UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

School of Geography, Archaeology and Environmental Studies

Towards a GIS Platform for monitoring the water quality of the Jukskei river catchment: defining a sampling network

A research report submitted for the fulfilment of the degree of Master of Science in GIS and Remote Sensing

Student: Reneilwe Malatji 813058

Supervisor: Dr. Stefania Merlo

Co-Supervisor: Prof. Chris Curtis

Declaration

The research report titled “**Towards a GIS Platform for monitoring the water quality of the Jukskei river catchment: defining a sampling network**” is conducted under the supervision of Dr. Stefania Merlo and Prof. Chris Curtis for the MSc in Geographical Information Systems and Remote Sensing at the University of the Witwatersrand, Johannesburg.

I, the undersigned, hereby declare that this research report is my own original work and all sources have been accurately reported and acknowledged and that this document has not been previously, in its entirety or in part, submitted at any university to obtain an academic qualification.

……………………………………………

Malatji Reneilwe

Student Number 813058

May 2019

Acknowledgments

This accomplishment would not have been possible without the contribution of several individuals.

Foremost, I would first like to express my sincere gratitude to my supervisor Dr. Stefania Merlo of the school of GAES, Faculty of Science at the University of the Witwatersrand, for her continuous support, patience, enthusiasm and immense knowledge. Her office was always open whenever I ran into difficulties on my research and constantly allowed this paper to be my own work but steered me in the right direction whenever she thought needed.

Besides my supervisor, I would like to thank my co-supervisor Prof. Chris Curtis for his encouragement, insightful comments, and hard questions.

I would also like to thank the experts who were involved in the validation survey for this research including the City of Johannesburg Municipality (COJ) for allowing me to dig into their datasets and systems and more specifically Mpho Zwane who held my hand throughout the fieldwork and data collection process.

To my Sponsor Department of Water and Sanitation (DWS) Learning Academy, for financial support, I am grateful.

I would also like to acknowledge Nompumelelo Lukhele of the CSIR DST, Drashti Naik, Yingisani Chabalala of GAES, and Calvin Opiti of Digital Mining as the second readers of this report, and I am grateful to them for their valuable comments on this report.

I thank my fellow labmates, Galaletsang Kebine, Tinobokashe Ncube, Khumalo Perseverance, Nkanyiso Cele and Heritage Fani for stimulating discussions, for the sleepless nights we were working together before deadlines and for all the fun and support we shared throughout the whole process.

I must express my very profound gratitude to my parents and siblings for providing me with unfailing support and continuous encouragement throughout my years of study and process of researching and writing this report.

Above all, I owe it all to the Almighty God for granting me the wisdom, health, and strength to undertake this research task and enabling me to its completion.

Reneilwe Malatji
May 2019

Table of Contents

List of Figures

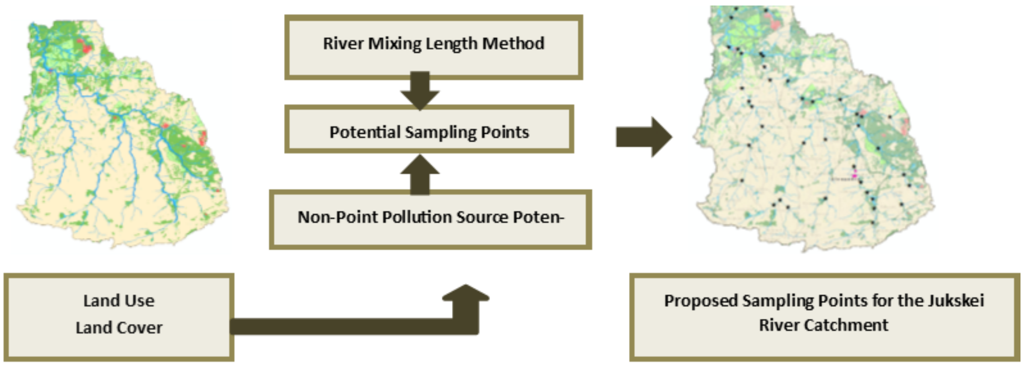
List of Tables

Appendices

Abbreviations and Acronyms

Acronym	**Abbreviations**
AHP	Analytic Hierarchical Process
ANP	Analytic Network Process
CoJ	City of Johannesburg
DEA	Data Envelopment Analysis
DEM	Digital Elevation Model
DRDLR	Department of Rural Development and Land Reform
DWA	Department of Water Affairs
DWS	Department of Water and Sanitation
EISD	Environmental Infrastructure and Services Department
EMC	Event Mean Concentrations
RQIS	Resource Quality Information Services
JW	Johannesburg Water
LULC	Land Use Land Cover
MOE	Municipal Entity
PPS	Potential Pollution Score
RML	River Mixing length
WB	Water Body/ Water Bodies

Graphical Abstract



Abstract

To obtain an understanding of issues relating to water quality monitoring in the Jukskei River Catchment a users' needs assessment is performed, through the use of literature review, interviews, and site visits. These are evaluated and used to develop a conceptual model giving a representation of a system that can address water quality concerns. The designed conceptual model for this research incorporates hydrological features, drainage systems, DEM extraction and thematic layers needed for the monitoring of water quality. The development of appropriate sampling locations is an essential step in the valuation, restoration, and protection of river water quality. Previous studies have focused on mathematical aspects and physical attributes of monitoring the quality of river water, however, human activities surrounding the river are often not considered in detail. In this study, an advanced water quality monitoring procedure is implemented using a combination of River Mixing Length (RML), information about human activities, and Geographic Information Systems (GIS) to identify representative sampling locations, with the purpose of optimizing the current locations of sampling points. The implementation of a new Potential Pollution Score (PPS) model of land use is used to prioritize the importance of each sampling point prior to selecting the most appropriate site for the entire catchment. Residential areas, mining and agriculture are amongst the 6 non-point sources integrated to develop representative sampling locations. Out of an optimal number of 11 242 sampling points, this research has identified 40 sites to propose. This number of sites is chosen because it is closer to the number of sampling points currently used (32) by the EISD for water quality monitoring of the Jukskei river catchment and thus does not stretch their funds.

Keywords: water quality; monitoring network; design methodology; sampling location; river water; GIS; RML

1. Introduction

1.1. Background of the study

Seventy-one percent (71%) of the earth's surface is covered by water, of which 2.52% is fresh water and less than 0.3% is in rivers, lakes and the atmosphere (Chen and Han, 2018). Access to good quality drinking water is a basic essential right, and an important component of an effective policy for health protection (WHO, 2004). This makes the development of surface water monitoring very critical for the protection and restoration of rivers (Ouyang, 2005).

Pollution from different and uncontrolled sources is a continuous water quality problem, creating the necessity for the detection, control, and analysis of spatial-temporal conditions of a water body (Lin *et al.*, 2017). Traditional water monitoring strategies have mainly relied on the collection of ground data, completed through the use of sampling points and their analysis (Alilou *et al.*, 2018). These water quality analyses, based on in situ monitoring approaches, have developed through the years to the use of electronic sensors and more recently, towards the use of remote sensing and the integration of these approaches using GIS (Chen and Han, 2018).

A GIS environment is considered by many as a viable solution to gathering and sharing information for the better management and interpretation of datasets by means of spatial thematic mapping tools (Gobakis *et al.*, 2017). GIS-based platforms and environments are adopted by organizations with spatial data as the main specialty for their area of interest, aiming at facilitating data access and visualization of spatial datasets and thus allow the simulation of results and differing conditions (Kulkarni *et al.*, 2014; Assaf *et al.*, 2008).

One of the crucial steps for the effective water monitoring strategy for a flowing water body is the selection of appropriate optimized sampling sites, which are chosen to fit certain criteria, such as suitable location. Whilst in the past, such locations were chosen using mere convenience, nowadays GIS and statistical analysis are employed to generate sampling points suitable for subsequent GIS-based analysis (Alilou *et al.*, 2018).

1.2. Problem statement

Water pollution remains the biggest concern for water quality in South Africa. In townships in urban areas, people occupy land illegally at times and thus water provision and distribution are not always organized (Amdany *et al.*, 2014). Sanitation is not provided at several locations within the country, in particular in the City of Johannesburg Municipality. There are many informally occupied areas which make up part of the Alexandra Township; these areas are considered to be some of the major causes of pollution of the Jukskei river catchment (Zwane, 2018).

The foundation of this research is based on information provided by the City of Johannesburg-Environmental Infrastructure and Services Department (EISD). Pressing issues relating to water quality in the municipality include the lack of usable data on the water quality of the Jukskei river catchment, particularly because of inefficient data capturing and poorly chosen sampling point locations.

Even though sampling points located along the river are monitored and collected monthly, their results portray no change in water quality (Zwane, 2018). However, residents alongside the river continuously raise alerts on the changes associated with the quality of water, which are not reflected in any of the sample analyses performed based on the in-situ data sampling and the information provided by the EISD (Zwane, 2018). Other than the issues with sampling and sampling strategy, there is a lack of efficient organization of the collected datasets, which makes continuous monitoring and comparative analysis difficult.

This research proposes the use of a GIS platform as a solution to the issues of water quality monitoring of the Jukskei river catchment.

1.3. Research questions

1. What are the limitations of the current sampling scheme of the Jukskei River?
2. What sampling strategy would be ideal for the improvement of sampling in the catchment?
3. What are the requirements for a GIS platform to be used by the city of Johannesburg EISD for the monitoring of water quality?

1.4. Aim and objectives

This research aims to design a GIS platform to be used by the Environmental Infrastructure and Services Department for monitoring the water quality of the Jukskei River Catchment. As a fundamental step towards improving the limitations of the current water samples collection system, GIS based analysis will also be used to suggest an optimized strategy for locating water sampling points along the catchment.

The following objectives will be pursued to achieve the aim of the research:

1. Determine the scope of a GIS platform for water quality monitoring using a user's needs assessment;
2. Design the conceptual model of a GIS for water monitoring in the Jukskei catchment;
3. Optimize the sampling strategy for water quality monitoring using GIS.

2. Review of the literature

2.1 Water quality monitoring

Water quality monitoring refers to the attainment of measurable and illustrative information on the physical, chemical, and biological characteristics of a water body over time and space (Sanders *et al.*, 1983). To understand the dynamics of a catchment, a well-designed water quality monitoring strategy is required to identify water quality issues, while creating the baseline values needed for trend analysis (Strobl and Robillard, 2008).

Clean water contamination and clean water scarcity are amongst the major challenges of the 21st century in developing nations like South Africa (Antonini *et al.*, 2017). According to a study by Antonini *et al.* (2017), 30% of world's population still do not have the privilege of clean water. Their study goes on to suggest the implementation of a GIS platform for the easy sharing of information and updates on the changing status of water bodies.

Qin *et al.* (2018) call attention to the fact that continuously keeping track of the state of water helps maintain water safety and water quality standards; particularly if the water from a river is used for consumption. However, he does not suggest ways to track the quality. Research by Sanders *et al.* (1983), Park (2006) and Alilou *et al.* (2018) highlights that water quality is a challenging subject specifically regarding the spatial and temporal variability of the moving water body.

The strategies of water quality monitoring require addressing of gaps in monitoring, which allows the ordering of priorities, in terms of the level of importance, to provide clear reasons for its establishment (Myers and Ludtke, 2017). This helps in the assessment of the authorities' commitment to water quality issues and addressing them through the basic act of response rate assessment on a platform accessible to everyone.

Optimal water monitoring design structure should satisfy monitoring objectives more efficiently and effectively than any other possible configuration (Telci *et al.*, 2009). The water quality monitoring program must also be flexible, due to the fact that the monitoring strategy may need to be modified at a later stage, as deemed appropriate (Steele, 1987). The monitoring design procedure is, therefore, an iterative process in which any information gained since the last iteration may lead to revisions in the network design of each new iteration (Qin *et al.*, 2018).

With each monitoring strategy, there is a need for a feedback loop, in which the effects of those decisions are evaluated; the success, as well as failure of previous management decisions can be best assessed in this way (Park, 2006). However, the monitoring strategy iterative process is commonly ignored once the water quality monitoring network has been established, leading to the collection of fixed and hence, unusable, inadequate and costly data (Qin *et al.*, 2018).

In the planning and implementation of water monitoring strategies, there is a fear of redesigning a water quality monitoring program owing to the fact that, should existing programs be abandoned for any reason, the ability to determine trends may be forfeited (DWS, 2017). Additionally, established water quality monitoring strategy often includes fixed sampling stations, which are only reluctantly deserted.

2.2. Spatial sampling of water quality

Despite several decades of operations and the increasing importance of water quality monitoring of rivers, municipalities still rely on experiential insights and subjective judgments in locating water quality monitoring points (Park *et al.*, 2006).

The placement of a permanent sampling station is described by Sanders *et al.* (1983) as the most critical design factor in any water quality monitoring design. However, research has shown that criteria to establish sampling locations for representative sampling have received relatively little attention from the government departments responsible for water quality monitoring worldwide (Sanders *et al.*, 1983; Harmancioglu *et al.*, 1999).

Previously, the participation of local citizens and the public was not of major concern and focus; however, more recently, citizen science has been used to raise awareness of water quality issues. Nonetheless, research has shown that with regards to crucial issues such as water quality, citizen science is not very effective. This is explained by Jollymore *et al.* (2017) as unreliable and mistrusted data, because the manner in which the general public understands and attach reasons and meaning to data is not the same manner in which it is scientifically perceived, on this topic. Jalbert and Kinchy (2016) elaborate that citizen science may cost more money than the results are worth.

Most monitoring strategy designs subdivide the sampling locations into two groups, namely macro-locations and micro-locations (Park *et al.*, 2006). Macro locations are usually sampled within a watershed, whereas micro-locations refer to critical sampling points, such as outfalls, determined relative to the macro location. Often in current water quality monitoring network design, macro-locations are allocated systematically, while micro-locations are functions of macro-locations and critical locations (Harmancioglu *et al.*, 1999).

Traditionally, water quality samples were often collected from points of mere convenience, such as bridges and areas near dumping sites. There however exist more structured approaches to designating sampling locations, such as the method explained by Horton (1945). This method defines stream order by specifying the smallest unbranched tributary in the headwaters of a watershed as first order and so on (Horton, 1945). A stream consisting solely of first-order tributaries is considered as second order. In 1971, Sharp (1971) used Horton's approach to systematically designate sampling station locations by subdividing the stream into equal segments of contributing tributaries. This procedure was developed to identify and isolate pollution point sources (Sharp, 1971).

Sharp's (1971) procedure also incorporated work by Shreve (1967), in which pollution point sources were used to identify sampling locations. Other approaches included selecting sampling locations as a function of possible stream standards violations or as a function of stream segments below outfalls (Beckers *et al.*, 1972). Ward (1973) suggested locating sampling points at critical quality points such as major pollution sources. Whichever approach is employed, a rational systematic approach, which includes a well-defined and outlined strategy and flexibility of iterations, is needed.

2.3. Sampling location design

Sharp's procedure (Sharp, 1971) and Sander's application (Sanders, 1980) have been utilized in several studies to locate representative sampling points as supported by studies by Park *et al.*, (2006) and Strobl and Robillard (2008). However, these methods can only be applied in large river basins and henceforth it is based on the number of tributaries. If a river system has few or no branches, and the length of each branch is substantially different, the method is not entirely relevant.

Because Sharp's (1971) and Sander's (1980) methods are dependent on the number of tributaries, sampling point locations cannot be identified in detail. This is because the position of an upstream, middle stream or downstream point for river tributaries cannot be accurately determined. To solve this limitation and to provide additional advantages, a new procedure, called River Mixing Length (RML), is developed in this study to locate river basin sampling points through the integration of GIS.

In the RML method, rivers and branches are divided into small segments, and the length of each segment is equal to the river's mixing length. Subsequently, each river's mixing length is considered a potential sampling point to monitor changes in water quality. Each tributary contributing to the main river has its own characteristics. For this reason, it is necessary to consider river characteristics and human activities when designing a sampling network. In the RML method, like in any other water monitoring method, a representative location of sampling points is arguably the most critical factor (Sanders *et al*., 1983).

In recent decades, the design of water quality monitoring sampling points has been studied by many researchers (Karamouz *et al.,* 2008). Some monitoring networks have been designed based on Sharp's procedure and Sanders' application for locating representative sampling points (Chilundo *et al.*, 2008, Park *et al.*, 2006, Strobl and Robillard, 2008).

Several researchers have used statistical methods (Ouyang, 2005), Genetic Algorithms (Chen *et al.*, 2008, Icaga, 2005, Karamouz *et al.*, 2008, Park *et al.*, 2006, Ning and Chang, 2004), the Kriging method and the Entropy method (Lo *et al.*, 1996; Karamouz *et al*., 2008). However, most of these studies focused on mathematical aspects.

Non-point source pollutants are introduced into rivers, streams, lakes, and oceans through rainfall, snowmelt, runoff, and groundwater which contribute to base flow. Evaluation of land use activities is required to identify and manage non-point source pollution, and land use is one of the most important factors in modelling non-point source pollution (Ierodiaconou *et al.*, 2005, Sivertun and Prange, 2003, Zhang and Huang, 2010).

2.4. GIS in water quality monitoring

2.4.1 Uses of GIS in water quality monitoring

This section attempts to provide a clear indication of the difficulty and challenge in organizing water quality monitoring procedures in terms of communication, reporting and transferring and sharing information. GIS can be useful to the Environmental and Infrastructure Services Department (EISD) data organization in relation to collecting, recording, reporting and sharing water quality data from the Jukskei river catchment.

The integration of GIS and hydrological modeling in water monitoring strategies offer the significant advantage of applying the full information content of spatially distributed data to analyse hydrologic processes (Craigie, 2018 and Vieux, 1991). Early efforts to utilize spatial information for the purpose of gaining insight into hydrologic processes in a river basin include hydrological modeling, which was conducted per watershed (Huggins and Burney, 1982).

Band (1986) and Jenson and Dominique (1988) described a method of dividing a watershed into areas bounded by drainage divides and stream channels, using raster digital elevation data. This is amongst the early uses of GIS in the monitoring and thus the modeling of water quality.

A GIS interface is described as a systematic tool used for the collection of sparse information, which varies in nature. GIS develops this raw information into static charts, dynamic maps and sometimes matrices for the illustration of the spatial and temporal evolution of real-life processes using statistical simulations (Repetto *et al.*, 2018 and Kulkarni *et al.*, 2014).

A GIS platform does not only integrate the products of GIS and Internet technologies with the advantage of ease of access, data transparency, better and easier visualization and most importantly, cost-effectiveness, but also incorporates a feedback component (Al-sabhan *et al.*, 2003).

An assessment of users' needs is the initial and most crucial stage in the development of a GIS platform as it will be utilized by many people including the EISD and public users (Yi, 2018). In support of this view Myers and Ludtke (2017) emphasize the need for a background check on the problems, concerns and thus needs of those directly affected by the quality of the water body in question.

The water quality data used in many GIS reporting tools especially in South Africa is collected by DWS as part of the body's task of managing the quantity, quality and reliability of water resources for the benefit of all users (DWAF, 1996). Resource Quality Information Services (RQIS) which is part of DWS is responsible for water quality analysis and scientific valuation; which are significant factors required for trend analysis and monitoring water quality patterns over time. Given their role and the effects of the results published, there is a need for clear establishment of the users' needs assessments (DWAF, 1996).

GIS training is costly, especially when considering the different software and add-ins that are required, for both the learning process and in practical work application. Regardless of the fact that it can take years for a person to become fully competent, with the right GIS resources Water Quality (WQ) specialists can focus on data analysis and presentation (Cobban, 1994). Given that GIS applications are usually limited to one or a few functions with each application requiring the full-time attention of people trained with advanced capability in GIS; there is a need for the government to invest in this department (Kühn *et al.*, 1996).

