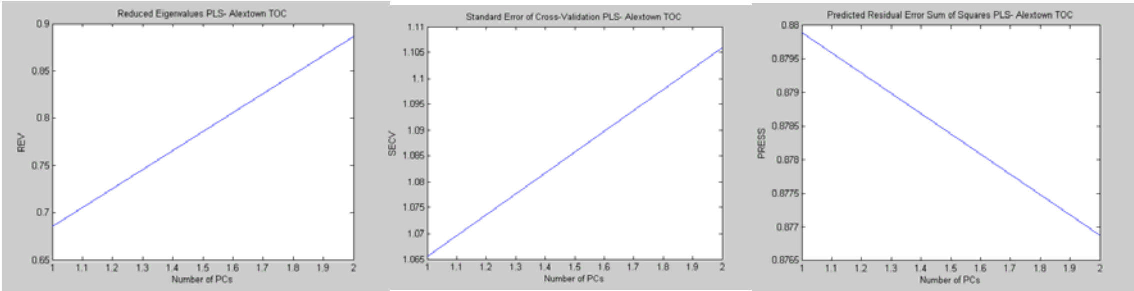
Calibration = 30 objects

Validation = 17 objects

PLS was performed on the data which was separated as:

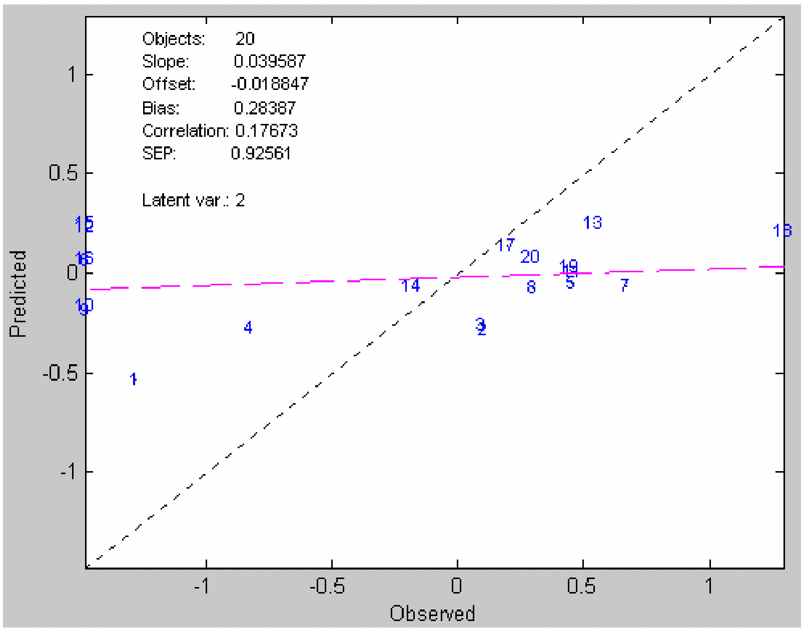
- Y = TOC, X = EC, TP
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the resulting analysis for TOC, the PLS model indicated that two significant factors out of the two variables used were required to model the predictions, as indicated in the predicted error plots shown in Figure 50. This is due to only two variables being used for prediction. The resulting Observed vs Predicted Calibration and Validation plots as given in Figures 51 and 52 shows that the model performed very poorly, with low $r^2$ values of 0.17 and 0.43 respectively. Substantial bias is also indicated in the calibration model for over predicting the values for TP. The validation data did not have much bias towards either over or under fitting the predicted values. However, due to the poor calibration model, it was only expected that the validation would not perform adequately. The resulting errors of fit indicated that most of the X data variance was utilised ($R^2X$=99.9%). The $Q^2$ (amount of variance explained in predicted Y via cross




n = 2

**Figure 50 – PLS analysis error plots for TOC for Alextown Catchment**




$R^2X$=99.96%
$R^2Y$=4.5955%
$Q^2$=91.771%
RMSEP=2.9908

**Figure 51 – TOC calibration plot for Alextown Catchment**

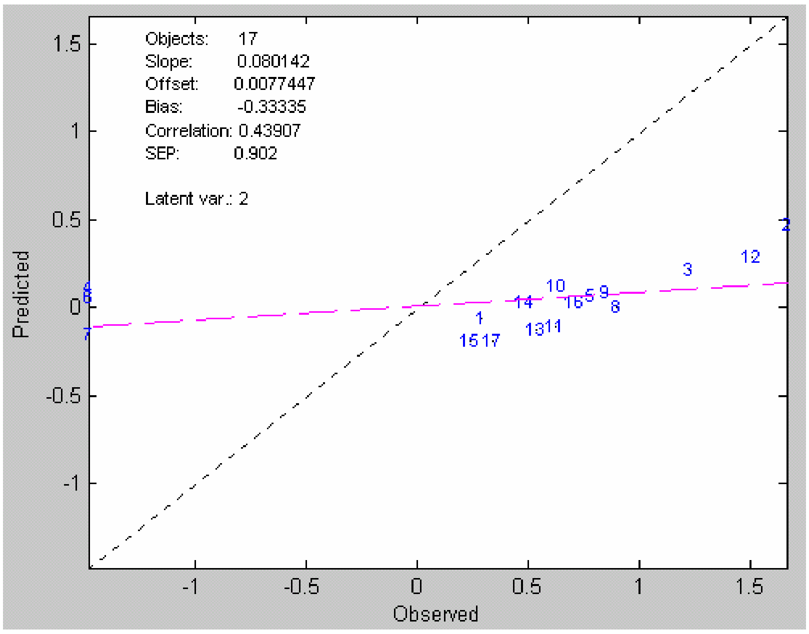



**Figure 52 – TOC validation plot for Alextown Catchment**

validation $Q^2$=91.7%) also indicated good predictability using cross validation. However, the $R^2Y$ value (amount of Y variance explained) was quite low (4.6%), indicating that the prediction was not acceptable. The resulting RMSEP was also high, highlighting poor model performance.

Analysis of TN was not undertaken as it is not closely correlated with other variables.

### 5.2.5 Gumbeel Catchment

**PCA Analysis**

The analysis for Gumbeel EMC data involved six variables; pH, EC, SS, TOC, TP and TN and nineteen objects, BiE01 - BiE19 (ie 19x6 data matrix). No outliers were found in the Gumbeel data. From the analysis and the corresponding Scree plot, as shown in Figure 53, it was determined that only the first two PCs were significant, contributing 71.4% of the data variance. Therefore PC1 and PC2 were investigated in the PCA analysis. The resulting Scores plot and Biplot are shown in Figures 54 and 55.

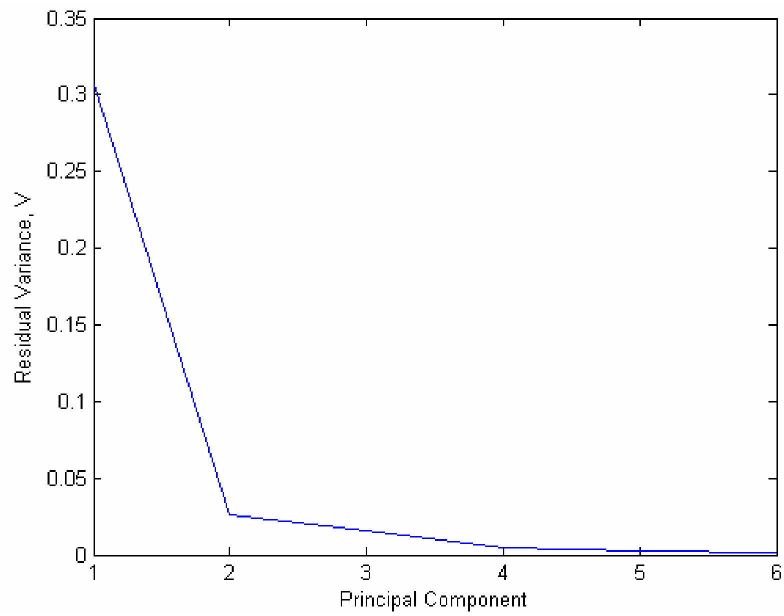



**Figure 53 – Scree Plot for Gumbeel Catchment**

Based on the Scores plot and the Biplot, the following conclusions can be derived:

- TP is very strongly correlated with SS indicating that it is in particulate form.

- TN and TOC shows some correlation with SS indicating that these parameters are partially in particulate form.