2.4.2 GIS conceptual models for water management and monitoring

For effective and efficient presentation, simulation and exploration of different components that contribute to water quality monitoring and management; different conceptual models have been established and developed over time. Amongst other initiators is the company ESRI which has a variety of representative diagrams of hydrological models, watershed models, stream delineation models, and water pollution models.

Figure 1 gives a few illustrations of conceptual models created for different purposes by Esri; taking into consideration the fact that each conceptual model is different because of the different ideas they try to simulate or emphasize to the user or viewer. These contain the different datasets that are needed, the different infrastructures that contribute to monitoring water quality, and most importantly the characteristics of the water body to consider. These were used as guides to the development of a conceptual model for this research.

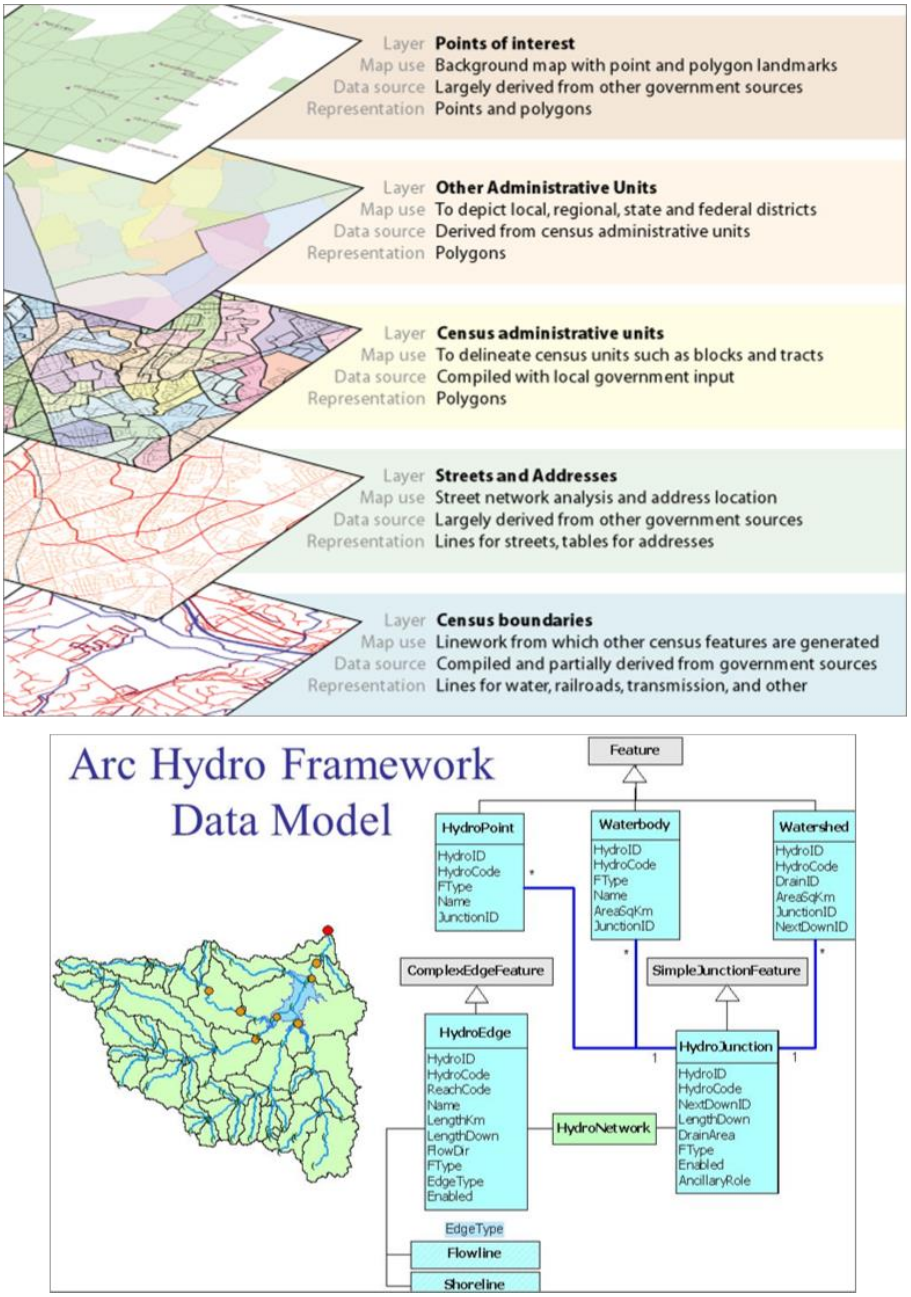


Figure 1: Examples of conceptual models from ESRI arc-hydro

Looking at the two examples of conceptual models (Figure 1); a similar process is used for their generation. The steps commonly used for the development of conceptual models are summarized and represented in Figure 2, beginning with the establishment of a problem statement which addresses the research question and the aim of the conceptual model, followed by the review of literature (scientifically published journal articles and relevant reliable sources), selection of important variables and lastly, generating the conceptual model.

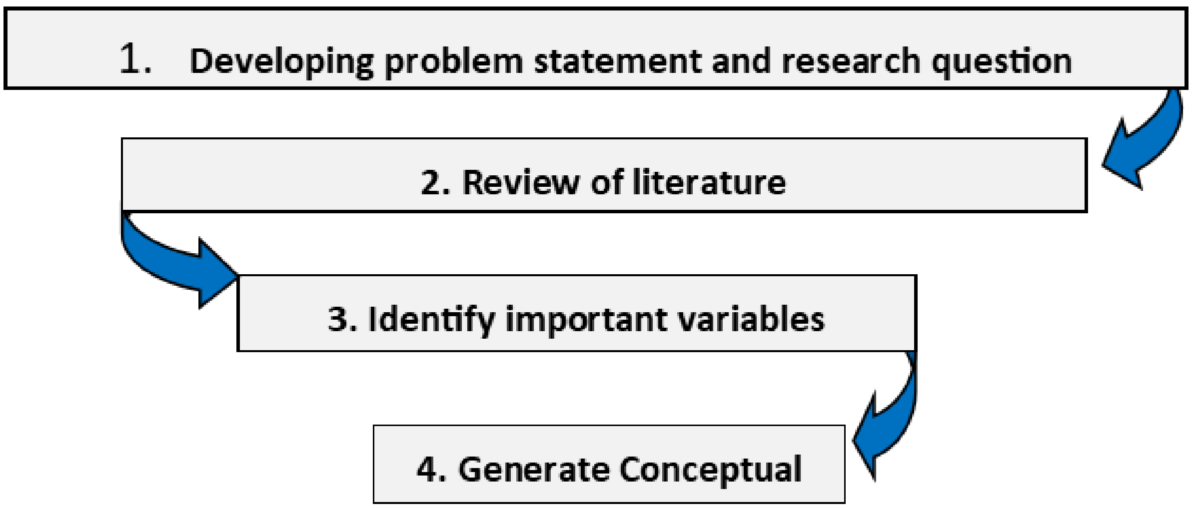


Figure 2: Process followed for the development of the conceptual

Data processing and statistics are critical parts of the conceptual model design. Here, the user selects the WQ variables to be displayed, the date range and map symbolization which will best present how data interconnects from its collection to its usage (ESRI, 1991).

Through the use of advanced GIS and web technology, regular time series data may be created. Various procedures and symbols can be tested on the computer display in combination with a selection of background geographical data such as geology; hydrology and land cover (Cobban and Silberbauer, 1993). With the help of the users' needs assessment, the conceptual model develops with depth and understanding of what attracts the user of the platform and consequently the data contained within it.

In 1992, the Council of Scientific and Industrial Research (CSIR) and RQIS worked together on a project titled "*Conceptual design report for a national river water quality assessment programme*" which was aimed at: i) determining data and the suitability thereof to meet the expectations of the information users; ii) developing a conceptual design of the operating measures; iii) developing criteria for selecting the monitoring design, and; iv) developing water quality assurance and thus; v) water quality control concepts (Harris *et al.*, 1992).

The Harris *et al.* (1992) project was influenced by Moore (1990), who motivated the need for DWS to release and regulate software, computer hardware, data acquisition, maintenance, and, most significantly, implement the conceptual model.

Tong and Chen (2002) illustrated the importance of integrating Land Use Land Cover (LULC) and water quality assessments through the use of technology, specifically, remote sensing and GIS analysis techniques. This way of thinking and modelling was later empowered and strengthened by the research of Rekolainen *et al.* (2003) which began to explain the idea and integration of the conceptual frameworks and modelling into water quality monitoring and resource management.

Lastly, research by Rekolainen *et al.* (2003) accentuated the importance of a conceptual framework and thus model inclusivity of the action planning, water body classification and characterization, monitoring, management planning, and implementation measures. This is one of the most critical points that GIS with consideration of LULC types and the functions of the water department should lean and move towards in the future.

3. Methodology

3.1 Study Area

The Jukskei River is one of the catchments of the Johannesburg Municipality. It covers an area of over 800km^2 (Campbell, 1996). This river empties its water into the Hartbeespoort Dam after merging with the Crocodile River (Figure 3). It meanders northwards through various residential areas including the densely populated Alexandra Township and only drains to the northern half of the city (Amdany *et al.*, 2014). This river catchment is monitored by the City of Johannesburg EISD once every month using a sampling strategy adapted from the Department of Water affairs, based on water quality standard (SANS 241) guided by the National Water Act (Act 36 of 1998).

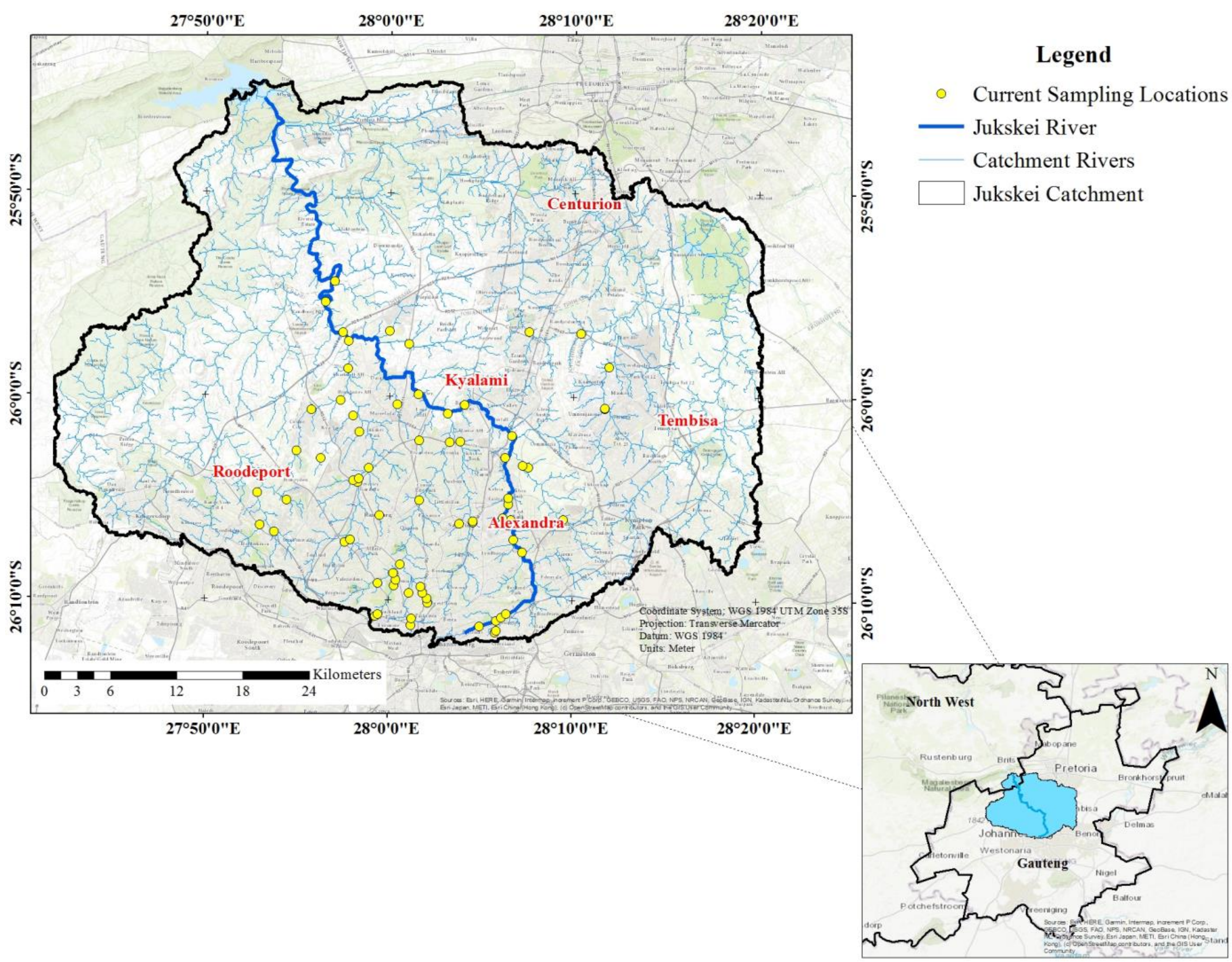


Figure 3. Map of the study area with the location of the water sampling locations used by the EISD

3.2 Current sampling system for Jukskei river catchment (as documented by the CoJ (2013)

The documentation provided by the EISD provides a background on how the current sampling locations were designed; it describes the issues and concerns related to water quality of the rivers in the metropolitan cities like Johannesburg in recent years. The document emphasizes the direct relationship between water scarcity and water pollution. Polluted water by various catchment-based activities means that water is not available for the specific intended use. Subsequently, the Jukskei and Rietspruit rivers are identified as the major rivers that drain the whole CoJ municipality.

Like other urban rivers, the CoJ river systems are highly polluted by various contaminants originating from a wide range of catchment activities within and outside CoJ borders. Pollution of CoJ rivers can be attributed to various sources and because of this, monitoring and reporting programs and strategies have become a priority to the EISD through the years.

The water quality monitoring program of the CoJ is run by Environmental Infrastructure and Services Department Management (EISD). This program is a critical tool that serves a very important role in directing catchment management interventions such as pollution control and law enforcement by various city departments. The monitoring program is comprised of 120 points located in rivers and tributaries throughout the whole city. The data and information generated from this program are circulated to among others, Environmental Health Department, Johannesburg Water (JW), Johannesburg Road Agency, and Johannesburg City Parks. According to the monitoring data (CoJ, 2013), there has been a consistent increase in sewage pollution since the beginning of the monitoring program, shown by the use of an indicator of faecal contamination, *Escherichia coli*.

Various city departments and entities are responsible for operations and projects significantly impacting on water quality of CoJ water resources. The impacts originating from these operations and projects are better managed by various forums established between and among relevant stakeholders. This, according to the EISD, ensures coordinated planning and better management of related impacts to water resources.

Amongst the different city departments, there are fora which are initiated to address a variety of water resource pollution challenges. These fora integrate the water management department and relevant MOEs responsible for any operation impacting on the environment. The formation of these fora is necessary due to the variability of operations and related environmental degradation associated with the specific operation.

The main issue associated with these fora is that they were never formalized and as a result, stakeholders are not committed, often not attending meetings, providing reports or implementing rehabilitation measures when environmental degradation occurs from their respective operations or projects (CoJ, 2013).

An example of these fora is the *Joburg Water - Water Quality Forum* which was initiated under the EISD (CoJ, 2013). Although well established, this forum is not immune to problems as it is currently non-functional, regardless of the fact that it is chaired by EISD Environmental Management represented by Water and Biodiversity Directorate Quality Management Unit of Natural Management Directorate (CoJ, 2013). This was in recognition that the city and its entities have a statutory responsibility towards the environment as required in the following legislation:

1. The Constitution of the Republic of South Africa, Act 108 of 1996
2. National Environmental Management Act, Act 107 of 1998
3. National Water Act, Act 86 of 1998
4. Environmental Conservation Act, Act 73 of 1989
5. Water Services Act. Water Services by-laws of the City of Johannesburg.

The primary objective of the *Joburg Water - Water Quality Forum* is to adopt a coordinated approach in managing environmental degradation emanating from Joburg Water operations and to provide strategic directions as well as the promotion of water resources protection approaches (CoJ, 2013). The second objective is to proactively inform operations, maintenance and land use practices that contribute to improved or declining water quality (CoJ, 2013).

The structure of CoJ/JW water quality fora consists of regional managers of all Johannesburg Water Regional offices with respect to their defined operational areas (Figure 4).

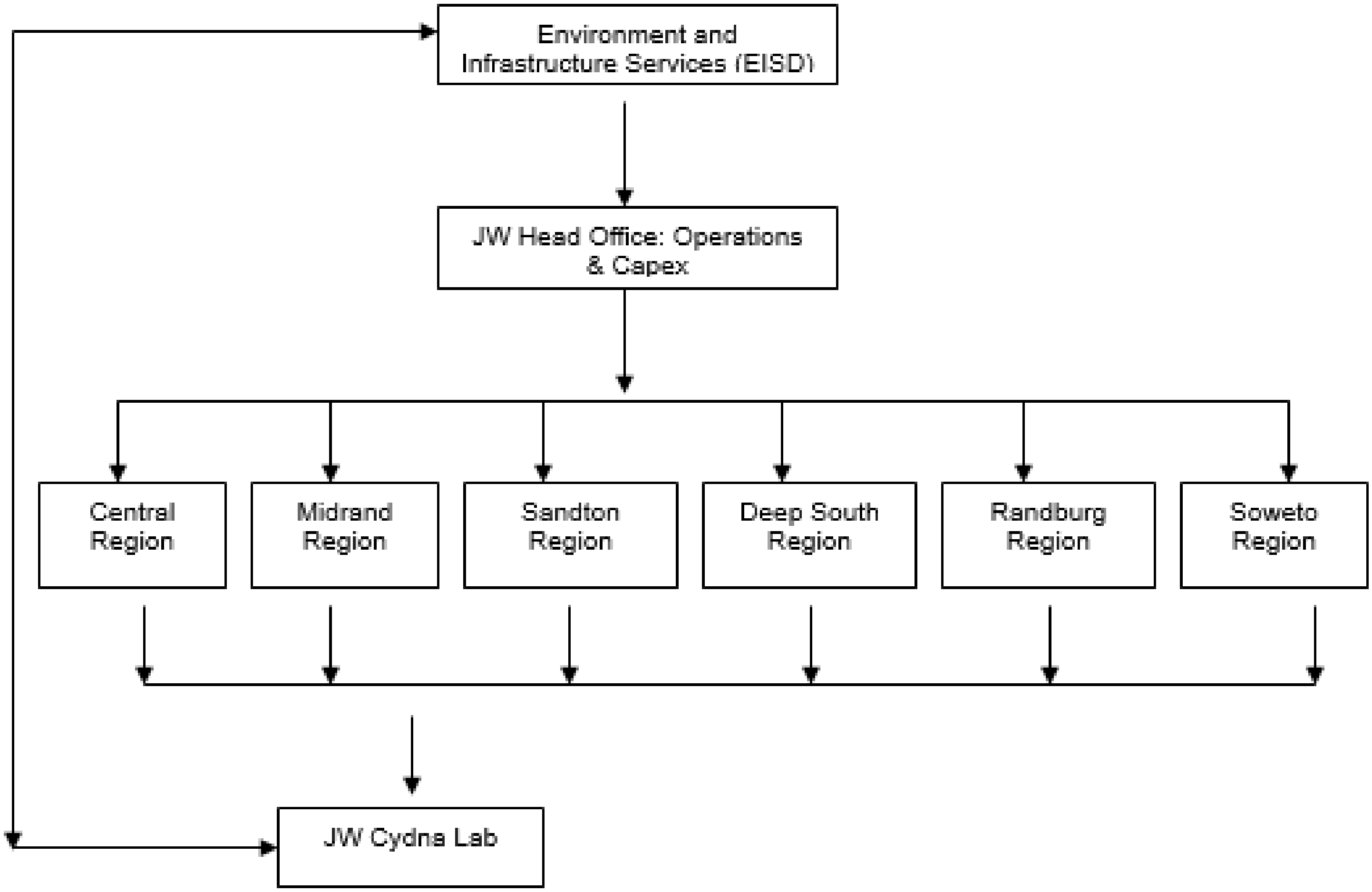


Figure 4. Water Monitoring Structure (COJ, 2013)

3.3. Data acquisition

3.3.1 User needs assessment

In determining the scope of this research, a review of relevant literature was conducted to identify the common needs associated with different users, followed by a users' assessment, which is conducted through face to face interviews with the municipality- EISD, river catchment region managers and potential users of the GIS platform. The questionnaire in Appendix A is used only as a guide for open-ended questions and discussion.

The most important component for the achievement of the conceptual model for the monitoring of the Jukskei River Catchment is the physical interviews; these include the main brief of this research report. These were carried out at the EISD of the CoJ in Braamfontein on the 25th June 2018 and 1st November 2018 (Traduna House); the site visits were then carried out during the monthly schedule of water quality sampling (6th December 2018). The interview details are provided in Table 1.

Table 1. Interview Details

<table>
<tr><th>Interviewee</th><th>Role</th><th>List of Questions</th></tr>
<tr><td>Nandy Manzini</td><td>EISD-Assistant Director (Water Quality Management – CoJ)</td><td rowspan="5">1. What is the sampling Strategy?
2. When were the sampling points established?
3. How were the sampling points established?
4. Has there been a change in sampling points since established?
5. Main Challenge with Data Capturing
6. Backup Database (X)
7. What kind of platform would you prefer for the GIS platform (X)
8. What main function will the GIS platform perform? (X)
9. What are the major Water Quality Concerns?</td></tr>
<tr><td>Emma Netshihulu</td><td>GIS Specialist and Information Management (CoJ)</td></tr>
<tr><td>Mpho Zwane</td><td>EISD</td></tr>
<tr><td>Clarence</td><td>Water Sampler</td></tr>
<tr><td>Sammy Katuma</td><td>Junior Researcher- GIS (GCRO)</td></tr>
</table>

3.3.2 Catchment verification and optimization of sampling

The GPS locations of the currently used sampling points were provided by the City of Johannesburg (Table 2). This includes their current sampling point's location, characteristics of the river, width and depth (these are useful in determining the kind of dataset that can be utilized).

Table 2. Coordinates of the currently used sampling locations provided by COJ

DWJ MONITORING PROGRAMME			
		Coordinates	
Site	**Location**	**Degrees, Minutes, Seconds**	
DWJ1	Downstream Sebenza	26 07.143S	28 10.423E
DWJ2	Downstream kelvin	26 07.131S	28 10.448E
DWJ3	Modderfontein Rd, near golf course	26 06.199S	28 09.964E
DWJ4	Modderfontein/ Midrand boundary		
DWJ6	Bruma Lake outflow	26 10.694S	28 06.444E
DWJ7	Upstream Gilloolys Farm	26 10.130S	28 07.362E
DWJ10	Lombardy East	26 06.987S	28 06.814E
DWJ11	North end far east Bank Rd, Alexandra	26 06.341S	28 06.520E
DWJ12	Marlboro bridge, Alexandra		
DWJ13	Bridge Rd, Buccleuch	26 03.012S	28 06.291E
DWJ14	Off Katherine Dr, Strathavon	26 06.210S	28 04.566E
DWJ15	Downstream Wynberg	26 06.143S	28 04.578E
DWJ16	Sunninghill	26 02.257S	28 03.864E
DWJ17	Rustenberg Rd, Parkhurst		
DWJ18	Forest Farm, Sandton	26 05.137S	28 01.641E
DWJ19	Downstream N1, Paulshof	26 02.190S	28 01.641E
DWJ20	Windsor		
DWJ21	Downstream Strydom Park		
DWJ22	Downstream Kya sands	26 00.832S	27 58.280E
DWJ23	Chartwell	25 58.692S	27 57.736E
DWJ24	After confluence with little Jukskei	26 01.888S	28 06.649E
DWJ26	Leeukop Prison	26 00.860S	28 03.146E
DWJ27	Upstream Northern Works	25 57.394S	27 59.376E
DWJ29	Downstraem NW & Klein kskei	25 57.342S	27 57.780E
DWJ33	Windsor Golf Course		
DWJ34	Modderfontein factory site		
DWJ36	First Ave Edenvale	26 07.947S	28 08.470E
DWJ37	Leeukop Golf Course	26 00.432S	28 04.083E
DWJ38	North of London Rd bridge, Alex	26 06.582S	28 06.768E
DWJ40	near Altrek centre, Alex	26 06.086S	28 06.640E
DWJ41	stormwater channel at Hofmeyer Bridge, Alex	26 05.958S	28 06.196E
DWJ42	Opposite S'tsewtla, Alex	26 05.683S	28 06.407E

The characteristics of all the datasets used for the research are listed in Table 3.

Table 3. Dataset description

Dataset	Format	Source	Date of Acquisition	Scale
Sampling Locations	Excel spreadsheet MXD (ArcMap as points –vector data)	EISD	25 June 2018	1:50 000
Jukskei River Catchment Characteristics	Excel spreadsheet	EISD	25 June 2018	1:50 000
South African Rivers	Shapefile (polyline-vector data)	SRTM DEM	04 July 2018	Threshold 500
Land Use Land Cover	Shapefile (tiff image- raster data)	DRDLR	02 December 2018	1:50 000
Stream Width	Coordinates (Latitude & Longitude)	Google Earth	25 June 2018	1: 4 200

3.4 Software

The software chosen for this research is listed in Table 4, including examples of studies that used it for similar research to this one.

Table 4. Software used

Name	Use	Source
ArcGIS (ArcMap)	Catchment Delineation Environmental Modelling	Ames *et al.* (2009), Chen *et al.* (2010), Jolma *et al.* (2010), Tischler *et al.* (2007)
SuperDecision	ANP Scoring	Kucukaltan *et al.* (2016)
Google Earth	Ground Data (Stream Width)	Telci *et al.* (2009)

ArcGIS (ArcMap)

ArcMap is the main component of ArcGIS, developed by Esri. It is mainly used to view, edit, create, and analyze geospatial data. It allows users to explore datasets and create new datasets. In this study, it is mainly used for calculations and for data presentation (maps).

SuperDecison

SuperDecision is decision-making software, used commonly for the implementation of the Analytic Hierarchy Process (AHP) and the Analytic Network Process (ANP). Its main purpose is to integrate judgment and data to effectively prioritize options and predict outcomes. For this study, the SuperDecision software is used for prioritizing the contribution of the LULC classes on water quality.

Google Earth

Google Earth, which is a geo-browser that accesses satellite and aerial imagery, ocean bathymetry, and other geographic information, is used in this research for measuring the stream width. It is used specifically due to its good spatial resolution of 15 m to 1 m which allows clear and accurate measurements.

3.5. Methods

This section describes the methodology followed in the research; it first gives a summary of the overall methodology (Figure 5), followed by an explanation of how the conceptual model is designed (section 3.5.1). Furthermore, the ways in which the hydrological networks were extracted is stated (section 3.5.2) and the optimization of the sampling points carried out in this study is explained (section 3.5.3).

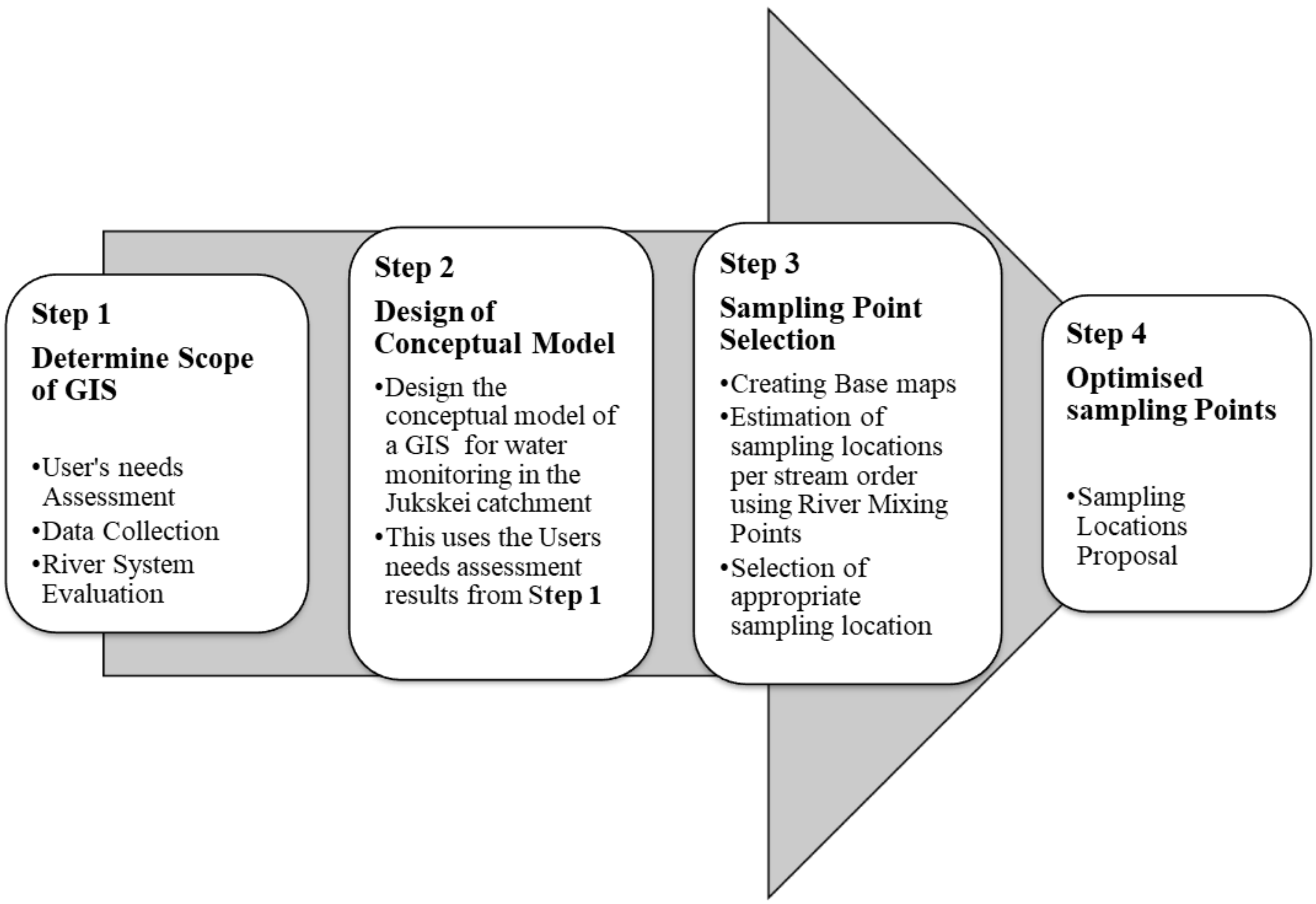


Figure 5. Summary of the Methodology adapted from Hoover *et al*. (2014)

3.5.1 Conceptual Model

A conceptual model is a representation of a system, this is to say, a platform made up of different concepts which are used to assist with the simulation and thus understanding of a subject of concern. For the satisfaction of objective 1 and 2 of this research, a conceptual model is developed; providing, first a briefing and secondly a hosting environment for all information related to the suitability analysis of geo-processes related to water quality monitoring.

The model incorporates all the datasets and interactions between different datasets and departments which are needed to make the EISD strategize and work better when it comes to the protection and conservation of good water quality standards for the Jukskei River Catchment.

The overall idea and structure of the conceptual model is developed from ideas gathered from Esri; among other examples, the Arc Hydro Water Utility Model stood out to the needs of this research and this example can be accessed from the following (Djokic, 2010); http://www.caee.utexas.edu/prof/Maidment/esri/ArcHydroRiver/Dec2010/Day2/Djokic.ppt

To effectively develop a conceptual model in the scope of this research, interviews and site visits were conducted to understand how the sampling is performed at the sampling locations and hence the procedures conducted (Figure 6). A- showing the Bruma Lake Outer sampling location, B- an example of how samples are collected from the Rhodes Part Dam sampling location and lastly C-the use of a bailer to sample water where it’s not physically possible to manually do so (Bruma lake inner sampling location).

Figure 6. Site Visit to Sampling Locations of Region F- the Jukskei River Catchment

This played a major role in the development of the conceptual model, as it provides more understanding of how the sampling procedure works and its overall requirements from tools to administration and labor that go into water quality monitoring and overall water resources management and protection.

3.5.2 Design of optimized sampling locations

To decide on representative sampling sites, the RML method introduced by Do *et al.* (2012) is used together with topographical features. In the RML method, rivers are divided into sections, which are equal to the river mixing lengths. The relocation of the sampling points will form the last part of this research. New sampling points will be established based on information from the brief by the EISD, the RML estimations; taking into consideration the cost constraints, land use land cover, hydrological processes and pollution sources (Wang and Yin, 1997; Telci *et al.*, 2009).

3.5.2.1. Hydrological Network Extraction

The Arc Hydro tool is used to extract the Jukskei river catchment rivers; the algorithms used are all under terrain processess. The specific steps followed are described in detail below:

1. Sink Fill

Firstly, the sink fill algorithm is used to fill or cover small imperfections that might exist in the DEM (Digital Elevation Model). This step modifies the elevation value to eliminate problems such as having a trapped cell without a value amongst other pixels with digital numbers.

2. Flow direction

Secondly, the flow direction computation is performed; this step calculates the flow direction for each grid cell. The digital value grid indicates the direction of flow of the river channel.

3. Flow Accumulation

Thirdly, after determining the direction of the river flow, the flow accumulation is computed. The flow accumulation algorithm calculates the accumulated number of cells upstream of a cell, for each cell in the input grid. In this step, the rivers are defined from the DEM by applying a threshold value using the raster calculator. A smaller threshold value results in denser rivers and usually in a greater number of delineated catchments.

4. Stream Definition

Fourthly, for the definition of the river appearances i.e. the delineation of the river; the stream definition function in the Arc Hydro toolbar is used. For this research, a 500 threshold (defines the cell values that are determined to have enough accumulation to be classified as a stream) is used.

5. Stream to Feature

Lastly, the streams extracted are converted from raster to vector data structure.

The SRTM DEM tile used was downloaded from the USGS Earth Explorer website (https://earthexplorer.usgs.gov/) and the streams are extracted using the Arc-hydro tool following a standard procedure (Figure 7).

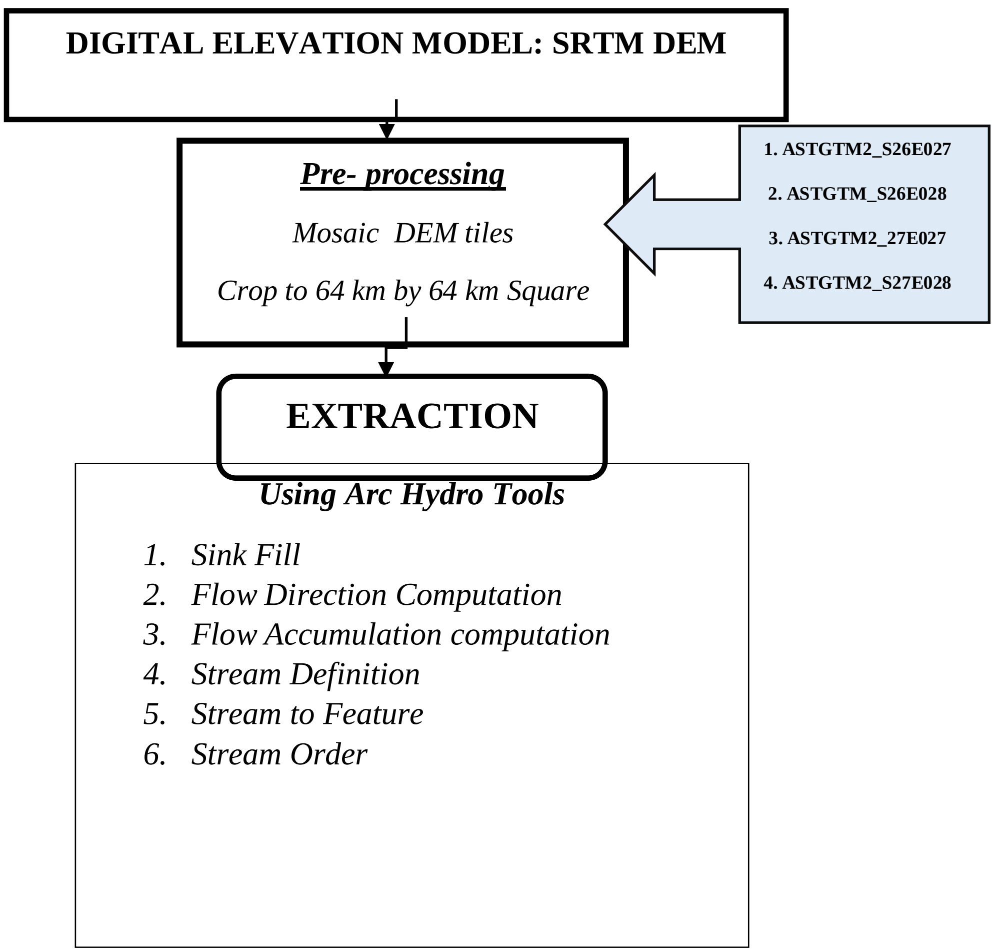


Figure 7. DEM extraction procedure

3.5.3 Sampling scenario
3.5.3.1 Non-Urbanized Scenario

a. Sharp's procedure and Sander's application

The Sharp's procedure (Sharp, 1971) and Sanders' application (Sanders, 1980) methods have been used in many studies to locate representative sampling points (Park *et al.*, 2006; Strobl and Robillard, 2008). However, these methods can only be applied in large river basins and are based on the number of tributaries. If a river system has few or no branches, and the length of each branch is substantially different, these methods are not entirely relevant. In addition, due to the method's dependence on the number of tributaries, sampling point locations cannot be identified in detail. For each river or branch, the position of upstream, middle stream or downstream point cannot be determined.

To solve this limitation and to provide additional advantages by integrating with GIS, a new procedure is developed to locate river basin sampling points (Do *et al.* 2012). In this method, rivers and branches are divided into small segments, and the length of each segment is equal to the river's mixing length. Therefore, each river's mixing length is considered a potential sampling point to monitor changes in water quality.

b. Stream Order

Stream order is described as a hierarchy of networks from the river source which was proposed in 1952 by Arthur Newell Strahler (Shreve, 1996); he explained it as a way to define the size of a river network (ThoughtCo, 2017). The aim of stream ordering is to allocate a numeric sequence to links in a river (Pro.arcgis, 2017). Water is a crucial resource yet limited, thus increasing the need for accurate measures of its availability and analysis. This is to say that stream order is a very important algorithm in the analysis of water extent using sub-watersheds (Mishra and Nagarajan, 2010).

The Strahler Method is used for assigning stream order to a river system. Stream order only increases when streams of the same order intersect. Hence, the intersection of a first-order and second-order link will remain a second-order link, rather than creating a third-order link. The stream ordering process is carried out in ArcMap 10.5 software under the hydrology add-in function (Figure 8).