- An appreciable fraction of TOC is in dissolved form or as DOC. Hence the comments made in relation to Bonogin catchment are also relevant here.

- Structural pollutant abatement measures such as sediment traps will be particularly effective in the case of SS and TP and partially effective in the case of TN and TOC.

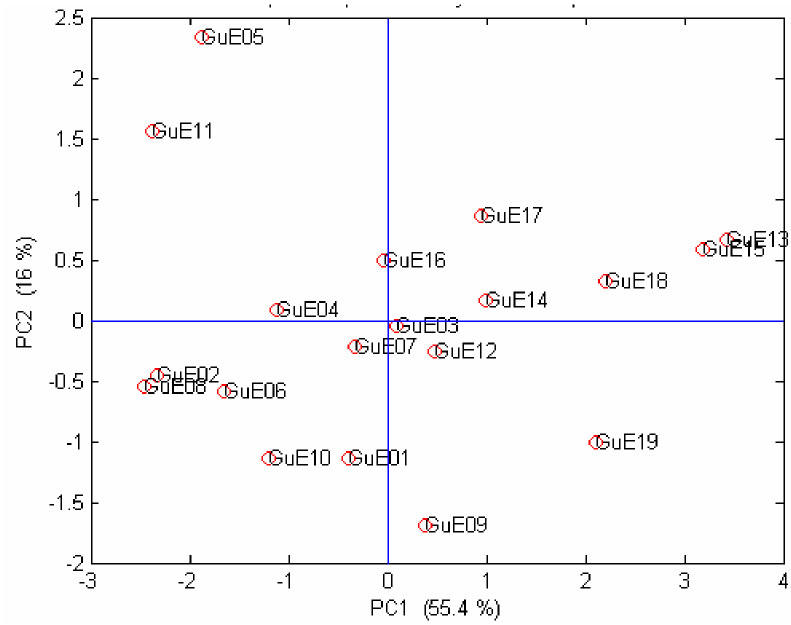



**Figure 54 –Scores plot for Gumbeel Catchment**

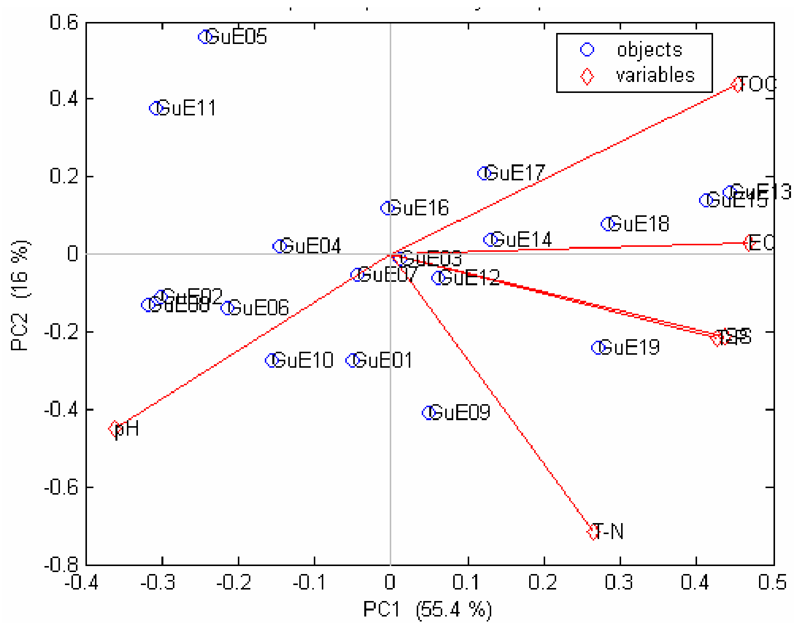



**Figure 55 –Biplot for Gumbeel Catchment**

**PLS Regression**

*PLS – TP regressed* (excludes pH and SS)

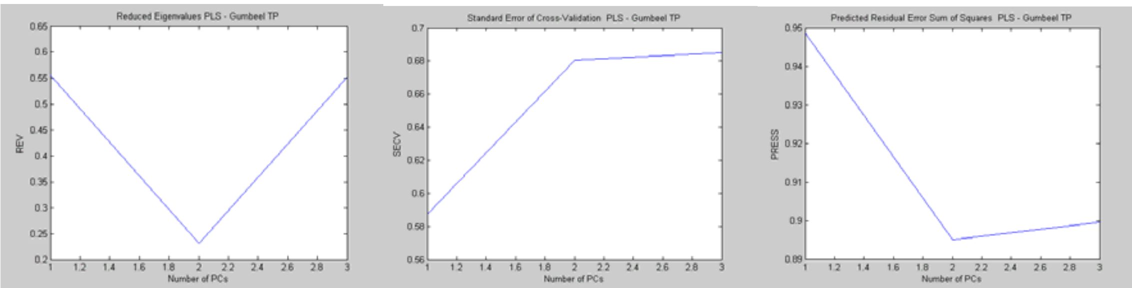
Calibration = 23 objects

Validation = 23 objects

PLS was performed on the data which was separated as:

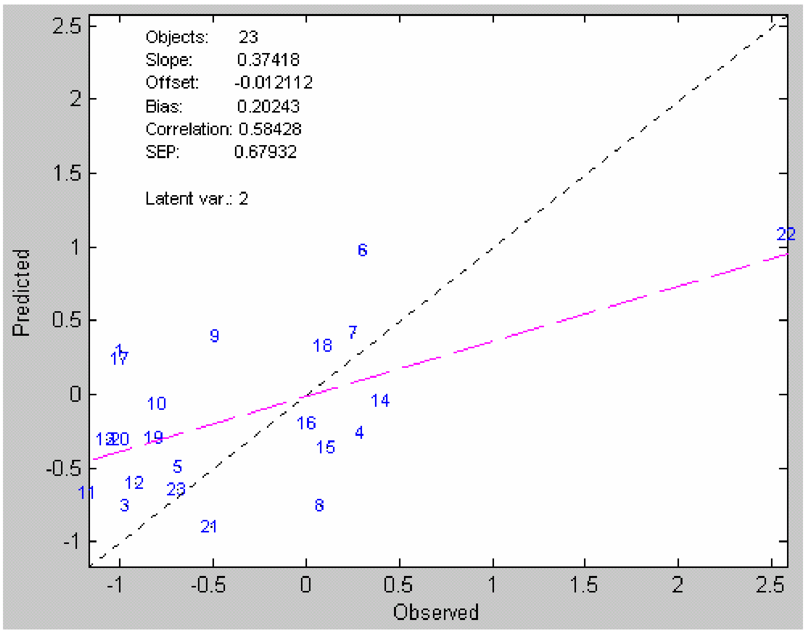
- Y = TP, X = TOC, TN and EC
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TP, the PLS model indicated that two significant factors out of three variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 56. The resulting Observed vs Predicted Calibration and Validation plots shown in Figures 57 and 58 indicated that the model is able to provide only a limited prediction, with $r^2$ values of 0.58 and 0.45 respectively. The main reason for this is related to the correlation between the variables. TOC is not strongly correlated with the other variables, indicating that a poor model performance will exist. Both the calibration and validation model were biased with the model under predicting values for both calibration and validation. The errors of fit also indicated that the model is not predicting very well. Most of the X data variance was extracted ($R^2X$=85%) in the models, and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=86%) also indicated reasonable predictability using cross validation. However, the $R^2Y$ value (amount of Y variance explained), was poor (38%), indicating that the prediction of TP values is considerably poor. The resulting error of prediction (RMSEP=2.3), indicated some inherent errors in the prediction model.