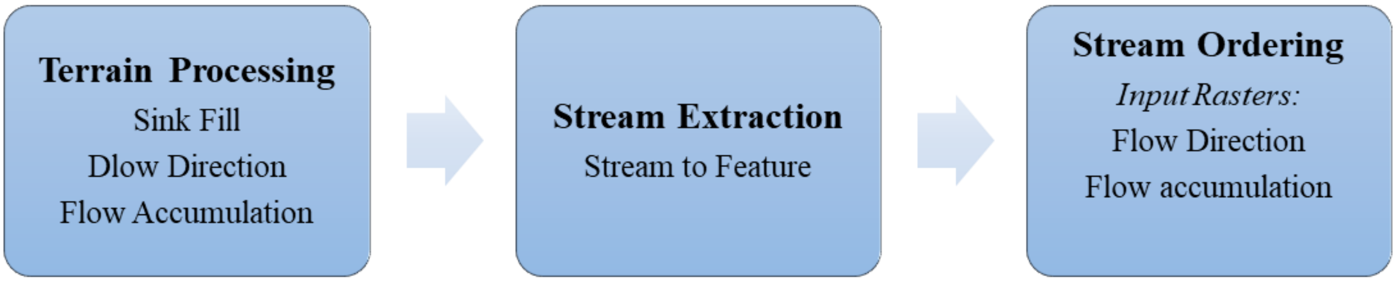
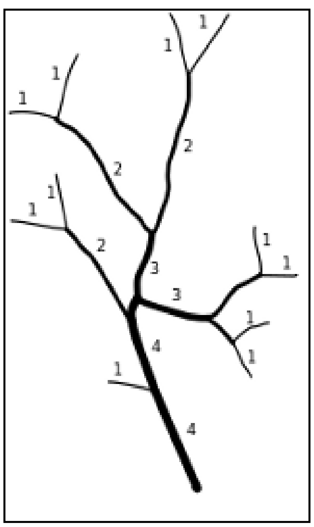


Figure 8. The process followed to obtain Stream Order in ArcMap



In this research report the Strahler stream ordering method is used:

The Strahler's stream order is a modification of Horton's stream order (1945); in this method the main river is not determined, but the ordering is based on the order of the tributaries. This ordering method proceeds sequentially downstream (Anon, 2017).

c. River Mixing Length

The River Mixing Length (RML) theory was first proposed and developed in the early 20th century by Prandtl (1926). RML was described as the distance over which a fluid parcel (upstream) will maintain its properties before it mixes with surrounding fluid (downstream) (Prandtl, 1926; Alilou *et al.,* 2018). In water resource research, determining mixing length and mixing zones is a very important task for identifying representative water quality sampling points (Sanders *et al.*, 1977).

To determine representative sampling locations for the non-urbanized Jukskei River Catchment, the RML method, which of Do *et al*. (2012), was used. In this approach, rivers and branches are divided into sections, which are equal to the river mixing lengths. For this method, it has been proposed amongst others, by Day (1977), that the centre of each section can be considered a sampling location.

A river system layer was added to a study area map, and the river system was divided into small segments using the split function in ArcMap. River network must be decided with high precision, any major branches should not be omitted so that all major branches are treated equally (Day, 1977). The length of each segment is equal to the river mixing length (Day, 1977). It is more appropriate to use fixed mixing lengths for river channels with similar hydraulic characteristics, yet when these characteristics change considerably along the river, the computations should be revised to conclude with a different length (Do *et al.*, 2012).

The RML of each branch of a river or river is determined by using a single geometric parameter, the mean flow width, using the equation, L = 25 W (Day, 1977; Do *et al.*, 2012). The first step, then, is to use Google Earth to measure the stream width because of its good spatial resolution of 15 m–1 m (Telci *et al.*, 2009). It should be noted that the width of a river varies with discharge; this is to say the wetted area of the channel is much wider during a stormflow. Secondly, to estimate the total number of sections of a branch or river; equation (1) was then used. In this research, ArcGIS 10.5 was used to segment the river into different lengths which are equal to the river mixing length. Finally, using equation (2) sections for an entire river network or the total number of potential points is achieved (Do *et al.*, 2012).

$$N_j = \frac{l_j}{L_j} = \frac{l_j}{25W_j} \qquad (1)$$

$$N = \frac{1}{25}\sum_{j=1}^{n}\frac{l_j}{W_j} \qquad (2)$$

Where N_j is the total Number of segments of a river j; l_j is the total length of river j; l_j shows river's mixing length of each segment; W_j is the Width of the stream and N is the total number of potential sampling points of an entire river system.

Thirdly, based on current sampling locations and budget constraints, a method introduced by Sanders (1983) is used to determine the number of sampling locations needed for the Jukskei River Catchment (this method is mainly used to determine the minimum viable number of sampling locations).

$$S_i = 2^i\text{-}1 \qquad (3)$$

Where S_i represents the number of stations and i is a natural number (positive integer) representing the hierarchy of sampling points. A high hierarchy value point has a lower priority than a low hierarchy value point in selecting sampling points (Sanders, 1983)

Calculation:

$$40 = 2^i\text{-}1$$

$$41 = 2^i$$

$$i = \frac{log 41}{\log 2}$$

$$i = 5$$

In the fourth step, the location of the 40 points with different i^{th} hierarchy values was determined. Henceforth, equations (4) and (5) are used to identify major centroids where i^{th} hierarchy point is located in a segment whose magnitude is closest (Do *et al.*, 2012).

$$M_i = \frac{N - k - 1}{2} = \frac{(\frac{1}{25}\sum_{j=1}^{n}\frac{l_j}{W_j} - k + 1)}{2} \qquad (4)$$

$$M_i + 1 = \frac{M_{i+1}}{2} \qquad (5)$$

Where k is the total number of junctions and M_i is the RML magnitude at the ith hierarchy….

Lastly, potential sampling location layer was added to the current study area map. Efforts for determining sampling locations more precisely inside the divided sections would be more time consuming and would need more detailed data for better positioning. Hence the approach of locating potential points in the middle of segments was proposed as a good estimate at the given level of precision (Do *et al.*, 2012).

A potential sampling point was located in the middle of each divided section, and each point was designated with an identifier name, and a type of point. Two types of points are used: NM points, located in the main river, and NS points, located in tributaries. These points are thus placed onto the stream using the length of the stream.

d. Stream width average

The measured stream widths measured on Google Earth imagery are averaged. The standard deviations are calculated; mainly to show how much the data is spread from the mean. A low standard deviation shows that the data points tend to be very close to the mean. A high standard deviation shows that the data points are spread out over a large range of values. This calculation helps to gauge the accuracy of the proposed model.

e. River Segmentation and Number of Points Estimation

Following the method used and by Alilou *et al.* (2018), using all the stream orders extracted for the Jukskei River Catchment, the number of sampling points for monitoring of the respective stream orders and thus the entire catchment is estimated for all the delineated stream orders. The initial step of this method is the estimation of the number of sampling locations required for the Jukskei river catchment; this is mainly based on the topography and river characteristics. This is followed by determining the river flow using the RML equations as discussed in section 3.6.1.2 above.

3.5.3.2 Urbanized Scenario

a. Integrating human activities

The GIS-based suitability model is used as the foundation for integrating LULC into the main model for optimizing sampling location. Criteria such as population density, major roads, bridges, and manufacturing areas are often used to assign optimal sites for water quality monitoring. The suitability criterion is thought to be strongly related to water pollutants and consequently the most important in assigning optimal sites for monitoring points (Chang and Lin, 2014).

The approach of this research is to identify sites where most of these criteria have a great potential influence using GIS suitability analysis.

Suitability analysis is described as a raster-based GIS analysis made up of five main steps (Alsahli and Al-Harbi, 2018):

1. Criteria determination
2. Input layer determination and preparation
3. Layer rank and weight
4. Suitability map production
5. Model output evaluation.

The criteria are determined and accomplished before any analysis. Selecting criteria for assigning water quality monitoring points is grounded on previous studies that explored the role of numerous variables in water quality decline most especially in built-up areas (Do *et al.,* 2012). The necessary layers for suitability analysis are specified based on the selected criteria under criteria determination. The layers of major interest are graded and then weighted based on their priorities (Alsahli and Al-Harbi, 2018). Though the priority of layers is subjective, this hierarchy is chosen based on the expected influence of these criteria on water pollutants within the catchment, as reviewed from previous studies and from analyses documented by the department responsible (Larssen *et al.*, 1999; Raffuse *et al.*, 2007).

A land use map layer acquired from DRDLR (2017) is divided into 10 categories: barren land, built up, cultivated, grassland, mines and quarries, natural wooded land, planted forest, scrubland, waterbodies, wetland. However, for the scope of this project, the cultivated land and planted forest are grouped under agriculture and grassland, while scrubland and naturally wooded land are grouped under vegetation.

b. LULC within the Jukskei river catchment

The urbanized scenario takes into consideration the spread of the population and all the social activities that are related to it and hence contribute to the decrease in water quality; the following LULC types will be explored in order of significance; high to low (Figure 9 and Figure 10).

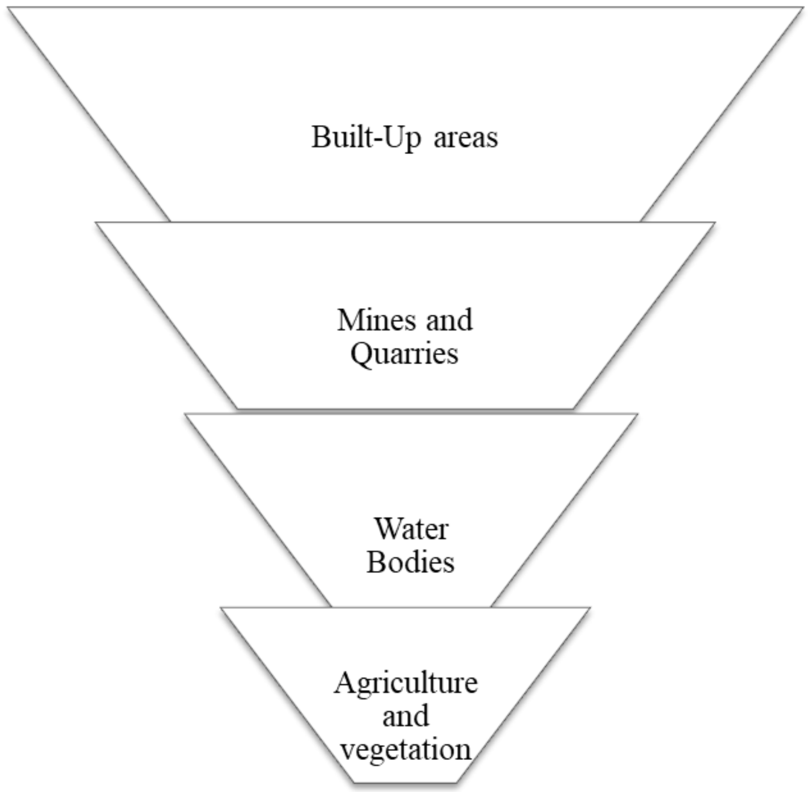


Figure 9. Factors considered in order of importance

Based on visual knowledge of the study area and information provided by the EISD, the order of priority was decided (Figure 9), with the main concern raised being the built-up area; made up of residential areas, industries and active infrastructures (roads and railways). Figure 10 shows a detailed map of the LULC within the delineated Jukskei River Catchment obtained from the Department of Rural Development and Land Reform LULC maps for 2017.

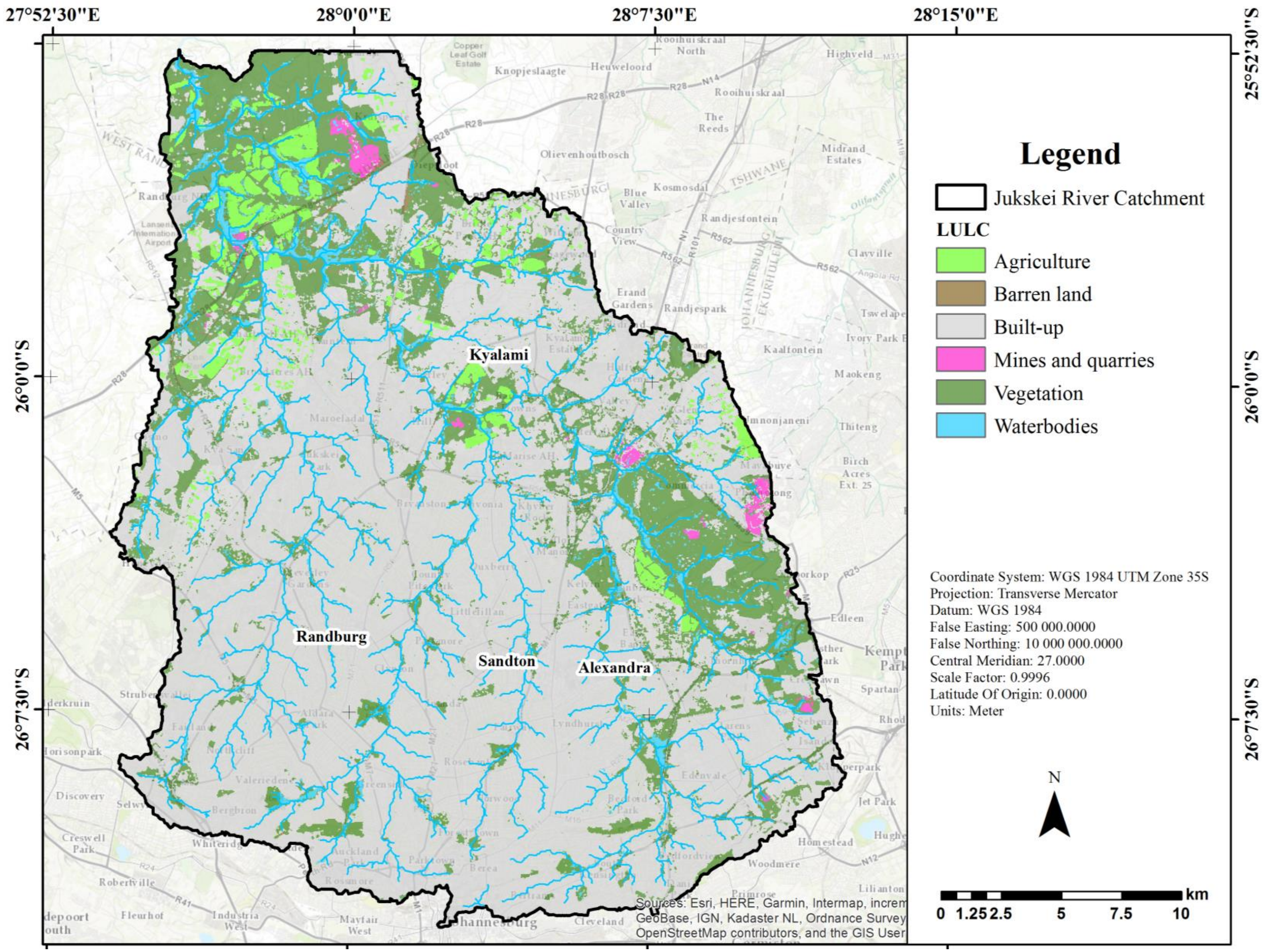


Figure 10. Land Use Land Cover Map (DRDP, 2017)

c. Contributing Area

Sivertun and Prange (2003) suggested that pollutants produced at more than 1000 metres from the river banks cannot reach the river and thus cannot have any influence on the water quality of the river. Consequently, for this research, the contributing area is taken as 1000 m from the river. This area is determined using a buffer (Figure 11). This assumption of a 1000 m buffer is controversial mainly because urban areas often have existing underground stormwater drainage, and this is not easily accounted for by the 1000 m buffer. For this study, the buffer is extended 1000 m from the Jukskei catchment to avoid a sharp cut of the LULC classes because what happens in one class affects the activities in neighbouring LULC classes.

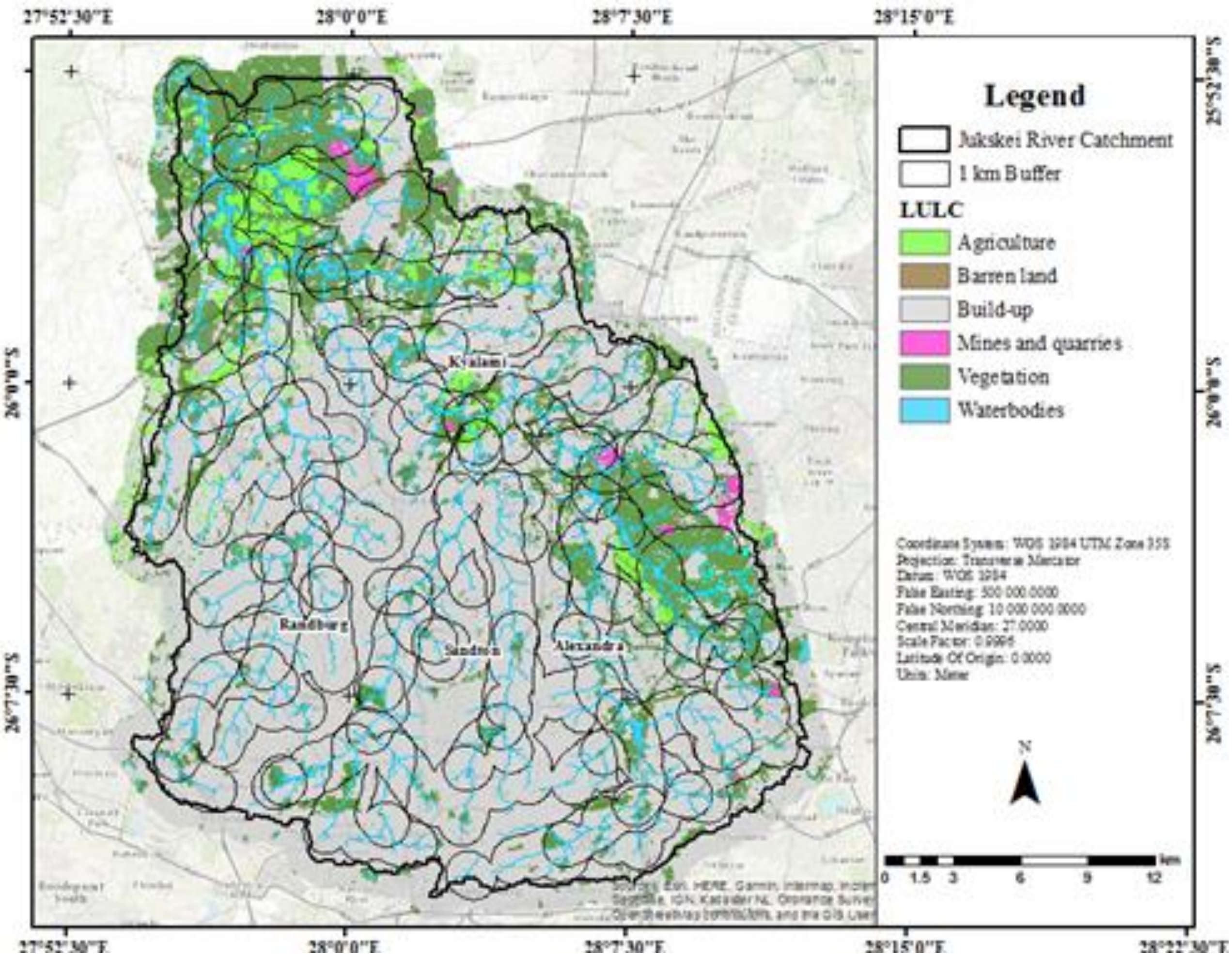


Figure 11. Land use land cover within the 1000 m buffer (adapted from DRDP, 2017)

d. Relative potential pollution weight for non-point source

Non-point pollution sources have a major influence on the quality of water (Baird *et al.*, 1996). The areas with higher potential pollution may affect water quality and should be strictly monitored (Chang and Lin, 2014). Here, the Analytic Network Process (ANP) method is applied as a multi-criteria assessment to determine relative potential pollution weights for non-point sources. Among multi-criteria decision-making approaches like the Data Envelopment Analysis DEA, the ANP method is the most appropriate method (Kucukaltan *et al.*, 2016), as it considers the criteria's dependencies and the calculation of their relative weights (Lin *et al.*, 2009).

Subsequently, the event mean concentrations of each non-point source are used to precisely calculate and determine the relative potential pollution weight for non-point sources, which are located in the contributing land use type (Table 5).