n=2

**Figure 56 – PLS analysis error plots for TP for Gumbeel Catchment**




R2X=84.977%
R2Y=38.702%
Q2=86.085%
RMSEP=2.3553

**Figure 57 – TP calibration plot for Gumbeel Catchment**

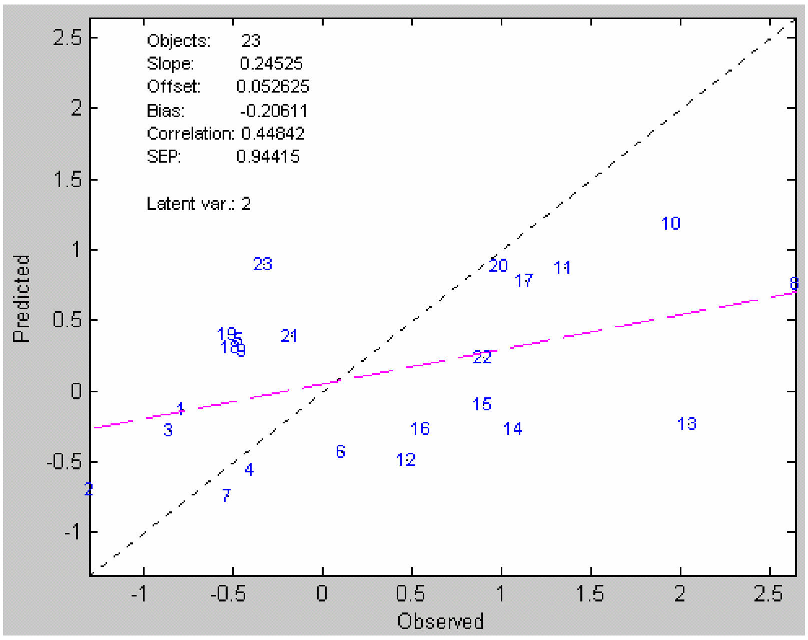



**Figure 58 – TP validation plot for Gumbeel Catchment**

*PLS – TOC regressed* (excludes pH)

Calibration = 23 objects

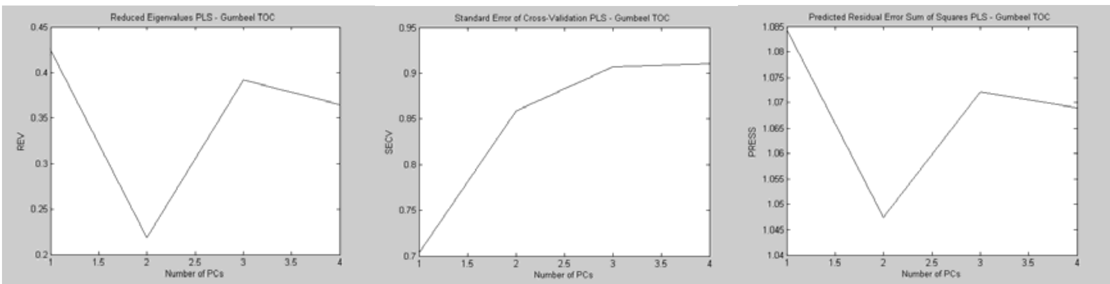
Validation = 23 objects

PLS was performed on the data which was separated as:

- Y = TOC, X = TOC, TN, SS and EC

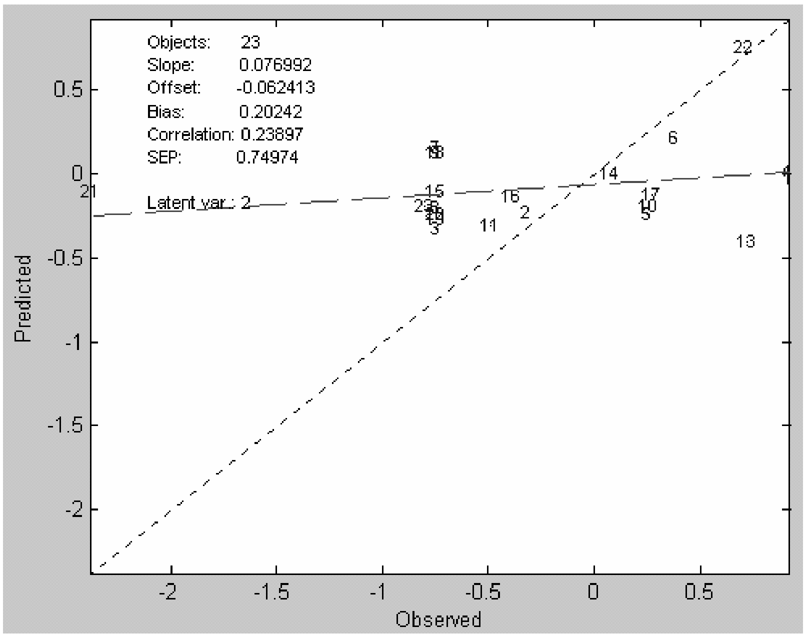
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TOC, the PLS model indicated that two significant factors out of four variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 59. The corresponding Observed vs Predicted Calibration and Validation plots shown in Figures 60 and 61 indicated that the model is poor, with $r^2$ values of 0.24 and 0.49 respectively. The validation provided marginally better results. The main reason for this is related to the correlation between the variables. Both the calibration and validation models retained some bias in the respective predictions, with the calibration model over predicting the values. Only a minor bias was observed for the validation. The errors of fit also indicated that the model is predicting poorly, although most of the X data variance was extracted ($R^2X$=87%) in the models, and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=70%) also indicated reasonable predictability using cross validation. However, this has reduced slightly compared to other models, most probably due to the arrangement of values for X-Y data sets. The $R^2Y$ value (amount of Y variance explained), was very poor (11%), indicating that the prediction of TP values is not suitable. Similarly, the resulting error of prediction (RMSEP=3.6), also indicated that errors in the prediction model are becoming more prominent.



n = 2

**Figure 59 – PLS analysis error plots for TOC for Gumbeel Catchment**




R2X=86.72
R2Y=11.315
Q2=69.391
RMSEP=3.6481

**Figure 60 – TOC calibration plot for Gumbeel Catchment**

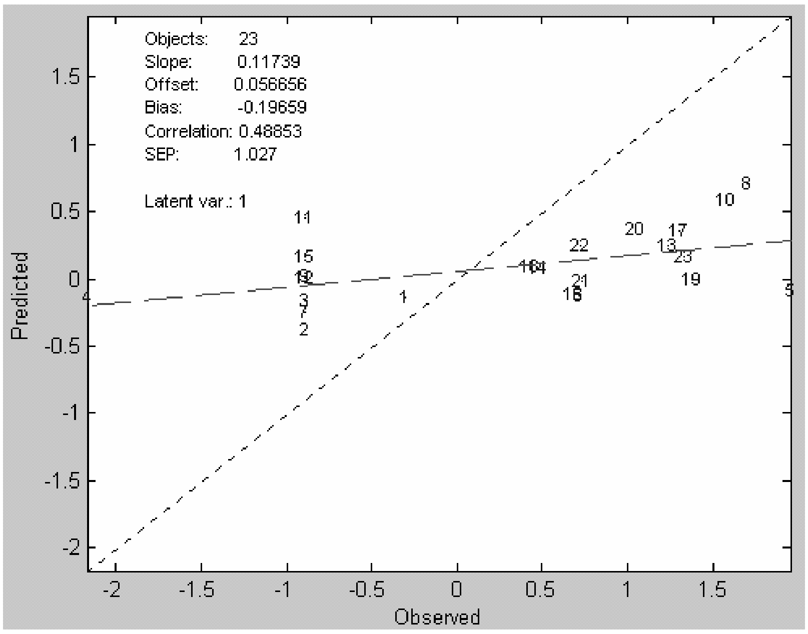



**Figure 61 – TOC validation plot for Gumbeel Catchment**

*PLS – TN regressed* (excludes pH and SS)

Calibration = 23 objects

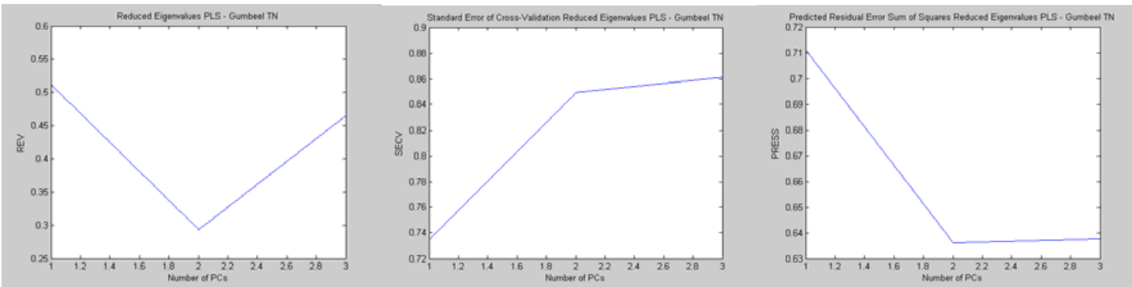
Validation = 23 objects

PLS was performed on the data which was separated as:

- Y = TN, X = TOC, TP and EC

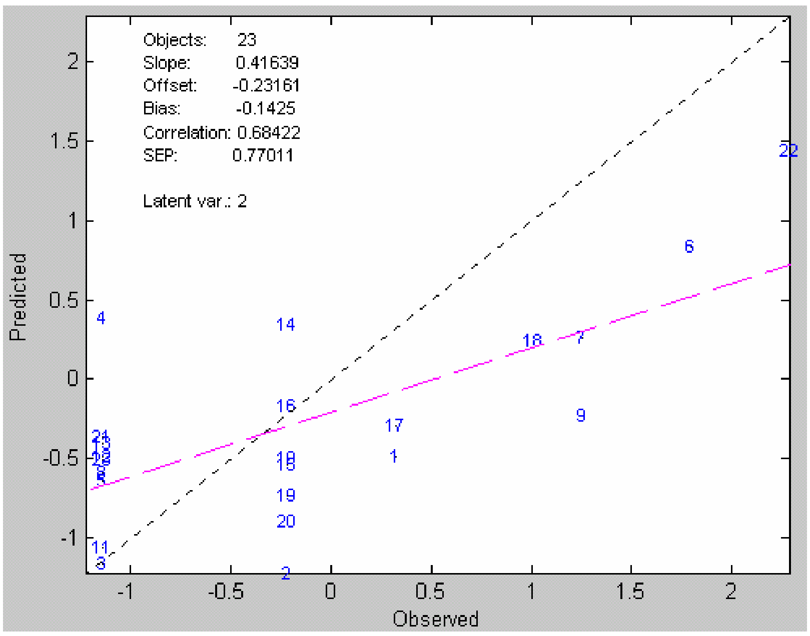
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TN, the PLS model indicated that two significant factors out of three variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 62. The corresponding Observed vs Predicted Calibration and Validation plots shown in Figures 63 and 64 indicated that the model performed slightly better than for TOC, with $r^2$ values of 0.68 and 0.62 respectively. The calibration model retained some bias towards under predicting the TN values. However, the validation produced minimal bias to either the observed or predicted values. The errors of fit also indicated that the model is predicting poorly. Some of the X data variance was extracted ($R^2X$=79%) in the models than has normally been found, although the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=95%) still indicated reasonable predictability using cross validation. The $R^2Y$ value




n=2

**Figure 62 – PLS analysis error plots for TN for Gumbeel Catchment**




$R^2X$=79.31%
$R^2Y$=45.763%
$Q^2$=94.693%
RMSEP=2.5995

**Figure 63 – TN calibration plot for Gumbeel Catchment**

(amount of Y variance explained), although better than for TP (46%), indicate that the prediction of TP values will be of limited accuracy. Similarly, the resulting error of prediction (RMSEP=2.6), also indicated that there are some inherent errors in the prediction model.

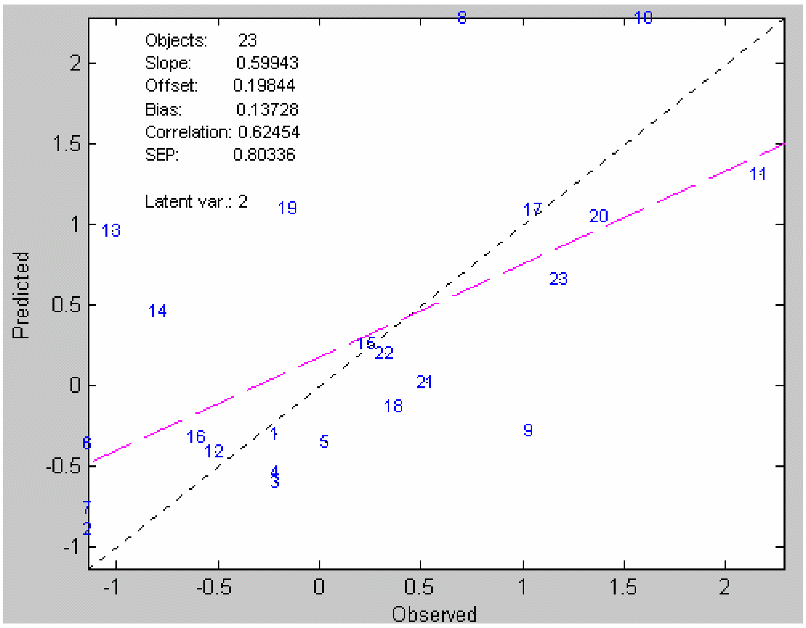



**Figure 64 – TN validation plot for Gumbeel Catchment**

### 5.2.6 Birdlife Catchment

**PCA Analysis**

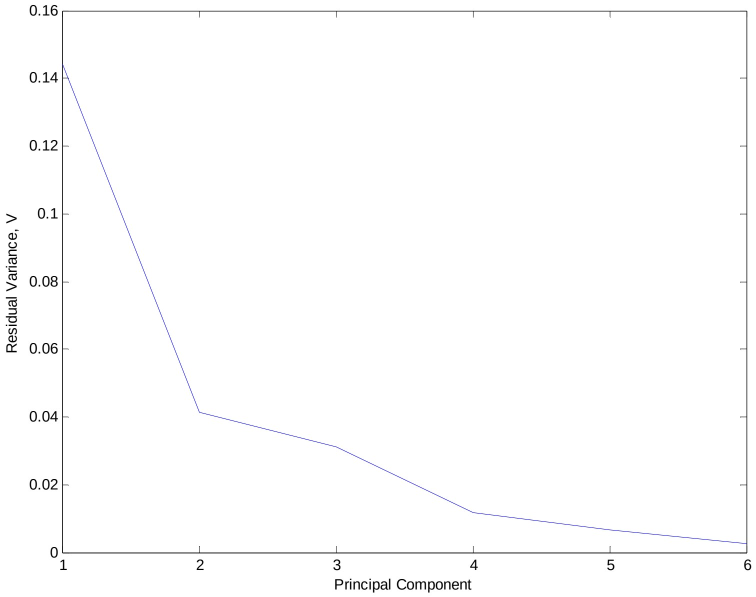



**Figure 65 – Scree Plot for Birdlife Catchment**

The analysis for Birdlife EMC data involved six variables; pH, EC, SS, TOC, TP and TN and twenty four objects, BiE01 - BiE24 (ie 24x6 data matrix). One potential outlier was also found, and subsequently removed providing a 23x6 data matrix. From this analysis and the corresponding Scree plot, as shown in Figure 65, it was determined that only the first two PCs were significant, contributing 58.3% of the data variance. The resulting Scores plot and Biplot are shown in Figures 66 and 67.

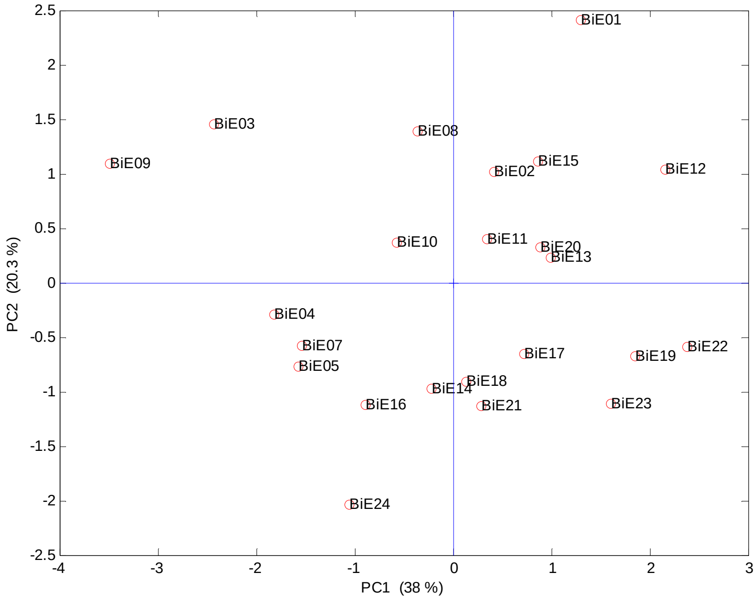



**Figure 66 –Scores plot for Birdlife Catchment**

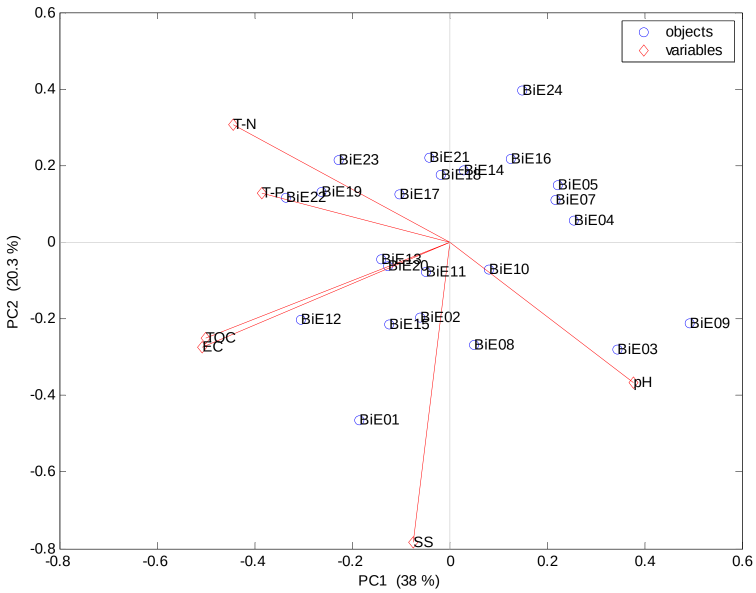



**Figure 67 –Biplot for Birdlife Catchment**

Based on the Scores plot and the Biplot, the following conclusions can be derived:

- TN and TP are strongly correlated with each other but is not correlated with SS. Hence TN and TP would be in dissolved form.