Table 5. Extracted weights based on AHP and fuzzy standardization for urban areas (weight based on study by Kucukaltan et al., 2016)

Factors	**Function**	**Control Point**	**Weight**
Slope	Sigmoid	0% highest suitability 0–15% decreasing suitability >15% no suitability	0.19
Distance from Road	J-shaped	0–10 m highest suitability 10–400 m decreasing suitability > 400 m no suitability	0.28
Distance from Water Bodies	Linear	0–10 m no suitability 10–1000 m increasing suitability > 1000 m highest suitability	0.15
Distance from built-up areas	Linear	0–100 m highest suitability 100–2400 m decreasing suitability > 2400 m no suitability	0.38

e. ANP Method

Non-point pollution sources have been proven to have a noteworthy impact on river water quality and thus require intensive monitoring (Baird *et al.*, 1996). To achieve this, the ANP method was applied to several studies and has proven to give accurate and useful estimates and predictions (Saaty and Vargas, 2006).

In this research, 6 non-point sources - agriculture, barren land, built-up area, vegetation, mines and quarries and water bodies (as illustrated in Figure 10) were used as criteria. The relative weight of each criterion is achieved by the ANP Method using ***Super Decision Software*** (Figure 12). The ANP method is one of the most powerful synthesis methodologies for combining judgment and data to effectively prioritize options and thus predict outcomes.

The Super Decision Software is decision support software that implements the ANP method amongst others, and it is mainly used for decision-making with dependence and feedback. This software provides tools for the creation and management of different models, prioritization, obtaining results and performing sensitivity analysis on the results. The ANP scoring is used in this research and the process followed is adopted from Alilou *et al.* (2017). The scoring process gives each LULC a level of importance or priority based on the relationship or impact it has on the water body compared to the other LULC classes; this is done on a scale on a 1 to 9.5, where 1 is a low level of importance and 9.5 is a high level of importance compared to the other class as demonstrated in Figure 12. The LULC hierarchy in the Super Decision software is determined by the user: this prioritizing is founded on the review of literature and most significantly from water quality concerns of the study area.

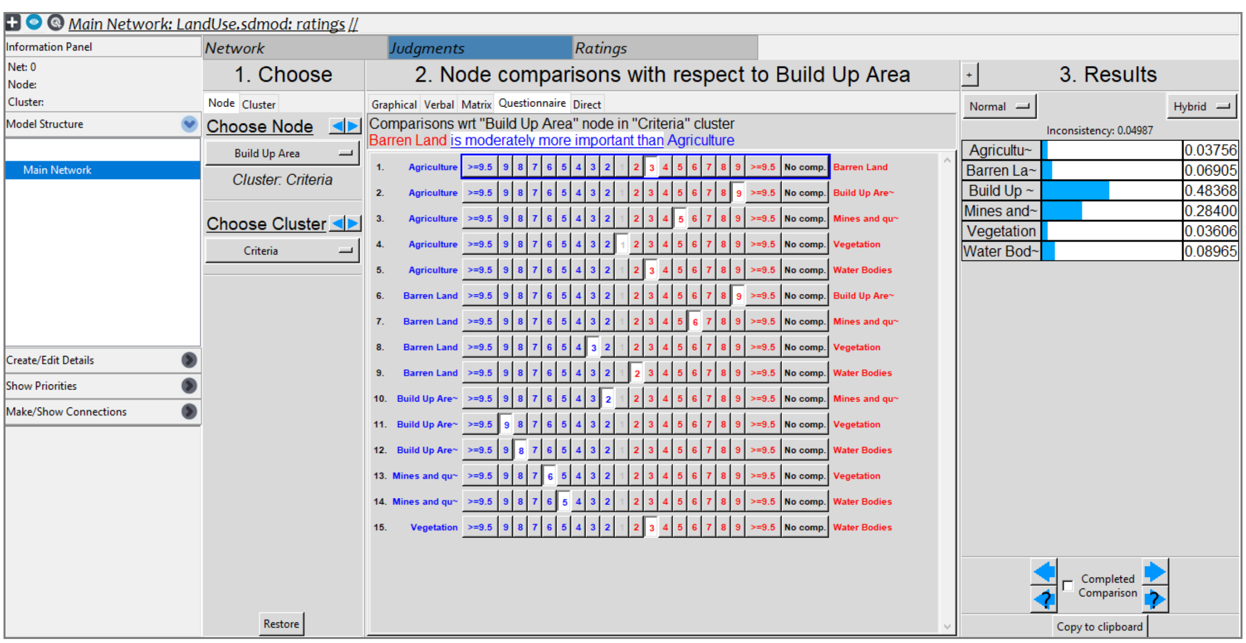


Figure 12. ANP Method Procedure (SuperDecision Software)

f. Scoring Candidate sampling points

From the estimated required number of sites, using the ANP weighting, potential sampling points are scored for better prioritization of contributing factors and thus achievement of optimal sampling locations for water quality sampling. Equation (6) is used.

$$NPPS = \sum_{i=1}^{6} W_i * A_i \qquad (6)$$

Where the NPPS is the Non-Point Source Pollution Scores, W_i and A_i are the potential pollution weight and percentage of each non-point source between candidate points respectively

Finally, the most appropriate points are selected based on hierarchy values (Sanders, 1983); where a low hierarchy value indicates higher priority just as a sampling point located in an area of high human activities has high priority and therefore must have more monitoring points (Do *et al.*, 2012).

The methodology outlined in this research proposes the use of the RML technique for addressing two vital components of pinpointing river water quality sampling locations, the systematic design, and human activities. RML is introduced to modify and further develop Sharp's procedure and Sander's application (1983) for selecting representative sampling points. To select the most appropriate sampling points, human activities and LULC layers are combined to optimize the locations of sampling points. Geographic Information Systems (GIS) are used to integrate human activities into a river basin through LULC mapping. The Potential Pollution Score (PPS) model is proposed in this research for locating high pollution catchments.

4. Results

4.1a Users' needs assessment

The results from the users' needs assessment survey are summarized in Table 6. Section 4.1b describes the conceptual model which was created through the answers received in the interviews conducted and site visit.

Table 6. Interview results

	Questionnaire	Responses	Respondent	Date
1	What is the sampling Strategy?	Delegation to different departments; JW, EISD	The brief provided by Mpho Zwane in a PDF Document.	01 November 2018
2	When were the sampling points established?	2013	Brief Provided by Mpho Zwane In a PDF Document	01 November 2018
3	How were the sampling points established?	The EISD selected the locations on mere convenience	Brief Provided by Mpho Zwane In a PDF Document	01 November 2018
4	Has there been a change in sampling points since established?	No, Insufficient funds	Mpho Zwane Nandy Manzini	25 June 2018
5	Main Challenge with Data Capturing	Misplaced Sheets Incomplete/ missing data Inconsistency in the format of capturing	Emma Netshihulu Mpho Zwane Nandy Manzini Pule Makena	25 June 2018 17 October 2018
6	Backup Database (X)	None	Mpho Zwane	25 June 2018
7	What kind of platform would you prefer for the GIS platform (X)	Web-based and public (with roles and privileges)	Mpho Zwane Emma Netshilulu	25 June 2018
8	What main function will the GIS platform perform? (X)	Record Data Share Information Perform statistical analysis Visualize and Model Data	Mpho Zwane Sammy Katuma	25 June 2018
9	Water Quality Concerns	Pollution The decline in water quality	Mpho Zwane Clarence Nandy Manzini	06 December 2018

4.1b Conceptual Model

The conceptual model created is designed to exhibit the different resources needed by the EISD for the management of water quality and the overall datasets that are needed for the creation of a water quality database (this includes the resources, sources, dataset types, the person responsible). This prototypical conceptual model gives an outline of the different functions that make the management of water quality possible and plausible.

Figure 13 of the model design represented graphically are the thematic layers needed for the management of water quality of the river catchments. These can be gathered from different institutions; both governmental and private as illustrated through the sources of this research.

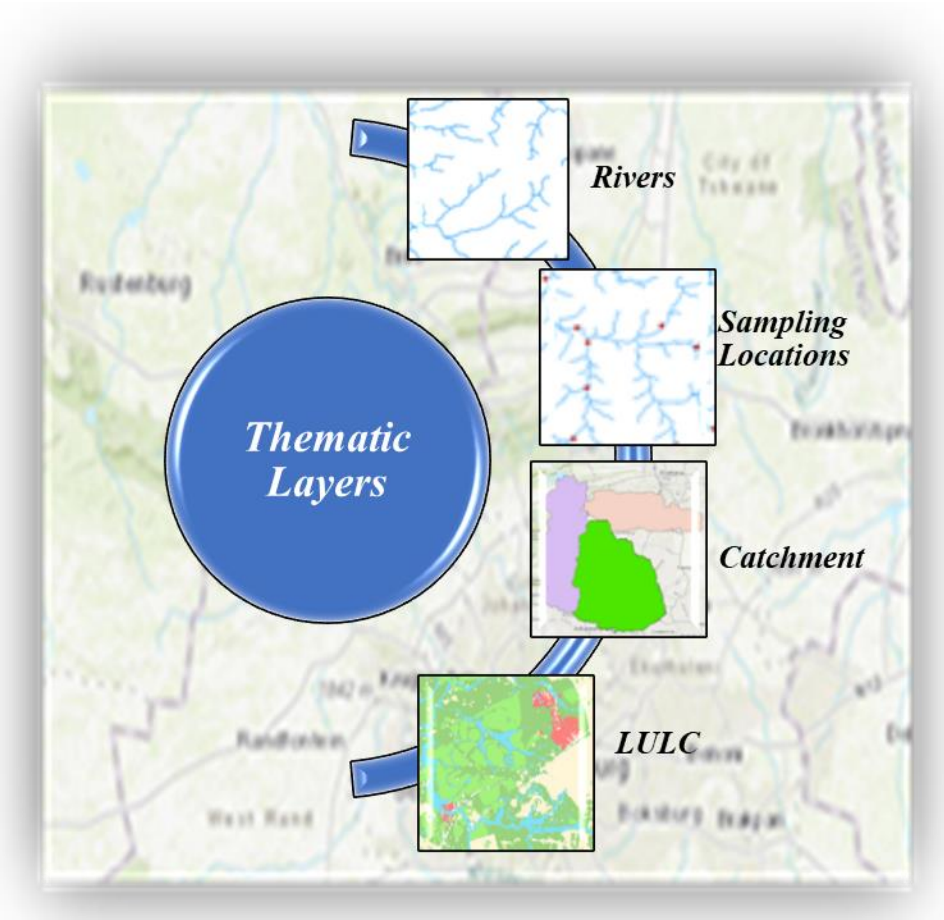


Figure 13. Thematic Layers required for the monitoring of water quality of rivers

Entity Relationship

Figure 14 shows the design of the relational database that can be used for the administration purposes and thus effective management of the water body. It gives an illustration of how the different datasets can be collected and organized in a database system for better relation, extraction, analysis, and sharing. It includes several tables thought to be significant in achieving its aim (managing water resources and thus quality).

- The infrastructure table will save information including that of road networks (which play a major role in the accessibility of sampling locations) and bridges (which have been the focus of the EISD in the location of sampling locations).
- In the database, GIS is incorporated in the table DEM where the water bodies will be delineated together with topography and terrain; with important information such as time and date which will assist in terms of seasonal analysis of water characteristics; these can be grouped and easily organized and related to the LULC and waterbody type.
- For the ease and quick access of data, the database structure is designed to allow searching per region or district and municipality, allowing ease of information capturing and hence analysis. Most useful to dataset source tables are the metadata tables which form one of the most crucial parts of secondary data.
- Additionally, more relevant to this current research are the sampling, observation and laboratory tables; these are responsible for keeping a record of the samples collected and tested and the person who handled their analysis.

This model illustrates the ease with which water quality information can be managed and thus easily simulated and distributed to both public and important stakeholders (Figure 14). It gives the different links that exist between different resources and infrastructures. In orange, are two tables that need special attention as they will give specific information on the changes of the water quality over time i.e. the sample results

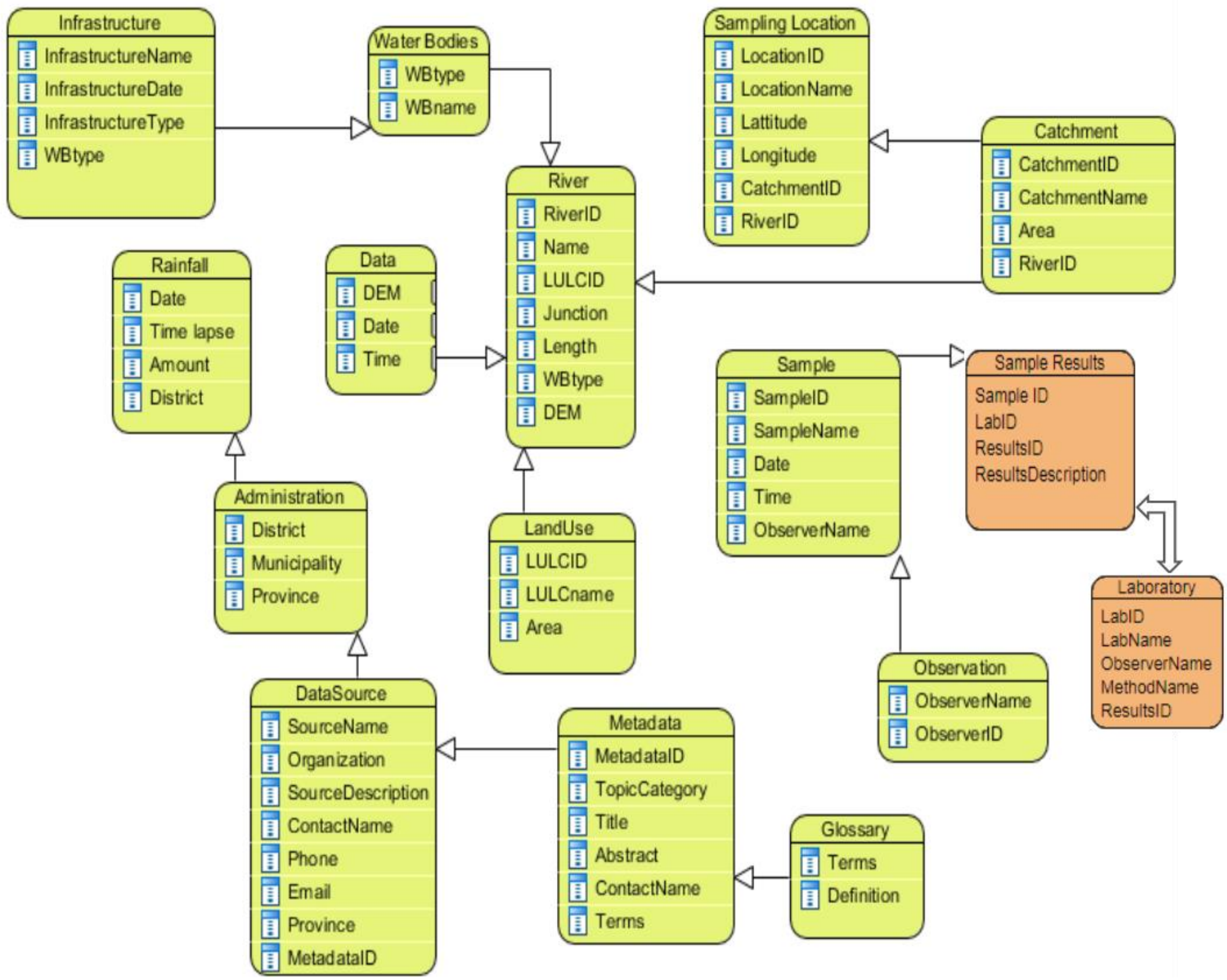


Figure 14. An ER diagram showing the requirements for monitoring river water quality

DEM Extraction

This process involves the delineation of river networks from a digital elevation model, chosen based on the needs of the projects and the characteristics aimed at for the project. Figure 15 shows the DEM (left) and the streams that were extracted from the model (right); the streams extracted can vary based on the threshold one uses.

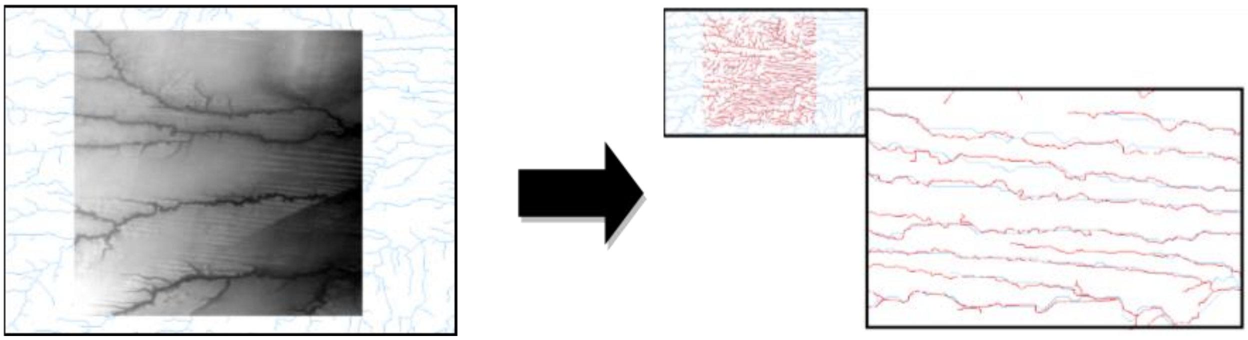


Figure 15. Input and Output of DEM Extraction

Drainage System

Figure 16 is an illustrative example of the process it takes to delineate catchment areas. This approach gives proper direction and helps with the planning of catchment management and thus accurate water quality monitoring. This process begins with the extraction of river networks from a dem (Figure 15 and from these the catchments can be delineated)

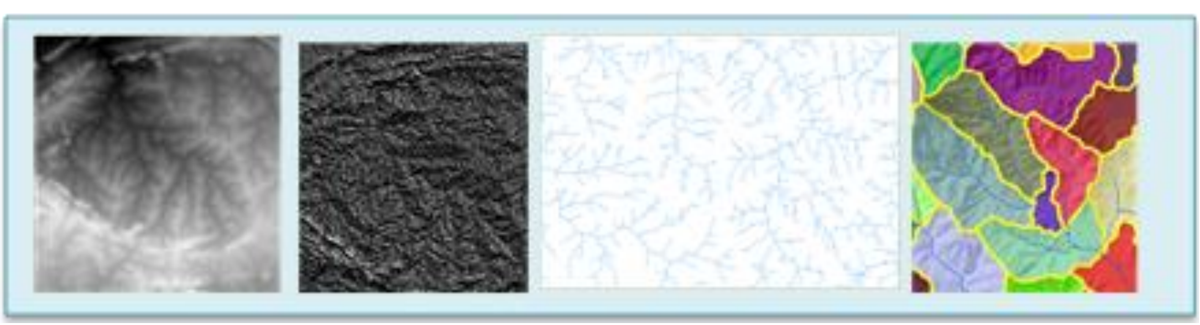


Figure 16. The delineation of river catchments

4.2 Optimization of sampling location

4.2.1. Delineated Jukskei River Catchment

The catchment provided by the CoJ is compared to the one delineated using an SRTM DEM with the Arc Hydro Tool (ArcMap Add-in). Based on the delineated catchment, on visual exploration, it is observed that the newly calculated Jukskei catchment is different from the one the CoJ has been using (Figure 17). From the appearance of these two maps, it appears that the catchment was previously delineated based on distance visualization; streams near the Jukskei river that are not necessarily tributaries connected to the river directly and thus do not feed into the Jukskei catchment were defined within the boundaries of the catchment, and because of this the outlined catchment area by CoJ is inaccurate. Again, on visual exploration, it is observed that the catchment currently being used is for the Crocodile River upstream of Hartbeespoort Dam and the Jukskei river is a tributary of the Crocodile River with its own sub-catchment.