- TOC is only weakly correlated with SS. Hence TOC would be primarily in dissolved form as DOC and the comments made in relation to the Bonogin catchment area also relevant here.

- The fact that TN, TP and TOC are primarily in soluble form would mean that structural pollutant abatement measures such as sediment traps will only be effective in the removal of SS and not the other primary pollutants.

**PLS Regression**

*PLS – TP regressed* (excludes pH)

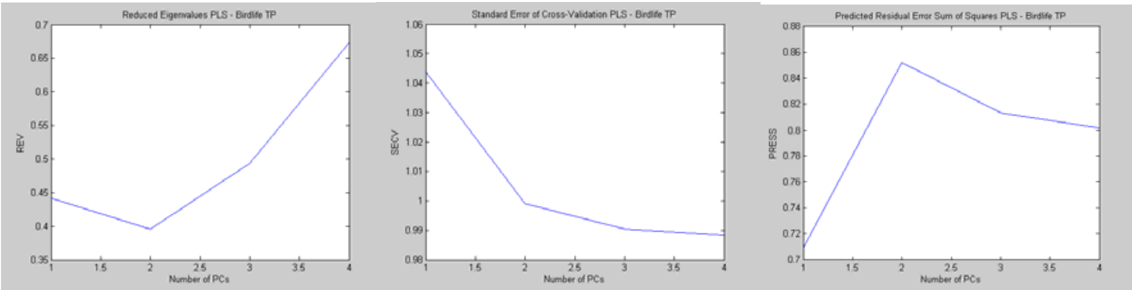
Calibration = 26 objects

Validation = 25 objects

PLS was performed on the data which was separated as:

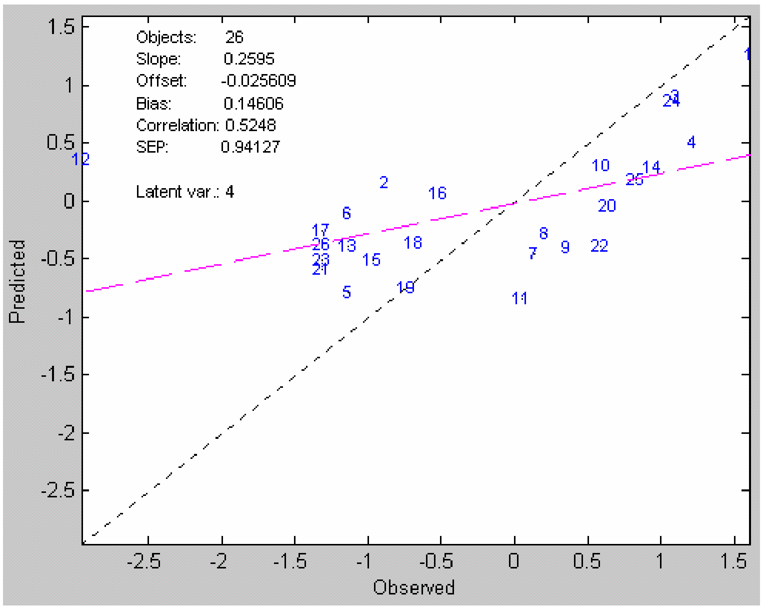
- Y = TP, X = EC, TN, TOC and SS
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the resulting analysis for TP, the PLS model indicated that four significant factors out of four variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 68. The resulting Observed vs Predicted Calibration plot as shown in Figure 69, indicated that the model provided a poor fit fo the data with an $r^2$ value of 0.52. This is most probably due to the calibration needing more variables to obtain the necessary data variance that retain some correlation between the variable. The validation model, as shown in Figure 70 too did not perform as well ($r^2$=0.2), and tended to retain some bias towards over predicting. It was also found to under predict the predicted values, with slightly more bias towards the observed. The resulting errors of fit indicated that most of the X data variance was extracted ($R^2X$=85%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=92%) also indicated good predictability using cross validation. However, the $R^2Y$ value (amount of Y variance explained) was much low (33%), indicating that the prediction model was not acceptable. The resulting error of prediction (RMSEP=2.4), indicated that errors in the prediction model were an issue.




n = 4

**Figure 68 – PLS analysis error plots for TP for Birdlife Catchment**




R2X=85.104%
R2Y=33.052%
Q2=92.061%
RMSEP=2.3825

**Figure 69 – TP calibration plot for Birdlife Catchment**

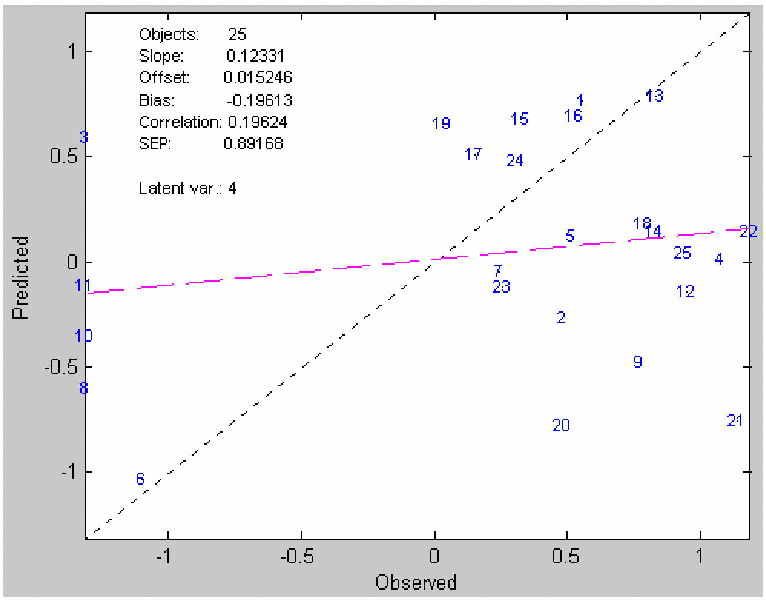



**Figure 70 – TP validation plot for Birldlife Catchment**

*PLS – TOC regressed* (excluding pH, SS, TP)

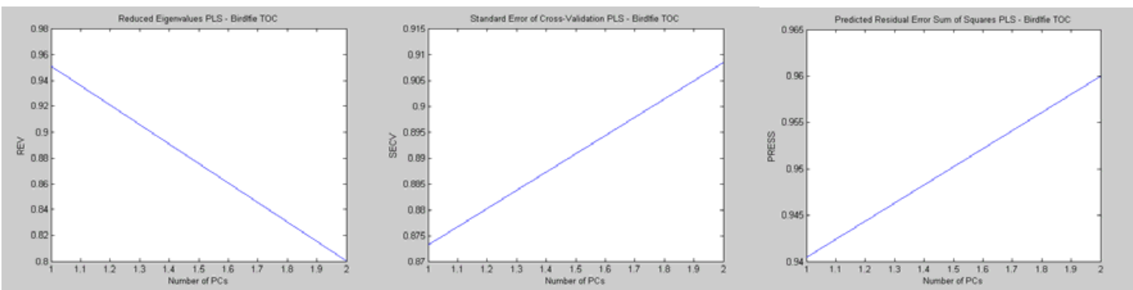
Calibration = 26 objects

Validation = 25 objects

PLS was performed on the data which was separated as:

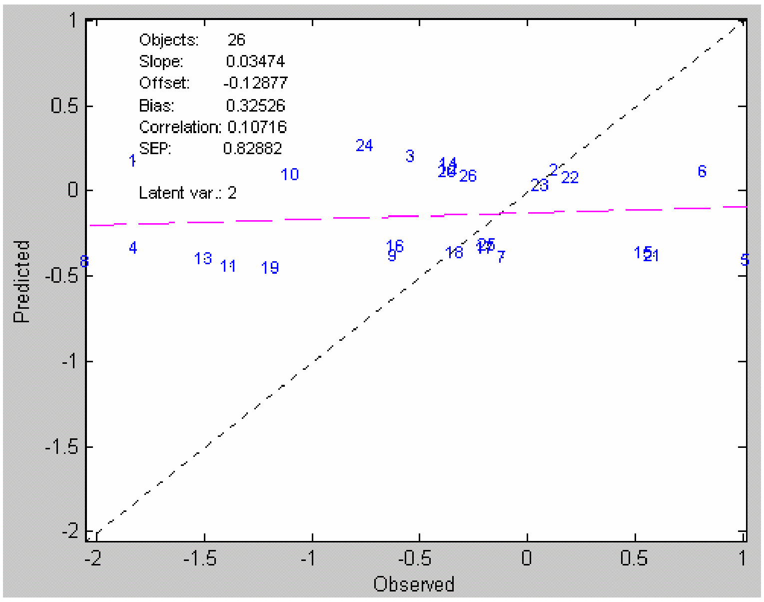
- Y = TOC, X = EC, TN
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TOC, the PLS model indicated that two significant factors out of two variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 71. The resulting Observed vs Predicted Calibration and Validation plots shown in Figures 72 and 73 performed poorly, as indicated by $r^2$ values of 0.1 and 0.2 respectively. This is primarily due to the use of only two variables for prediction, and although most of the variance is extracted for predicting the Y values, numerous errors will be induced. Both models were biased with the calibration model biased to over predicting the values and the validation model tending to under predict. The resulting errors of fit again indicated that most of the X data variance was extracted ($R^2X$=99%), and the $Q^2$ (amount of variance explained in predicted Y via cross validation $Q^2$=92%) also indicated good predictability using cross validation. However, the $R^2Y$ value (amount of Y variance explained) was very poor (11%), indicating that the prediction model was not acceptable. Similarly, the resulting error of prediction (RMSEP=3.2), is quite high indicating that substantial errors exist in the prediction model.




n = 2

**Figure 71 – PLS analysis error plots for TOC for Birdlife Catchment**




$R^2X$=98.716%
$R^2Y$=10.859%
$Q^2$=91.914%
RMSEP=3.1563

**Figure 72 – TOC calibration plot for Birdlife Catchment**

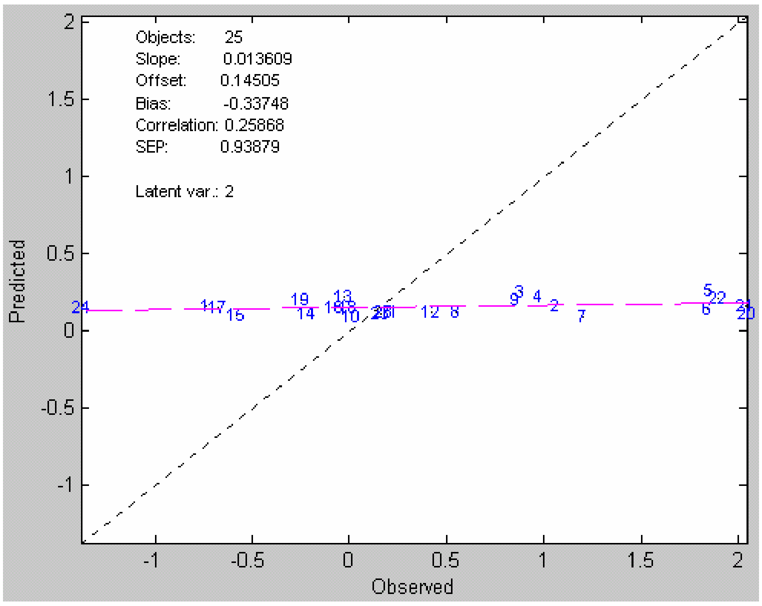



**Figure 73 – TOC validation plot for Birdlife Catchment**

*PLS – TN regressed* (excluding pH and SS)

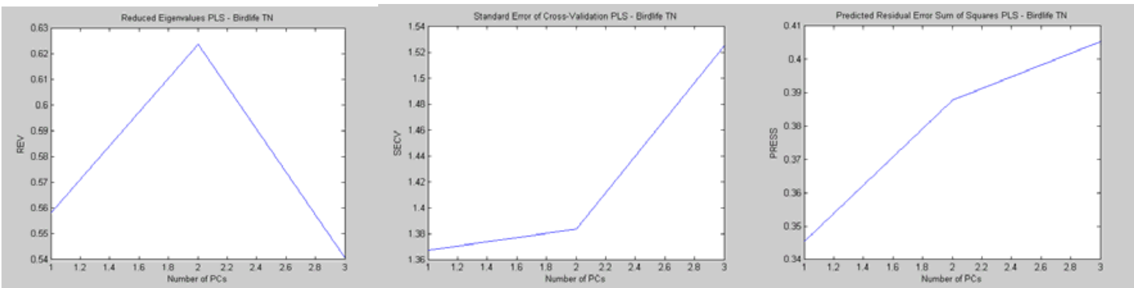
Calibration = 26 objects

Validation = 25 objects

PLS was performed on the data which was separated as:

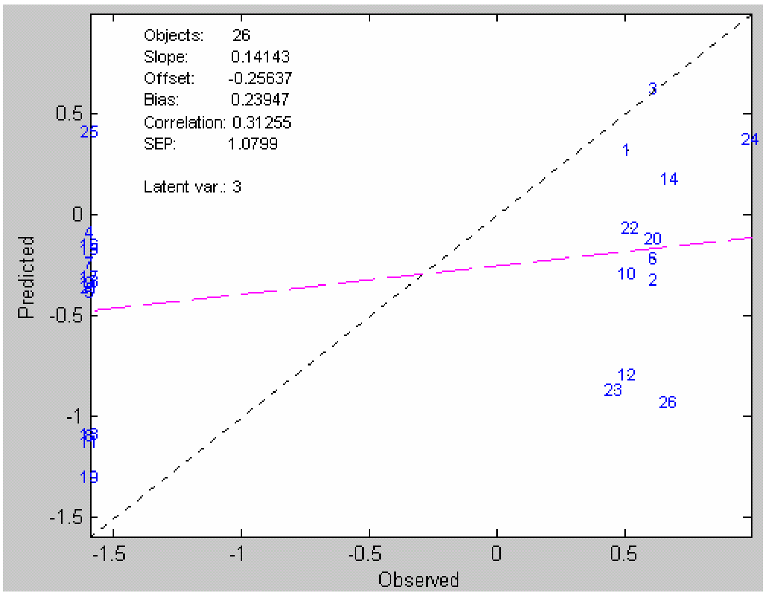
- Y = TN, X = TP, TOC, EC and TN
- Calibration/Validation matrices were separated according to the split rule (ie first half to calibration, remainder to validation)

In the analysis for TN, the PLS model indicated that three significant factors out of four variables used were required to model the predictions, as indicated by the predicted error plots shown in Figure 74. However, all three error plots indicated that significant errors are already present, and the resulting PLS model will be quite poor (all predicted errors are increasing). The resulting Observed vs Predicted Calibration and Validation plots as shown in Figures 75 and 76 perform poorly, as indicated by $r^2$ values of 0.37 and 0.21 respectively. This is related to the errors already predicted in the model. Same as for the TOC PLS, both models were again biased, with the calibration model being more biased to over predicting the values, were as the validation tended to under predict. Similarly, the resulting errors of fit again indicate that most of the X data variance was extracted ($R^2X$=92%), and the $Q^2$ (amount of variance explained in




n=3

**Figure 74 – PLS analysis error plots for TN for Birdlife Catchment**




$R^2X$=91.759%
$R^2Y$=25.664%
$Q^2$=98.923%
RMSEP=3.1961

**Figure 75 – TN calibration plot for Birdlife Catchment**

predicted Y via cross validation $Q^2$=99%) also indicated good data variance using cross validation. However, the $R^2Y$ value (amount of Y variance explained) was again poor (26%), indicating that the prediction model was not acceptable. Similarly, the resulting error of prediction (RMSEP=3.12, is also quite high indicating that substantial errors exist in the prediction model.