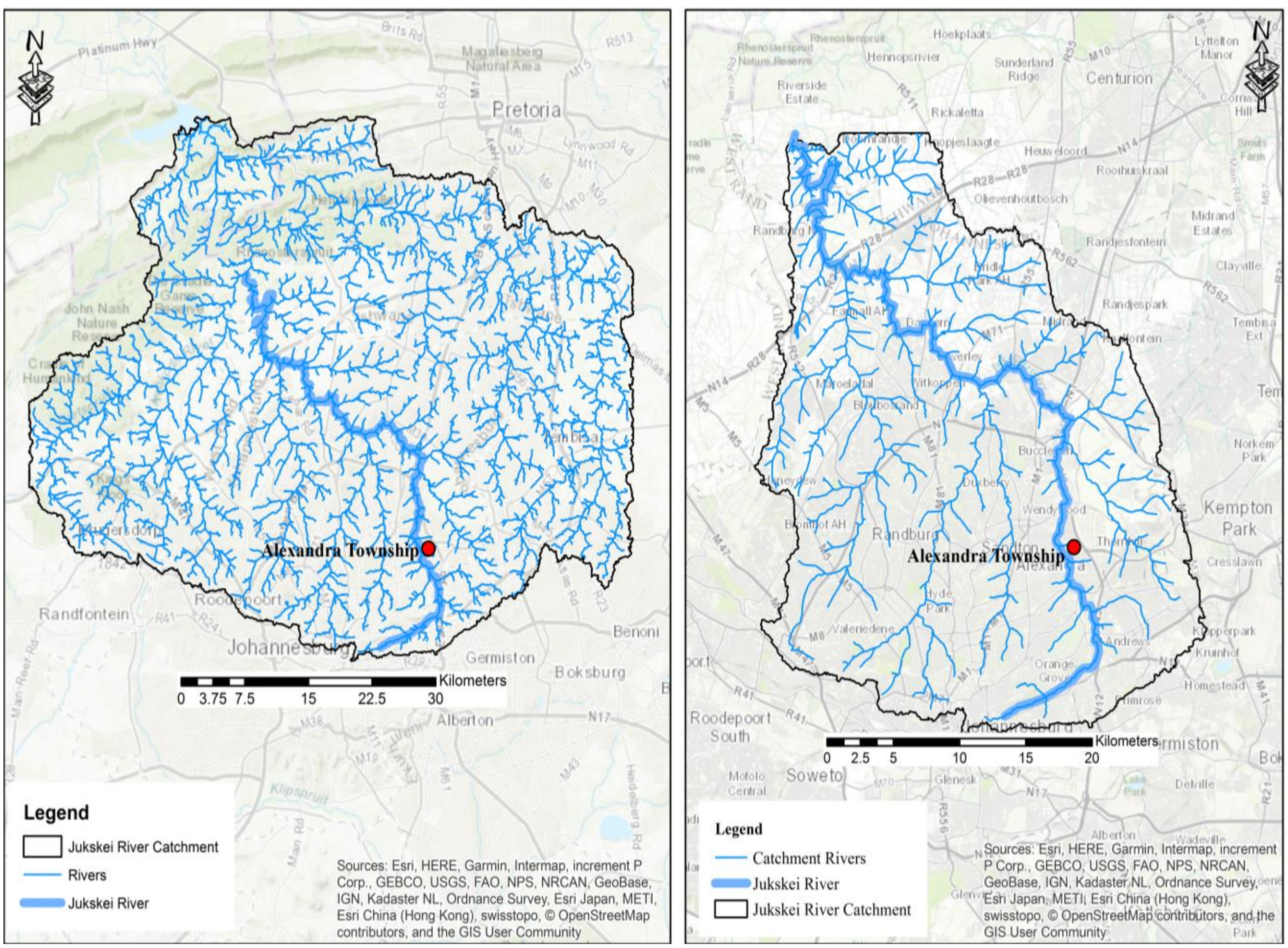


Figure 17. Jukskei river catchment (provided vs delineated)

Taking a closer look at the Jukskei River Catchment the CoJ is currently using for water quality sampling (Figure 18), it is noted that not only do the points cover the Jukskei but also cover streams and tributaries feeding into neighbouring streams; the Crocodile River and the Hennops River.

The results from the hydrological network delineation are one of the issues the EISD should be looking into as the reason behind erratic water quality data

This might be a reason why the sampling results show trends which do not correlate with what is observed on the ground.

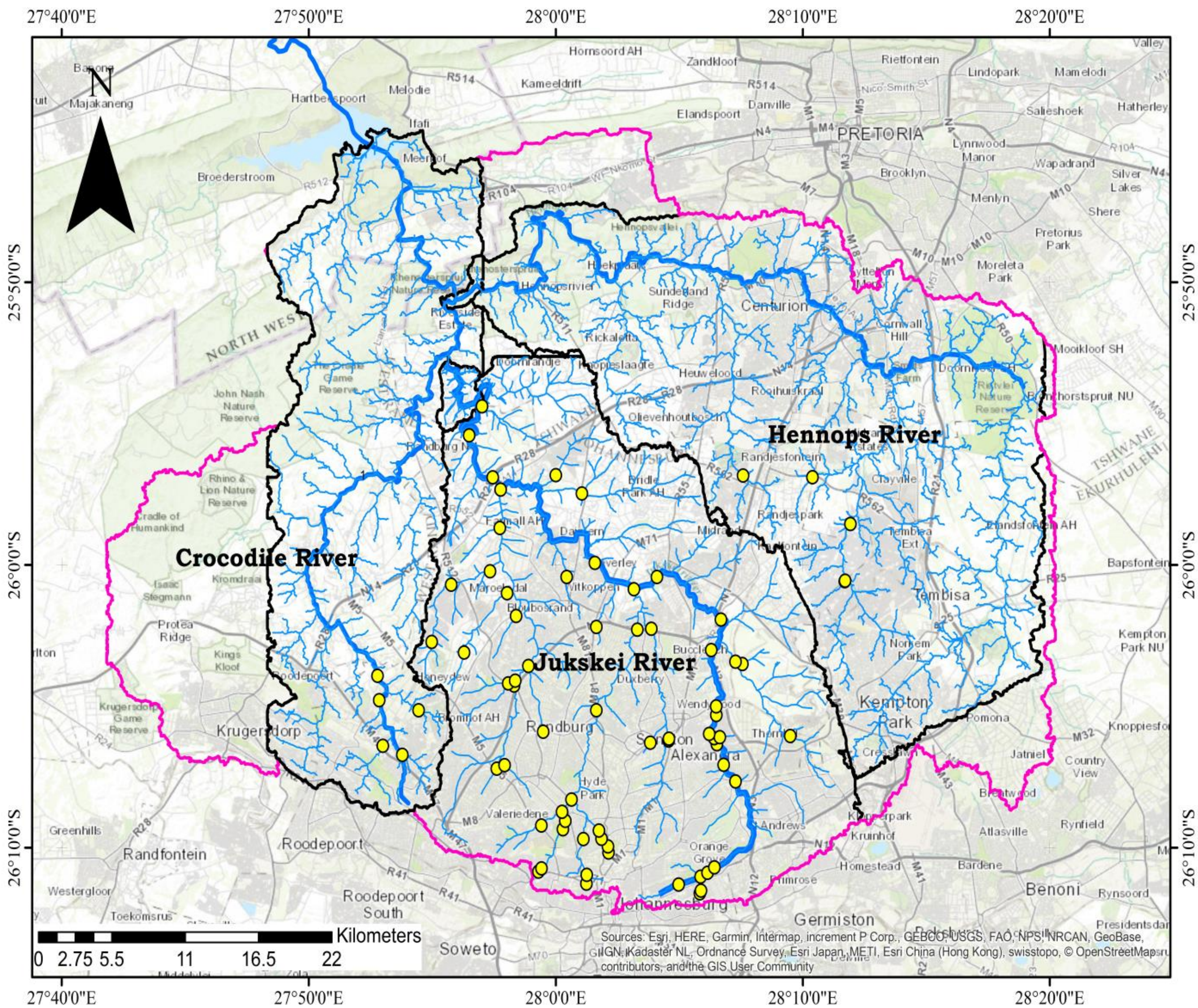


Figure 18. Catchments Surrounding the Jukskei River Catchment (Yellow dots represent the location of the current sampling points)

4.2.2 Location of sampling points

4.2.2.1 Stream Order

The Jukskei River catchment stream ordering resulted in 6 stream orders; the 6th order is the largest stream or the main Jukskei River. From this stream ordering only the major river and major tributaries, specifically four stream orders will be used to propose sampling locations, the 3rd, 4th, 5th and the 6th order (Figure 19).

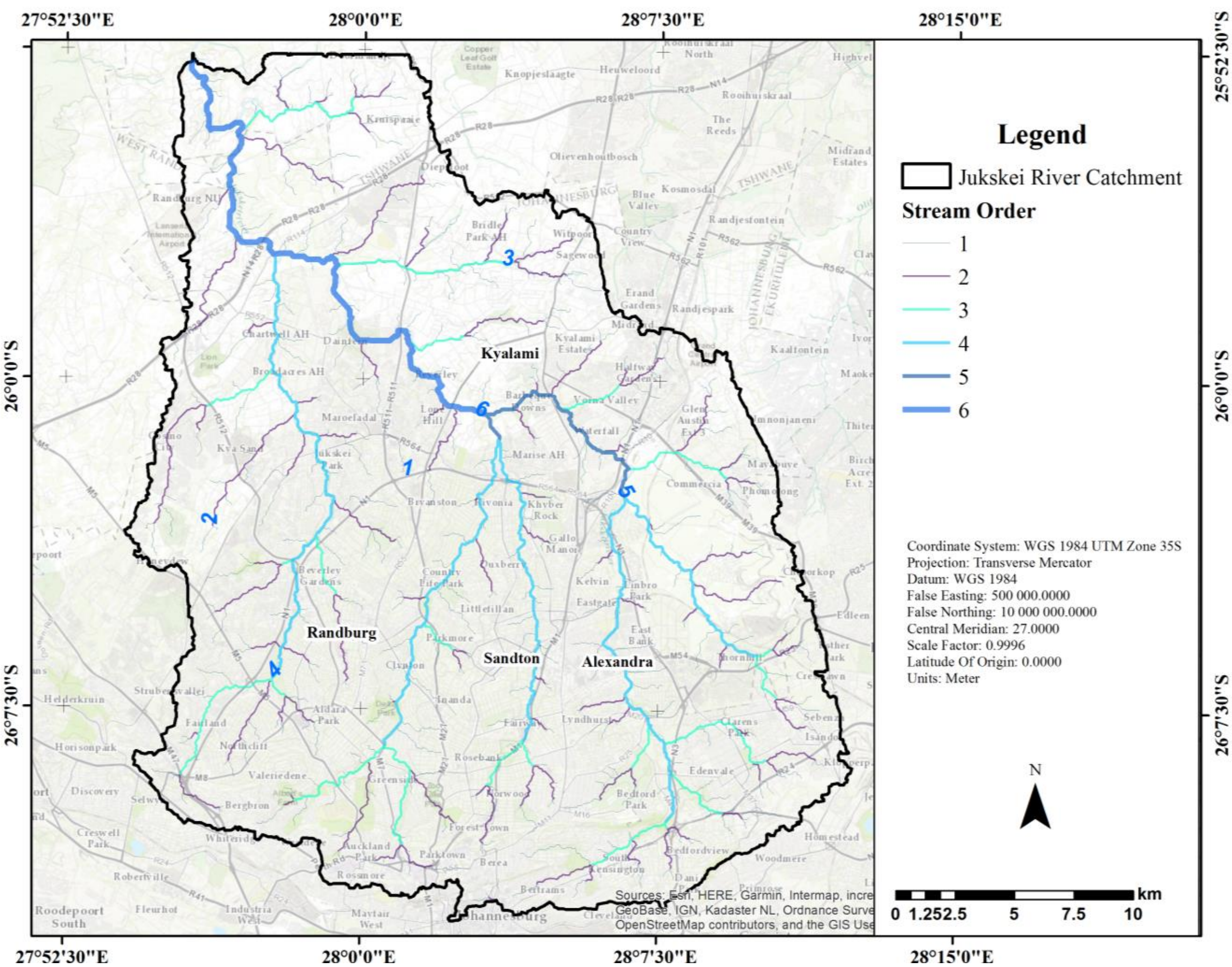


Figure 19. Stream order definition (numbers in blue indicate Strahler stream order)

Table 7 below shows the river mixing length magnitude at i^{th} hierarchy in the Jukskei River stream orders except for the 1st stream order; this is because this is the most upstream order and thus has no junctions or rivers and tributaries flowing into it.

Table 7. River Mixing Length magnitude at the numbered hierarchy

Stream Order	k	M_i
6^{th}	69	1 423
5^{th}	11	1 452
4^{th}	64	1 425
3^{rd}	34	1 440
2^{nd}	149	1 450

[*] ***k is the total number of junctions and M_i is the river mixing length magnitude at the i^{th} hierarchy***

Table 8 shows the measured stream width along the rivers of the Jukskei River Catchment, as per manual measurement using the Google Earth platform measurement tool. Overall the standard deviations acquired shows a wide spread of values along the different stream orders of the different streams feeding into the Jukskei River.

Table 8. Average Stream Width (measured from Google Earth in metres)

6th Order		5th Order		4th Order		3rd order		2nd Order		1st Order	
J11	20,6	J25	11,8	J7	7,2	J57	6,8	J66	1,2	J80	1,0
J12	32,8	J26	23,0	J8	8,6	J58	8,7	J67	1,3	j81	1,4
J13	34,4	J27	19,9	J9	10,6	J59	3,9	J68	4,8	J82	1,1
J14	18,8	J28	17,6	J10	13,6	J60	5,9	J69	3,4	J83	1,9
J15	25,5	J29	19,2	J31	9,2	J61	4,6	J70	4,8	J84	2,3
J16	34,2	J30	13,8	J32	16,1	J62	6,1	J71	2,4	J85	1,1
J17	22,4	**MEAN**	**17,5**	J33	12,5	J63	7,3	J72	1,9	J86	3,2
J18	35,5	**SDT_DEV**	**4,1**	J36	15,3	J64	5,9	J73	2,8	J87	2,2
J19	31,5			J37	7,9	J65	8,8	J74	4,7	J88	3,4
J20	28,4			J38	7,5	MEAN	6,4	J75	3,6	J89	1,7
J21	23,6			J39	8,7	SDT_DEV	1,6	J76	6,0	MEAN	1,9
J22	19,1			J40	9,8			J77	2,3	STD_DEV	0,8
J23	23,5			J41	9,3			J78	5,7		
J24	19,2			J42	9,4			J79	4,6		
J35	32,9			J44	21,0			MEAN	3,5		
MEAN	**26,8**			J45	7,0			SDT_DEV	1,6		
SDT_DEV	**6,3**			J46	10,8						
				J47	10,3						
				J48	9,2						
				J49	15,4						
				J50	9,2						
				J52	26,7						
				J53	11,2						
				J54	8,2						
				J55	9,0						
				J56	3,8						
				MEAN	**11,0**						
				SDT_DEV	**4,7**						

In Table 9 a detailed summary of the total number of points estimated for each stream order is presented as per the calculations suggested by Alilou *et al.* (2018); 42 points are estimated for the 6th stream order which makes up the main Jukskei River, 326 for the 5th, 292 for the 4th, 480 for the 3rd, for the 2nd stream Order 2 075 and lastly 8 327 for the 1st order, and in total the estimated number of potential sampling locations for the entire Jukskei River Catchment is 11 242.

Table 9. Total Number of Sampling locations estimated for the Jukskei River Catchment

Stream Order	Average Width (m)	Standard Deviation	Length (m)	Nj	N
6th	26.8	6.3	28194.5	42	
5th	17.5	4.1	11303.5	26	
4th	11.1	4.8	80789.2	292	
3rd	6.4	1.7	77311.7	480	
2nd	3.5	1.6	183636.5	2075	
1st	1.9	0.9	397610	8327	
					11 242

**N_j number of sampling locations estimated for that stream order*
**N Number of Sampling points for the Jukskei River Catchment*

4.2.2.1 LULC Percentage per Stream Order

For all the stream orders the percentage of each LULC type for a 1000 m buffer is calculated and represented using pie charts illustrated in Figure 20. Within all the stream orders a similar trend of dominance is observed as follows; built-up area, vegetation, water bodies (wetlands, dams, and rivers), agriculture, mines and quarries lastly barren land. Furthermore, it was detected in the literature that built up areas are the major contributors to degrading water quality in this catchment (Do *et al.*, 2012), and this is visually and statistically supported by the proportion of the built-up area to all the other LULC classes observed in the catchment.

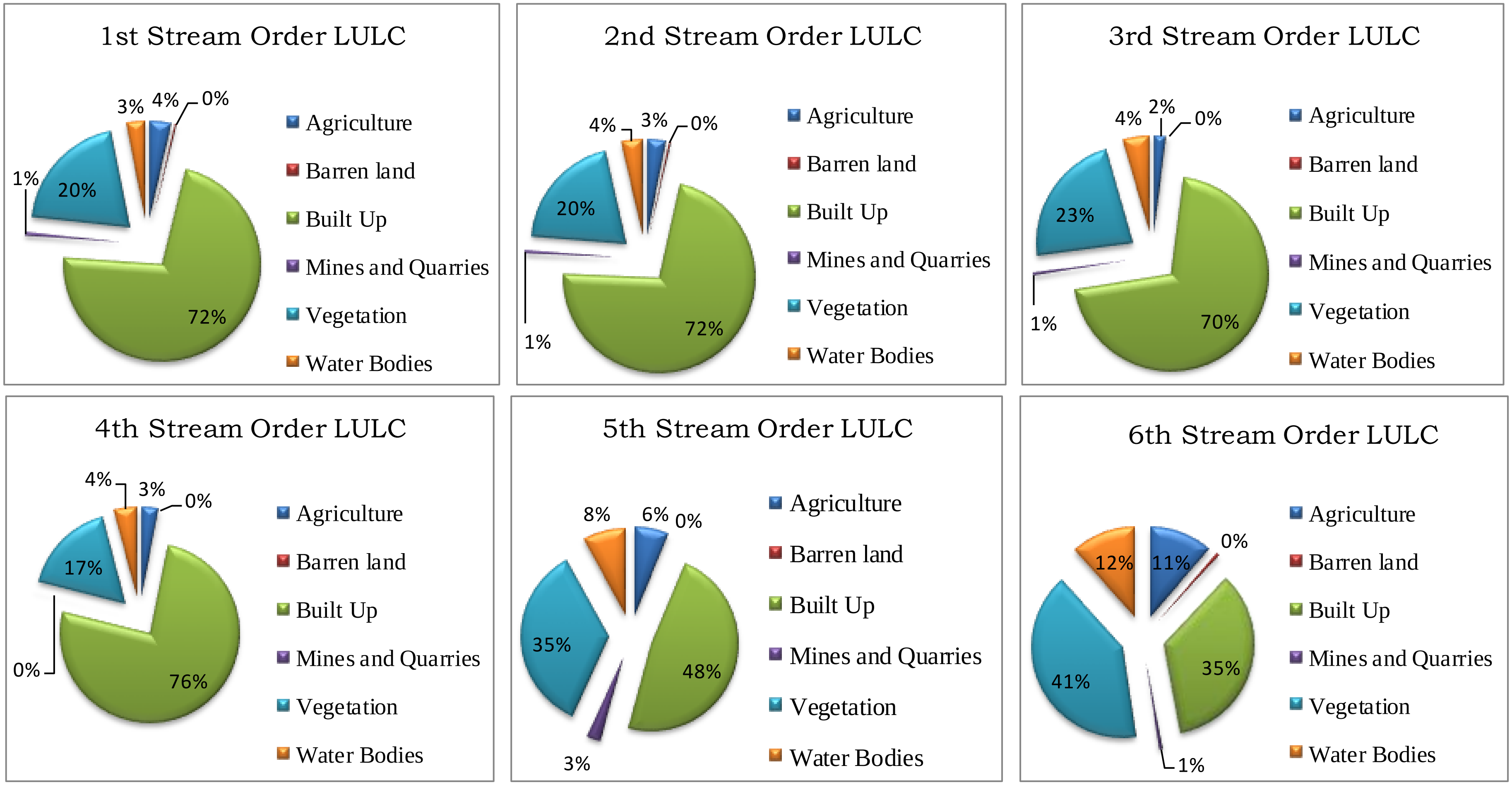


Figure 20. LULC per Stream Order

Alilou *et al.* (2018) performed a similar study using a study location (The Khoy Watershed located in the Azerbaijan province- West Iran) with characteristics similar to the one used in this research (Jukskei river catchment, Johannesburg, South Africa), even though they are similar in terms of being urbanized, differences do exist; Khoy is located in a cold semi-arid area with a mean annual rainfall of 281.94 mm whereas the Jukskei is located in the highveld with a mean annual rainfall of 700 mm. Alilou *et al.* (2018) results illustrated a higher weighting on built-up areas (0.38) which may be relevant to the context of the current study because the Jukskei River Catchment is made up of 72% built-up area (Table 10).

Table 10. Land Use Land Cover Percentage of the Jukskei River Catchment

Class Name	**Count (raster cells)**	**%**
Agriculture	31955	3.6
Barren land	2505	0.3
Build Up **Area**	649405	72.3
Mines and Quarries	5347	0.6
Vegetation	180873	20.1
Water Bodies	28059	3.1
	898144	**100**

4.2.2.2 Extracted Relative Weights Using ANP Method

In Table 11, the relative weights of the land use classes which are obtained using the pair-wise comparison matrix with the ANP method are depicted and in Table 12 the graphical expression of the relative weights of the land uses are displayed as an output from the super decision software.

Table 11. Alternative scoring of LULC

Alternatives	**Priorities**	**Totals**	**Agriculture (0.0188)**	**Barren Land (0.0345)**	**Built Up Area (0.2418)**	**Mines and Quarries (0.1420)**	**Vegetation (0.0180)**	**Water Bodies (0.0448)**
Vegetation	0.1475	0.1979	1.0000	0.3029	0.0917	0.1122	1.0000	0.3029
Built Up Area	0.4045	0.5428	0.0917	0.0917	1.0000	0.1331	0.0917	0.0917
Water Body	0.3228	0.4332	0.3029	0.0917	0.0917	0.9526	0.3029	1.0000
Bare Soil	0.1253	0.1682	0.3029	1.0000	0.0917	0.1122	0.0917	0.0917

Considering the relative PPS values, the built-up area showed a higher score, far more than the other LULC types; the mines and quarries followed (Table 12), the results differ from those of Do *et al.*, (2012), who reported high priority for built-up and agricultural areas.

Table 12. Priority scoring of LULC

PRIORITIES FOR COLUMN OF RATINGS SYSTEM	
Built Up Area	**0.483685**
Mines and Quarries	**0.284001**
Water Bodies	**0.089652**
Barren Land	**0.069046**
Agriculture	**0.037558**
Vegetation	**0.036058**

4.2.3 Mapping of point locations

a. Non-Urbanized Scenario

For the non-urbanized scenario, only the topographical features and the basics of river mixing length are considered. After performing the calculations which indicated 11 242 as the required number of points for the Jukskei River Catchment, 40 Sampling locations are proposed mainly because this number is closer to the number of sampling points used by the EISD for water quality sampling and thus fit within their budget limitations. These 40 proposed locations are divided throughout the catchment using proportional calculation. This is done with a specification of how many per stream order and thus their exact or approximate location (Do et al., 2012; Table 13 and Figure 21).

Table 13.Proposed number of sampling locations per stream order

Stream Order	Proposed Sampling Locations
6th	6
5th	3
4th	11
3rd	20
2nd	0
1st	0

Figure 21 shows the proposed locations of the water quality sampling points of the Jukskei catchment without factoring in the land use land cover associated with the Jukskei River. This is mainly based on the RML calculations which distribute sampling locations based on the segmentation of the rivers and tributaries.

Besides the RML, the number and location of the existing sampling locations are taken into consideration when proposing the sampling locations. From the estimated number of sampling locations for the catchment, a percentage is calculated for each of the stream orders; to determine how many points to allocate to each stream order based on the proposed number of 40 locations.

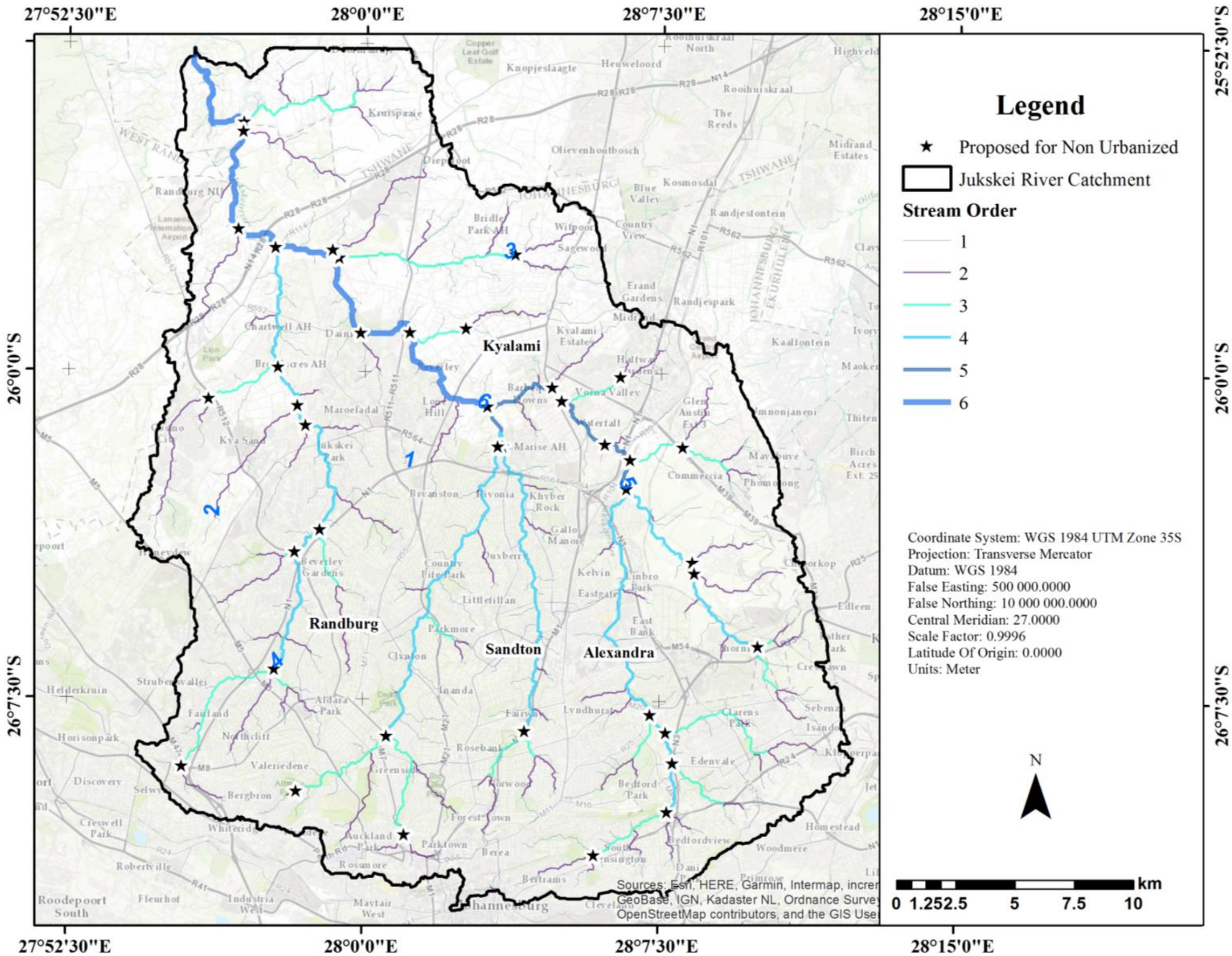


Figure 21. Proposed Sampling Locations for Non-Urbanized Environment

b. Urbanized Scenario

For the urbanized scenario, the land use land cover layer is used as a guide for priority setting to locate the new sampling points. From Table 14, it is observed that the Built-up area has the larger contribution when compared with the other classes in terms of count and thus proportion within the 1000 m buffer.

Table 14. LULC per stream order in meters (using the 1000m buffer)

	Stream Order					
	1st	2nd	3rd	4th	5th	6th
LULC						
Agriculture	31896	18501	4269	5579	1753	6881
Barren land	2484	1733	295	219	7	360
Built-Up	**642245**	**420271**	**156966**	**137196**	**13289**	21171
Mines and Quarries	5347	3383	1106	112	699	360
Vegetation	179466	117770	50282	30707	9785	**24565**
Water Bodies	27767	20527	10005	7467	2210	7042

Looking at the percentage contribution of each LULC type, cultivated land (agriculture) covers a small area (4.4%) which, when compared with the Built-up area, does not contribute much at 1000 m distance from the stream. Mining activities are not prominent (cover 0.4 %) in the 1000m area, but because this LULC type effect spreads easily and rapidly it is considered as one of the classes which may have a great influence on the quality of the water (Table 15).

Table 15. 1000m buffer LULC Percentage

Class Name	Count	%
Agriculture	11317	4.4
Barren	457	0.2
Build Up	166635	65.5
Mines and Quarries	1031	0.4
Vegetation	60103	23.6
Water Bodies	14998	5.9
	254541	**100**

The results from the NPPS are displayed in Table 16. These results show that the built-up area (95%) has the greatest effect on the quality of water of the Jukskei catchment compared to the other contributing classes explored in this study; and because high NPPS value means the LULC type has higher priority than a LULC type with a low NPPS, built-up should be the highly prioritized LULC type.

In this study, however, there is an exception for mines and quarries; they might cover a small area, but they may discharge effluent and are a high risk to water quality and qualify to be a high priority because of the potential outcomes of having a mining activity 1000 m from a flowing water body.

Table 16. Non-point pollution sources scores

Class Name	**Ai**	**Wi**	**NPPS**	**NPPS %**
Agriculture	4.5	0.037558	0.166984	0.50
Barren Land	0.2	0.069046	0.012396	0.04
Build Up	65.5	0.483685	31.66439	94.98
Mines and Quarries	0.4	0.284001	0.115033	0.35
Vegetation	23.6	0.036058	0.851413	2.55
Water Bodies	5.8	0.089652	0.528245	1.58
	100.0		**33.3**	**100.00**

**Ai is the % of each non-point source between candidate points*

**Wi is the potential pollution weight*

**NPPS is the non-point source pollution scores*

Proposed sampling locations

From the river mixing length method and the hierarchy calculations and non-point pollution sources, 40 new sampling locations are proposed and are distributed amongst the top three stream orders; 6th, 5th and 4th (Table 17 and Figure 18). However, this number is not the ideal because it is less than 1% of the optimal number (11 242); which is considered by several researchers including Alilou *et al.* (2018) as the best proportion to use for optimizing sampling locations.

Table 17. Proposed sampling locations per stream order

Stream Order	**Proposed Sampling Locations**
6th	13
5th	5
4th	11
3rd	9
2nd	0
1st	0

Taking a closer look at the positions of the proposed new points, it can be observed that these are more concentrated in land use with high NPPS value- the built-up area and the mines and quarries (Figure 22).

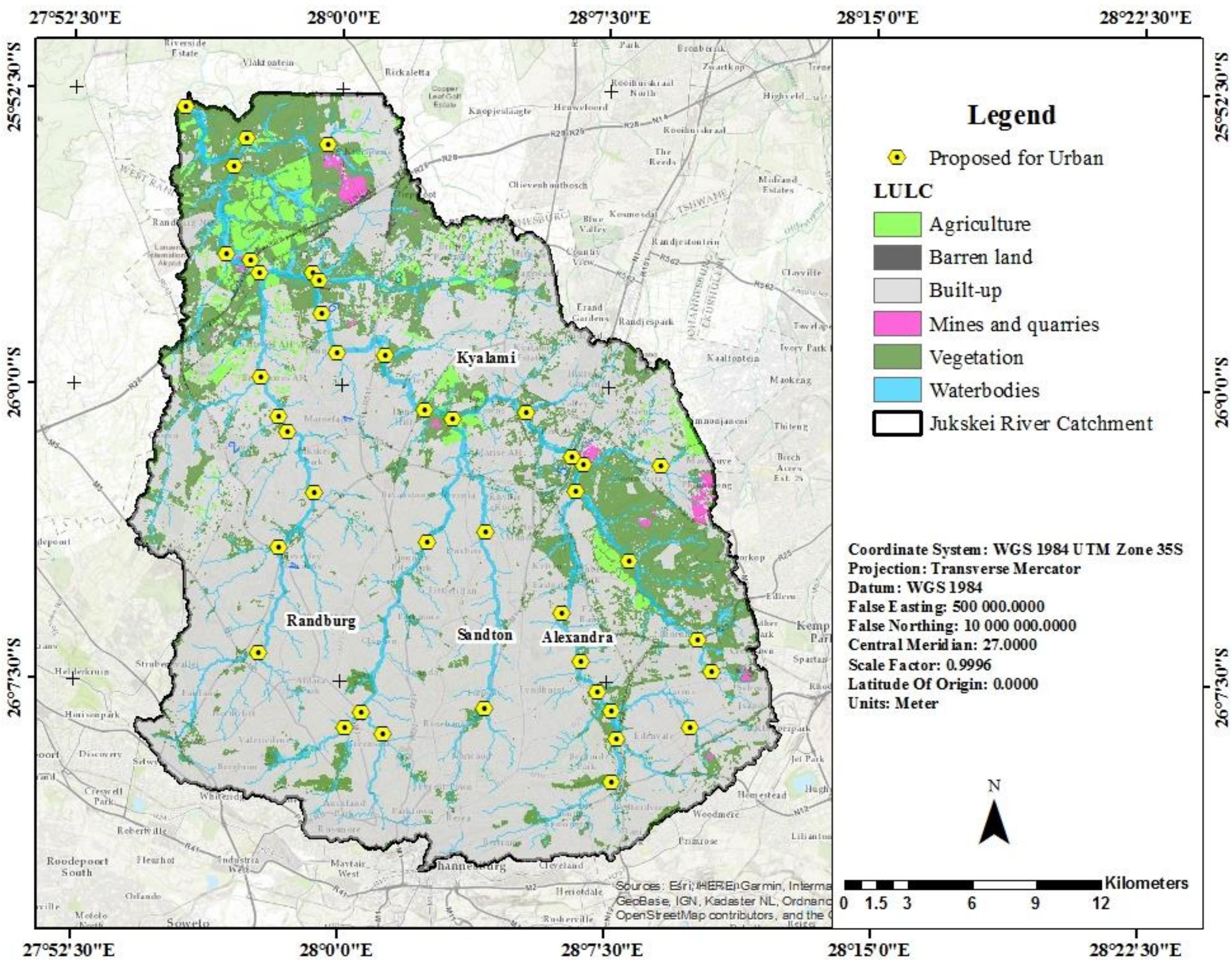


Figure 22. Proposed Sampling Locations for Urbanized Environment

Looking at the proposed number of sampling locations and the existing locations, the difference in location and number of sampling points per order is clear (Figure 23). EISD, CoJ shows 60 existing sampling locations however only 32 points are sampled monthly. Therefore, this may be one of the sources of inconsistencies of results, since recorded sampling information does not correlate. Table 18 and Figure 23 compare the sampling points of locations per order.

Table 18. Existing vs proposed sampling locations

Stream Order	Current	Non-Urbanized	Urbanized
6th	4	6	13
5th	3	3	5
4th	18	11	11
3rd	8	20	9
2nd	15	0	0
1st	12	0	0

The locations proposed for the urbanized scenario are closer to those the CoJ currently uses (within 500 m) distance and these are more correlated to the LULC and thus focus on the high pollution sources.

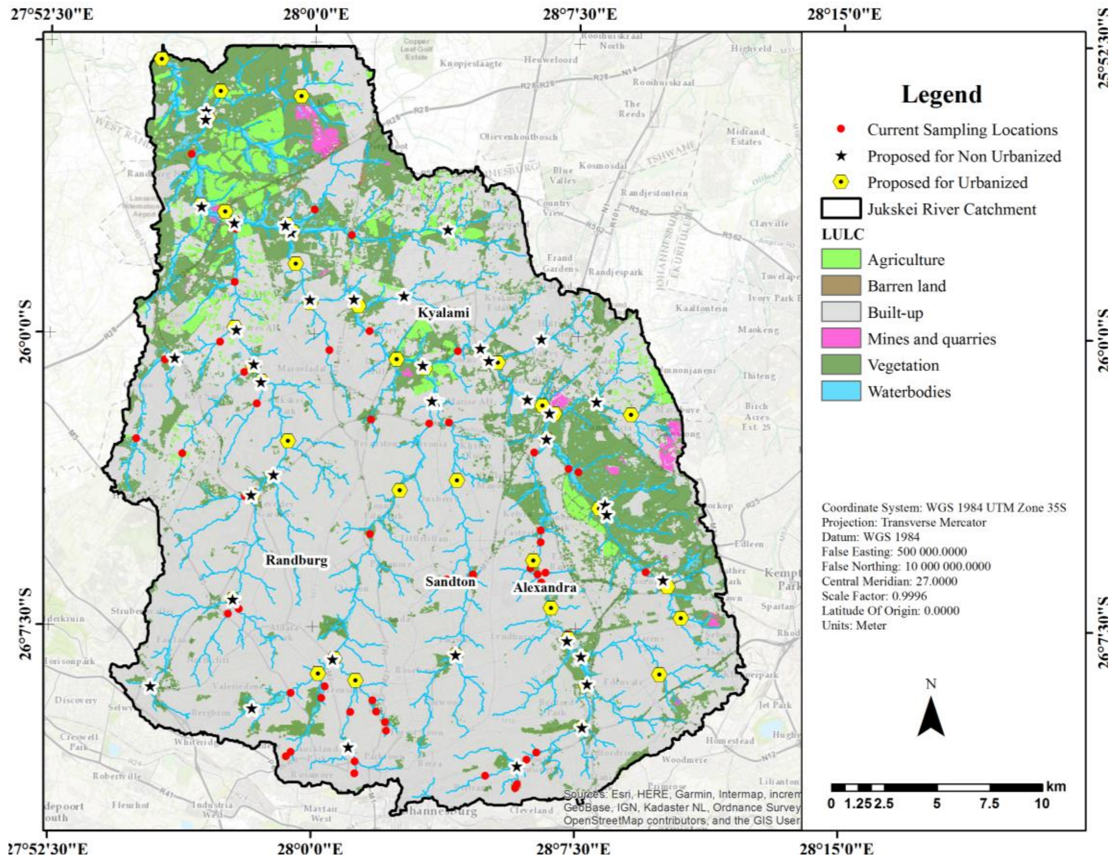


Figure 23. Comparison of existing vs proposed sampling

5. Discussion

The users' needs assessment conducted for the first objective, which was to determine the scope of a GIS platform for water quality monitoring, brought to light two main issues, the lack of iterations and revisits of the sampling locations in the water monitoring strategy. This is due to the fact that the current sampling points were placed in 2013, which means, it has been 6 years since 2013 without the re-evaluation of the used sampling locations. According to Park *et al.* (2006), this is a very long period, because rivers are not stationary but are continuously in motion and events in their surrounding change over time.

This lack of re-evaluation is due to the insufficiency of funds; which goes back to the research by Sanders *et al.* (1983), who raises the issue of water quality not being prioritized by government institutions to the level of its criticality.