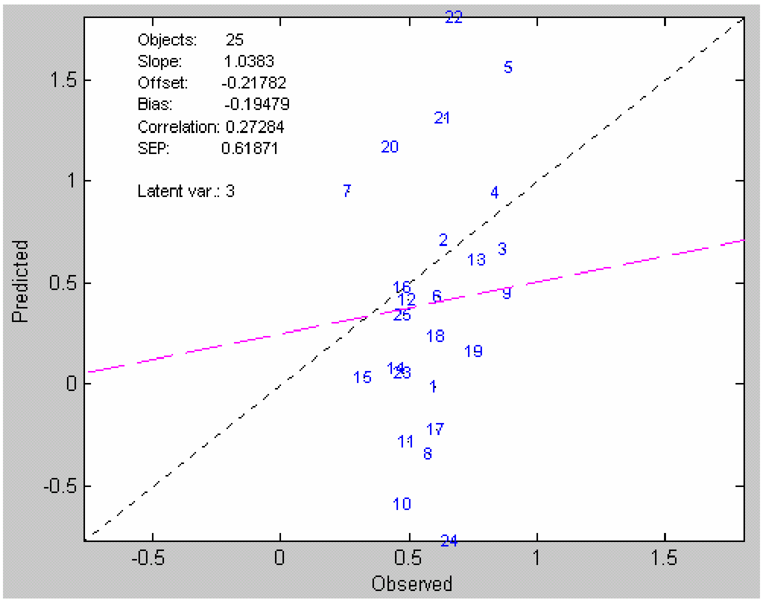



**Figure 76 – TN validation plot for Birdlife Catchment**

### 5.2.7 Summary of conclusions from the univariate analysis

Based on the conclusions derived from the analysis undertaken, it was evident that:

- There was no significant change in pH values between catchments. This could be related to the soil conditions in the study areas which are of similar characteristics.

- The mean values and standard deviations for the primary water quality parameters were generally found to increase with increasing urbanisation. The increase in mean values can be attributed to the increase in stormwater runoff pollution due to urbanisation. The increase in standard deviations indicates a high variability in runoff quality. This underlies the difficulties in developing predictive water quality models.

- For all three main catchments, the particulate bound component of heavy metals was significantly higher than the dissolved component. The pH value has a significant impact on the desorption of heavy metals adsorbed on particulates. The pH values were quite stable which could explain the fractioning of the heavy metals.

- Al, Mn and Fe are generally sourced from the soil and would be the result of erosion. The values obtained were found to increase with increasing urbanisation.

- For the study areas there was no appreciable difference in heavy metal concentrations. This runs counter to the general trend in increase in heavy metal concentrations with urbanisation as noted in numerous research studies. As the type of urbanisation in the study areas was residential, the primary source of heavy metals would be from roadways. Secondly, due to the fact that the dissolved fractions were below detection limit, it could be surmised overall that heavy metals are not a significant issue in these study areas. It is the dissolved component which is readily bioavailable.

### 5.2.8 Summary of conclusions from PCA analysis

Based on the conclusions derived from the analysis undertaken, it was evident that:

- The six study areas behaved quite differently to each other in terms of correlationships between different parameters which were not consistent for all catchments. This would make it difficult to develop stereotypical strategies for water quality management.

- It was quite common for TOC to be soluble in form (as DOC) in a number of study areas. This parameter can exert a significant influence on urban water quality particularly in relation to the bioavailability of heavy metals and hydrocarbons.

- Primary pollutants such as TN, TP and TOC were commonly in soluble form. This would mean that the effectiveness of structural pollutant abatement measures such as sediment traps is open to question as these measures are dependent on gravity settling.

### 5.2.9 Summary of conclusions from PLS regression

The calibration models derived for predicting various parameter values were of questionable value generally resulting in large errors of prediction. This in effect reflects the conclusions derived from the PCA where it was noted that strong correlationships between parameters were not particularly common. Secondly, the results obtained were not consistent across all the study areas. This in turn calls into question a common practice in urban water quality research where attempts are made to derive predictive models, quite often using univariate stratistical methods.

# 6. CONCLUSIONS

The outcomes from this study bring into question a number of fundamental concepts routinely accepted in stormwater quality management. The fact that the pollutant characteristics are not consistent across all the study areas would mean that the land use characteristics or the urban form is the overriding factor influencing water quality. This not only relates to the concentration but also to the chemical composition of stormwater pollutants. These conclusions would mean that the effectiveness of structural measures would not be universal and stereotypical solutions will prove inadequate. The common management technique of dealing with suspended materials as a primary treatment measure for urban stormwater quality may not always be successful as other pollutants are not necessarily in suspended form. It was repeatedly found that SS in most occasions is not correlated with TN, TP or TOC. Therefore as much of the pollution is moving in dissolved form, it is more bio-available and is therefore more likely to cause pollution in receiving waters. It could well be that this condition is linked to the climatic and rainfall conditions experienced in the study region which significantly influences pollutant composition, build-up and wash-off. Therefore it is important that predictive models developed have the versatility to take these characteristics into consideration.

The study confirmed that there is a general increase in pollutant concentrations with increasing urbanisation. However an increase in standard deviations was also observed which is significant. It indicates a high variability in stormwater runoff quality with increasing urbanisation. This underlies the difficulties in predicting the quality of urban runoff and the large margins error usually associated with predictive modelling.

The above findings underline the need to move beyond the dependency on customary structural measures and 'end-of-pipe' solutions and the key role that urban planning can play in safeguarding urban water environments. The univariate and multivariate statistical data analysis undertaken found that among the different urban forms, stormwater runoff from the area with detached housing in large suburban blocks exhibited the highest concentration and variability of pollutants. This is based on the concentration of various pollutants, their high variability and physico-chemical form. It is probable that these pollutants are being generated from the landscaped gardens and the relatively greater extent of road surface area.

Rural residential on large blocks were only marginally better. It could be concluded that in terms of safeguarding water quality, high density residential development which results in a relatively smaller footprint should be the preferred option. Based on the comprehensive study into correlating water quality to urban form, the important role that urban planning can play in safeguarding urban water environments was confirmed.

# 7. REFERENCES


1. Adams, M. J. (1995). Chemometrics in Analytical Chemistry. Cambridge, England: The Royal Society of Chemistry.
2. Ahyerre, M., Chebbo, G., Tassin, B. and Gaume, E., 1998, ‘Storm water quality modelling, an ambitious objective?’ Water Science and Technology, Vol. 37, No. 1, pp. 205-13.
3. Allison, R. A., Chiew, F. H. S. and McMahon, T. A., 1998, ‘Nutrient contribution of leaf litter in urban stormwater’, Journal of Environmental Management, Vol. 54, pp. 269-72.
4. Andral, M. C. (1999). Particles size distribution and hydrodynamic characteristics of solid matter carried by runoff from motorways. Water Environment Research, 71, 398-407.
5. Angino, E. E., Magnuson, L. M. and Stewart, G. F., 1972, Effects of urbanization on storm water quality runoff, Water Resources Research, Vol. 8, No. 1, pp. 135-40.
6. APHA. (1999). Standard methods for the examination of water and wastewater. Virginia, USA: American Water Works Association, Water Environment Federation.
7. ASCE 1975, ASCE Task Committee on the Effects of Urbanization on Low Flow, Total Runoff, Infiltration, and Ground-Water Recharge of the Committee on Surface-Water Hydrology of the Hydraulics Division, ‘Aspects of Hydrological Effects of Urbanisation’, Journal of the Hydraulics Division, ASCE, Vol. 101, No. HY5, pp. 449-68.
8. Barrett, M. E., Irish, L. B., Malina, J. F. and Charbeneau, R. J., 1998, Characterization of highway runoff in Austin, Texas, area’, Journal of Environmental Engineering, Vol. 124, No.2, pp. 131-37.
9. Bertrand-Krajewski, J-L., Chebbo, G. and Saget, A., 1998, Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon’, Water Research, Vol. 32, No. 8, pp. 2341-56.
10. Cattell, R. B., 1966, The Scree Test for the Number of Factors, Multivariate Behavioral Research, Vol. 1, No., pp. 245-76.

11. Cordery, I., 1977, 'Quality characteristics of urban storm-runoff water in Sydney, Australia, Water Resources Research, Vol. 13, No. 1, pp. 197-202.
12. Dabakk, E., Nilsson, M., Geladi, P., Wold, S. and Renberg, I., 1999, 'Inferring lake water chemistry from filtered seston usng NIR Spectrometry', Water Research, Vol. 34, No. 5, pp. 1666-72.
13. Deletic, A., 1998, 'The first flush load of urban surface runoff', Water Research, Vol. 32, No. 8, pp. 2462-70.
14. GCCC, 2000, Water and pollution movement: study catchments: base mapping summary report, Strategic and Environment Branch, Planning Directorate, Gold Coast City Council
15. Greb, S. R. and Bannerman, R. T., 1997, 'Influence of particle size on wet pond effectiveness', Water Environment Research, Vol. 69, No. 6, pp. 1134-8.
16. Guerin, T. F., 1999, 'The extraction of aged PAH residues from a clay soil using sonication and a Soxhlet procedure: a comparative study', Journal of Environmental Monitoring, Vol. 1, pp. 63-67.
17. Hall, K. J. and Anderson, B. C., 1986, 'The toxicity and chemical composition of urban stormwater runoff, Canadian Journal of Civil Engineering, Vol. 15, No. 1, pp. 98-105.
18. Hall, M. J. and Ellis, J. B., 1985, 'Water quality problems of urban areas', GeoJournal, Vol. 11, No. 3, pp. 265-75.
19. Hamers, T., van den Brink, P. J., Mos, L., van der Linden, S. C., Legler, J., Koeman, J. H. and Murk, A. J., 2003, 'Estrogenic and esterase-inhibiting potency in rainwater in relation to pesticide concentrations, sampling season and location', Environmental Pollution, Vol. 123, No. 1, pp. 47-65.
20. Hoffman, E. J., Latimer, J. S., Mills, G. L. and Quinn, J. G., 1982, 'Petroleum hydrocarbons in urban runoff from a commercial land use area', Journal Water Pollution Control Federation, Vol. 54, No. 11, pp. 1517-25.
21. House, M. A., Ellis, J. B., Herricks, E. E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H. and Clifforde, I. T., 1993, Urban drainage-impacts on receiving water quality, Water Science and Technology, Vol. 27, No. 12, pp. 117-58.
22. Isbell, R.F., 1996, The Australian Soil Classification, Vol4 of Australian Soil and Land Survey Handbook Series, CSIRO Publishing

23. Kokot, S., Grigg, M., Panayiotou, H. and Phuong, T. D., 1998, 'Data interpretation by some common chemometrics methods'. Electroanalysis, Vol. 10, pp.1081-8.
24. Librando, V., Magazzu, G. and Publisi, A., 1995, 'Multivariate micropollutants analysis in Marine waters', Water Science and Technology, Vol. 32, No. 9-10, pp. 341-48.
25. Lopes, T. J., Fossum, K. D., Phillips, J. V. and Monical, J. E., 1995, 'Statistical summary of selected physical, chemical, and microbial characteristics and estimates of constituent loads in urban stormwater, Maricopa County, Arizona, US Geological Survey, Water-Resources Investigations Report 94-4240, Tucson, Arizona.
26. Marengo, E., Gennaro, M. C., Giacosa, D., Abrigo, C., Saini, G. and Avignone, M. T., 1995, 'How chemometrics can helpfully assist in evaluating environmental data. Lagoon water', Analytica Chimica Acta, Vol. 317, pp. 53-63.
27. Marhaba, T. F., Bengraine, K., Pu, Y. and Arago, J., 2003, 'Sepctral fluorescence signatures and partial least squares regression: model to predict dissolved organic carbon in water', Journal of Hazardous Materials.
28. Marsalek, J., Brownlee, B., Mayer, T., Lawal, S. and Larkin, G. A., 1997, 'Heavy metals and PAHs in stormwater runoff from the Skyway Bridge, Burlington, Ontario', Water Quality Research Journal of Canada, Vol. 32, No. 4, pp. 815-27.
29. Massart, D. L., Vandeginste, B. G. M., Deming, S. M., Michotte, Y. and Kaufman, L., 1988, Chemometrics - A Text Book, Elsevier, Amsterdam.
30. Mein, R. G. and Goyen, A. G., 1988, 'Urban Runoff', Civil Engineering Transactions, Vol. CE30, No. 4, Institution of Engineers Australia, pp 225-38.
31. Novotny, V., Sung, H. M., Bannerman, H. and Baum, K., 1985, 'Estimating nonpoint pollution from small urban watersheds', Journal Water Pollution Control Federation, Vol. 57, pp. 744-8.
32. Parker, J. T. C., Fossum, K. D. and Ingersoll, T. L., 2000, 'Chemical characteristics of urban stormwater sediments and implications for environmental management, Maricopa County, Arizona', Environmental Management, Vol. 26, No. 1, pp. 99-115.
33. Parks, S. J. and Baker, L. A., 1997, 'Sources and transport of organic carbon in an Arizona river-reservoir system', Water Research, Vol. 31, No. 7, pp. 1751-9.

34. Pechacek, L. D., 1994, Urban runoff based on land use and particle size', Proceedings of the 1994 National Conference on Hydraulic Engineering, pp. 1243-6.
35. Pitt, R., 1979, Demonstration of nonpoint pollution abatement through improved street cleaning practices, U.S. Environmental Protection Agency, Report No. EPA/600/2-79-161.
36. Rahman, A., Thomas, E., Bhuiyan, S. & Goonetilleke, A., 2002, 'Modelling pollutant washoff from south-east Queensland catchments Australia', Enviro 2002 and International Water Association 3rd World Water Congress, Melbourne, 7-12 April.
37. Roger, S., Montrejaud-Vignoles, M., Andral, M. C., Herremans, L. and Fortune, J. P., 1998, 'Mineral, physical and chemical analysis of the solid matter carried by motorway runoff water', Water Research, vol. 32, No. 4, pp. 1119-25.
38. Saget, A., Chebbo, G. and Bertrand-Krajewski, J-L., 1996, 'The first flush in sewer systems', Water Science and Technology, Vol. 33, No. 9, pp. 101-8.
39. Sartor, J. D., Boyd, G. B., 1972, Water pollution aspects of street surface contaminants, Report No. EPA-R2-72/081, US Environmental Protection Agency, Washington, DC, USA.
40. Shinya, M., Tsuchinaga, T., Kitano, M., Yamada, Y. and Ishikawa, M., 2000, 'Characterization of heavy metals and polycyclic aromatic hydrocarbons in urban highway runoff', Water Science and Technology, Vol. 42, No. 7-8, pp. 201-208.
41. Tai, Y-L., 1991, Physical and chemical characterization of street dust and dirt from urban areas, Master of Science Thesis, The Graduate School, Department of Civil Engineering, The Pennsylvania State University.
42. US EPA, 1986, Analysis of Polycyclic Aromatic Hydrocarbons Method 610, 3rd edition, US EPA, OH.
43. US EPA, 2002, Peer consultation workshop on approaches to polycyclic aromatic hydrocarbon (PAH) health assessment, Report No. EPA/035/R-02/005.
44. Vazquez, A., Costoya, M., Pena, R. M., Garcia, S. and Herrero, C., 2003, 'A rainwater quality monitoring network: a preliminary study of the composition of rainwater in Galicia (NW Spain)', Chemosphere, Vol. 51, pp. 375-86.
45. Warren, N., Allan, I. J., Carter, J. E., House, W. A. and Parker, A., 2003, 'Pesticides and other micro-organic contaminants in freshwater sedimentary environments – a review', Applied Geochemistry, Vol. 18, pp. 159-94.

46. Westerhoff, P. and Anning, D., 2000, 'Concentrations and characteristics of organic carbon in surface water in Arizona: influence of urbaniszation', Journal of Hydrology, Vol. 236, pp. 202-22.
47. Wold, S., Sjostrom, M. and Eriksson, L., 2001, 'PLS-regression: a basic tool of chemometrics', Chemometrics and Intelligent Laboratory Systems, Vol. 58, pp. 109-30.
48. Wunderlin, D. A., Díaz María del Pilar, A. M. V., Pesce, S. F., Hued, A. C. and Bistoni, M., 2001, 'Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A Case Study: Suquía River Basin (Córdoba–Argentina)', Water Research, Vol. 35, No. 12, pp. 2881-94.