Besides the lack of budget for the re-evaluation of current sampling locations; there is an enormous problem in the administration of data which is associated with the sampling locations and the lab results. Furthermore, there are also challenges in areas of data capture, data storage, data retrieving and thus data representation and simulation.

Taking into consideration existing sampling locations and budget limitations indicated in the brief by the EISD, the number of sampling locations needed for the Jukskei catchment is estimated using equation (3) introduced by Sanders (1983) as outlined in section 3.5.3.1 of the methodology. In this research, 40 sampling locations were chosen for the optimization (which is 0.36% of the optimal points). This number was chosen based on the currently used locations (32). However, this number is not ideal for optimized sampling locations since it is much less than the 10% suggested by Alilou *et al.* (2018) as the acceptable proportion to use for optimizing sampling sites.

The Jukskei River catchment as currently defined by the CoJ for sampling water quality was found to include the Crocodile and Hennops river catchments. An SRTM DEM was used to delineate the actual Jukskei River catchment boundary (Hayakawa *et al.*, 2008 and Weepener *et al.*, 2015). Topographically and hydrologically it can be suggested that the inconsistency and unexpected results noted by the CoJ water sampling is due to the inclusion of sampling points outside the catchment area.

To illustrate the impact that LULC has on the location of water quality sampling two scenarios (i) the non-urbanized and (ii) urbanized are used in this research. The locations of the existing sampling sites that are used by the CoJ EISD department were awarded as a tender to a company, which became responsible for the allocation. Furthermore the information provided on the exact procedure is not in detail as it mainly focused on the water sampling strategy overall goals and what the reasoning behind their placement was.

5.1 Non-Urbanized Scenario

To find the optimal number of sampling points for water quality of a non-urbanized Jukskei river catchment, the main river channel with different stream width was divided into segments. The stream widths were averaged and used to estimate the number of sampling points required based on their width and river channel properties. To acquire the width, Google Earth was used because of its high spatial resolution and thus accuracy (15m-1m). The greatest number of sampling points was estimated for the 3rd stream order with 20 points optimized.

For the non-urbanized scenario, the major guideline was the location of the existing set of sampling points. The percentage of the measurement (estimated potential sampling location) was used to place the optimized water quality sampling points. With the consideration of the currently used sites; the new points are placed within 1km from the current points.

5.2 Urbanized Scenario

The optimal sampling locations are widely distributed over the entire study area, with some attention to locations with high NPPS. Two factors showed the highest suitability score recorded for the LULC within the Jukskei river catchment: Built-up area and Mines and Quarries.

Do *et al. (*2012) use a similar environment (southern region of Taipei, Taiwan) to the Jukskei river catchment and their study obtained similar results to this research. Amongst the other LULC variables used, agriculture is spatially less important as it only covers 4.4% of the area and hence, does not have much effect on water quality in the 1000 m buffer used.

Mining activities are not prominent in the Jukskei river catchment as they only cover 0.4 % in the 1000 m buffer. Nonetheless, mining and quarries may release harmful effluents and due to this fact, they may have a great influence on the quality of the water (Sivertun and Prange, 2003), consequently they are considered a major contributor. The overall results obtained demonstrate a similar trend as the study conducted by Alilou *et al.* (2018) and Do *et al.* (2012), where the built-up areas show greater impact on water quality.

Although comparing two pairs of nearby LULC types seems insufficient to draw a clear statistical conclusion about the relationship of pollutant levels and suitability scores and thus priority level (Chang and Lin, 2014), such evaluation helps score priority which is very essential for the allocation of optimum locations for water quality monitoring for specific issues, for example, pollution related to mine dumps.

In this scenario, in contrast to the non-urbanised scenario, the 6th stream order has the largest number of sampling locations (13) followed by the 4th stream order. The 6th order stream having a much higher number of points can be justified by the LULC types present along the 6th order stream i.e. built-up areas and mining activities which have scored high in the PPS scores.

When comparing the proportional LULC values for the whole Jukskei catchment to the LULC of the 1 km buffer, it is noted that built-up and vegetation classes both have higher values in terms of the predominant LULC in the catchment and in the buffer. Overall the proportions are the same with the exception of water bodies and agriculture, where agriculture is more dominant than water in the entire catchment but alternates position with water in the buffer LULC (in the buffer water is more dominant than agriculture).

The distance from these optimum sites to the nearest existing station, overall varied by a few meters to a little over a kilometre distance. The main difference is the integration of the LULC layer, which adds a bit more complexity and dynamics related to the different classes of LULC and the level of impact they actually have on the river channel considering their location relative to the main channel (Jukskei river) and thus its tributaries.

6. Conclusions

This research describes a practical methodology to propose optimized sampling locations for the Jukskei river catchment, where cost and data constraints exist. Results of this research demonstrate that combining the river mixing length, LULC (considering human activities), multi-criteria evaluation and hierarchy value is the most feasible way to identify representative river sampling locations.

First, a user's needs assessment is performed to get an overview of the water quality concerns and a conceptual model is used to address observable issues to help improve the organization, planning and thus the monitoring of the water quality of the Jukskei river catchment. From here, sampling optimization is considered and recommended.

Secondly, to identify representative sampling locations the RML method is used in a series of steps. To incorporate the different levels of pollution from different LULC the PPS scoring approach is used through the ANP method; which classifies different pollution point sources into priorities. Based on the existing points of the CoJ (32) and the number of water monitoring regions that exist in the catchment, 40 points are decided upon.

Results from this research illustrated that the location of water sampling points depends on the characteristics of the river and much more significantly on the LULC and anthropogenic activities connected to the river system.

The main implication of switching to new sampling locations is budget, which takes into account the funds for the testing of the results obtained by the newly proposed sampling points. This is to see if they show any differences to the ones currently used by the department and whether they give better and thus useful readings regarding the changes in water quality over a short period of time. The general cost of travelling to and from the sampling locations is one of the main consequences of switching the location of sampling and will thus have implications on the budget restrictions and number of people responsible for taking samples.

Another often overlooked yet important factor to consider is the iterative process; since these proposed sampling locations integrate LULC the iterative process will have to take into consideration the rate of change of the LULC before establishing a period to revisit the locations of the sampling.

The overall recommendation from this research is to use the existing points (only those within the newly delineated Jukskei river catchment) as additional points to the idealized and proposed sampling point's location. This can be useful to validate the results of the proposed sampling points. If the results show a high correlation between the existing and the present, then some of the old ones can be kept. This will save the department money and resources as the establishment of new points comes with some metadata generation and validation which can sometimes be tedious.

Going forward, more research needs to be considered and addressed in water sampling and the reasons behind the development of the existing sampling point or any proposed sampling sites. As may have been noted in section 3.2, the report provided by the EISD did not provide much information on the reasons behind the location of the existing sampling locations. However, for the 40 locations proposed, a reason for their locations is explained and supported by the users' needs assessment results; this gives value to the water quality sampling locations and thus gives confidence in achieving better quality results.

Limitations

The data provided by the municipality shows inconsistencies between the number of points on the ground and the ones which are being monitored (60 points were given in the form of a shapefile but only 32 show any detail of the GPS locations with both latitude and longitude values.

Information on how the existing water quality sampling points were placed, is not well documented making it difficult to understand the thought process behind the selection of their existing points. This thus makes visual exploration and the interviews the only point of reference as to which places are regarded as priority points for the placement of sampling points in this specific river catchment.

7. References

Alilou, H., Moghaddam Nia, A., Keshtkar, H., Han, D., Bray, M., 2018. A cost-effective and efficient framework to determine water quality monitoring network locations. T*he science of The Total Environment*, 624, 283–293. https://doi.org/10.1016/j.scitotenv.2017.12.121.

Al-sabhan, W., Mulligan, M., Blackburn, G.A., 2003. A Real-time Hydrological Model for Flood Prediction Using. The GIS and the WWW. *Computers, Environment and Urban Systems* 9–32.

Alsahli, M.M. and Al-Harbi, M., 2018. Allocating optimum sites for air quality monitoring stations using GIS suitability analysis. *Urban climate*, *24*, 875-886.

Amdany, R., Chimuka, L., Cukrowska, E., Kukučka, P., Kohoutek, J., Tölgyessy, P., Vrana, B., 2014. Assessment of bioavailable fraction of POPS in surface water bodies in Johannesburg City, South Africa, using passive samplers: an initial assessment. *Environ Monit Assess* 186, 5639–5653. https://doi.org/10.1007/s10661-014-3809-3.

Ames, D.P., Rafn, E.B., Van Kirk, R., Crosby, B., 2009. Estimation of stream channel geometry in Idaho using GIS-derived watershed characteristics. *Environmental Modelling & Software*, 24 (3), 444-448.

Anon. 2017. DEM Data Users Guide Available at:
https://agdc.usgs.gov/data/usgs/geodata/dem/dugdem.pdf [Accessed 13 Aug. 2018].

Antonini, K., Langer, M., Farid, A., Walter, U., 2017. SWEET CubeSat – Water detection and water quality monitoring for the 21st century. *Acta Astronautica* 140, 10–17. https://doi.org/10.1016/j.actaastro.2017.07.046.

Assaf, H. and Saadeh, M., 2008. Assessing water quality management options in the Upper Litani Basin, Lebanon, using an integrated GIS-based decision support system. *Environmental Modelling & Software*, *23*(10-11), 1327-1337.

Band, L.E., 1986. Topographic partition of watersheds with digital elevation models. *Water resources research*, *22*(1), 15-24.

Baird, C., Jennings, M., Ockerman, D., Dybala, T., 1996. *Characterization of Nonpoint Sources and Loadings to the Corpus Christi Bay National Estuary Program Study Area.*

Beckers, C.V., Chamberlain, S.G., Grimsrud, G.P., 1972. *Quantitative methods for preliminary design of water quality surveillance systems*. Socioeconomic Environmental Studies Series Report No. EPA-R5-72-001, US EPA, Washington, DC.

Bresler, E., and R. E. Green, Soil parameters and sampling scheme for characterizing soil hydraulic properties of a watershed, Tech. Rep. 148, 42, *Water Resour*. Res. Cent., Univ. of Hawaii, Honolulu, 1982.

Burgess, T. M., R. Webster, and A. B. McBratney, Optimal interpolation and isarithmic mapping of soil properties, 6, Sampling strategy, *J. Soil Sci.*, 31, 643-659, 1981.

Campbell, L.A., 1996. A study of stormwater runoff from Alexandra township in the Jukskei River. Masters Thesis. University of the Witwatersrand, Johannesburg, South Africa.

Chang, C.L. and Lin, Y.T., 2014. A water quality monitoring network design using fuzzy theory and multiple criteria analysis. *Environmental monitoring and assessment*, *186*(10), 6459-6469.

Chen, Y., Han, D., 2018. Water quality monitoring in the smart city: A pilot project. *Automation in Construction* 89, 307–316. https://doi.org/10.1016/j.autcon.2018.02.008.

Chilundo, M., Kelderman, P., 2008. Design of a water quality monitoring network for the Limpopo River Basin in Mozambique. *Phys. Chem. Earth*, Parts A/B/C 33 (8), 655–665.

Cobban, D. A. .1994. Assessment of water quality through a menu-driven user interface on a geographic information system platform. In: *Computer Graphics '*94 (Proc. 14th Ann. South African Computer Aided Design Symp.), session 1, 1-23.

Cobban, D. A. & Silberbauer, M. J. 1993. Water quality decision-making facilitated through the development of an interface between a geographic information system and a water quality database. In: Hydrology in Developing Regions... the Road Ahead (Proc. Sixth South African *National Hydrological Symp. ed.* by S. A. Lorenz, S. W. Kienzle & M. C. Dent), vol. II, 523-530.

Craigie, D. (2018). The *GIS and integrated water resource management*. Available at: http://www.ee.co.za/wp-content/uploads/legacy/GIS%20and%20integrated.pdf [Accessed 15 Aug. 2018].

Davis, B., and L. Borgman, Some exact sampling distributions for variogram estimates, *J. Int. Assoc. Math. Geol*., 11, 643-645, 1978.

Day, T.J., 1977. Observed mixing lengths in mountain streams. *Journal of Hydrology*, 35 (1–2), 125–136.

Do, H.T., Lo, S.L., Chiueh, P.T. and Thi, L.A.P., 2012. Design of sampling locations for mountainous river monitoring. *Environmental Modelling & Software*, 27, 62-70.

Gobakis, K., Mavrigiannaki, A., Kalaitzakis, K., Kolokotsa, D.-D., 2017. Design and development of a Web-based GIS platform for zero energy settlements monitoring. Energy Procedia, Sustainability in Energy and Buildings 2017: Proceedings of the Ninth KES International Conference, Chania, Greece, 5-7 July 2017 134, 48–60. https://doi.org/10.1016/j.egypro.2017.09.598.

DWAF.1996. *Water Law Principles - Discussion Document.* Department of Water Affairs and Forestry, Pretoria.

DWS. 2017. *Review, Evaluation and Optimization of The South African Water Resources Monitoring Network,* Pretoria.

ESRI. 1991. A*RC/INFO Data Model, Concepts & Key Terms*. Environmental Systems Research Institute, Inc., Redlands, California, USA.

Harris, J., Van Veelen, M. and Gilfillan, T.C., 1992. *Conceptual design report for a National River Water quality assessment programme.* Water Research Commission.

Harmancioglu, N.B. Fistikoglu, O., Ozkul, S.D., Singh, V.P., Alpaslan, M.N., 1999. *Water Quality Monitoring Network Design.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hayakawa, Y.S., Oguchi, T. and Lin, Z., 2008. Comparison of new and existing global digital elevation models: ASTER G-DEM and SRTM-3. *Geophysical Research Letters*, 35(17).

Hoover, J.H., Sutton, P.C., Anderson, S.J. and Keller, A.C., 2014. Designing and evaluating a groundwater quality Internet GIS. *Applied Geography*, *53*, 55-65.

Horton, R.E., 1945. Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin*, *56*(3), 275-370.

Huggins, L. F. and Burney J. R. 1982. 'Surface runoff, storage, and routing, Chapter 5'. In Haan, C.T. (Ed.). Hydrologic Modeling O/ Small Watershed, ASA E Monograph No. 5 *Agricultural Nonpoint Source Pollution Model*, American Society of Agricultural Engineering, St. Joseph, Michigan, 160 - 170.

Icaga, Y., 2005. Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) river basin. *Environmental Monitoring and Assessmen*t 108 (1), 261-277.

Ierodiaconou, D., Laurenson, L., Leblanc, M., Stagnitti, F., Duff, G., Salzman, S., Versace, V., 2005. The consequences of land use change on nutrient exports: a regional scale assessment in south-west Victoria, Australia. *Journal of Environmental Management* 74 (4), 305e316.

Jalbert, K., Kinchy, A.J., 2016. Sense and Influence: Environmental Monitoring Tools and the Power of Citizen Science. Journal of Environmental Policy & Planning 18, 379–397. https://doi.org/10.1080/1523908X.2015.1100985.

Jenson, S.K. and Dominique, J.O., 1988. Extracting topographic structure from digital elevation data for geographical information system analysis. Photogrammetric Eng. *Remote Sensing*, 54:1593-1600.

Jollymore, A., Haines, M.J., Satterfield, T., Johnson, M.S., 2017. Citizen science for water quality monitoring: Data implications of citizen perspectives. Journal of Environmental Management 200, 456–467. https://doi.org/10.1016/j.jenvman.2017.05.083.

Jolma, A., Altartouri, A., Ferencik, I., 2010. GIS, environmental modeling and engineering. *Environmental Modelling & Software* 25 (12), 1915-1916.

Karamouz, M., Kerachian, R., Akhbari, M. and Hafez, B., 2009. Design of river water quality monitoring networks: a case study. *Environ. Model. Assess*. 14 (6), 705.

Kühn, A. L., Kempster, P. L. and Everett, A. 1996. Drought Relief Program: *An Assessment of the Water Quality of Boreholes for Potable Use, January 1994 to October 1995.* IWQS report N/0000/00/GPQ1895, Department of Water Affairs and Forestry, Pretoria.

Kulkarni, A.T., Mohanty, J., Eldho, T.I., Rao, E.P., Mohan, B.K., 2014. A web GIS-based integrated flood assessment modeling tool for coastal urban watersheds. Computers & Geosciences 64, 7–14. https://doi.org/10.1016/j.cageo.2013.11.002.

Kucukaltan, B., Irani, Z. and Aktas, E., 2016. A decision support model for identification and prioritization of key performance indicators in the logistics industry. *Computers in Human Behavior*, *65*, 346-358.

Lin, Y.-H., Tsai, K.-M., Shiang, W.-J., Kuo, T.-C., Tsai, C.-H., 2009. Research on using ANP to establish a performance assessment model for business intelligence systems. *Expert Syst. Appl.* 36, 4135–4146.

Lin, B., Xu, Z., Wang, J., Lu, M., 2017. A low-cost water quality monitoring prototype device with embedded chromogenic reagent capsules and dynamic colorimetric detection. Sensors and Actuators B: Chemical 252, 24–29. https://doi.org/10.1016/j.snb.2017.05.056.

Lo, S.L., Kuo, J.T., Wang, S.M., 1996. Water quality monitoring network design of Keelung River, Northern Taiwan. *Water Science and Technology* 34 (12), 49-57.

McBratney, A. B., and R. Webster, How many observations are needed for regional estimations of soil properties?, *Soil Sci.*, 135, 177-183, 1983.

Mishra, S.S. and Nagarajan, R., 2010. Morphometric analysis and prioritization of sub-watersheds using GIS and remote sensing techniques: a case study of Odisha, India. *International Journal of Geomatics and Geosciences*, *1*(3), p.501.

Moore, C A, 1990, *Classification system and indices for reporting and determining the effectof management on water quality in South African water bodies,* (In preparation),DWA, Pretoria.

Myers, D.N., Ludtke, A.S., 2017. Chapter Two - Progress and Lessons Learned from Water-Quality Monitoring Networks, in Ahuja, S. (Ed.), Chemistry and Water. Elsevier, 23–120. https://doi.org/10.1016/B978-0-12-809330-6.00002-7.

Nielsen, J. (1997). Usability testing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed.) (pp. 1543e1568). New York: J. Wiley

Ning, S.K., Chang, N.B., 2004. Optimal expansion of water quality monitoring network by Fuzzy optimization approach. *Environmental Monitoring and Assessment* 91 (1), 145-170.

Nivala, A.-M., Brewster, S., & Sarjakoski, L. T. 2008. Usability evaluation of web mapping sites. The Cartographic Journal, 45(2), 129-138.

Ouyang, Y., 2005. Evaluation of river water quality monitoring stations by principal component analysis. *Water Research* 39, 2621–2635. https://doi.org/10.1016/j.watres.2005.04.024.

Park, S.Y., Choi, J.H., Wang, S. and Park, S.S., 2006. Design of a water quality monitoring network in a large river system using the genetic algorithm. *Ecological modeling*, 199(3), 289-297.

Pérez, C.J., Vega-Rodríguez, M.A., Reder, K., Flörke, M., 2017. A Multi-Objective Artificial Bee Colony-based optimization approach to design water quality monitoring networks in river basins. *Journal of Cleaner Production* 166, 579–589. https://doi.org/10.1016/j.jclepro.2017.08.060.

Pro.arcgis. 2017. *How Stream Order works—Help | ArcGIS Desktop.* Available at: http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/how-stream-order-works.htm [Accessed 11 Oct. 2018].

Qin, Y., Alam, A.U., Pan, S., Howlader, M.M.R., Ghosh, R., Hu, N.-X., Jin, H., Dong, S., Chen, C.-H., Deen, M.J., 2018. Integrated water quality monitoring system with pH, free chlorine, and temperature sensors. *Sensors and Actuators B: Chemical* 255, 781–790. https://doi.org/10.1016/j.snb.2017.07.188.

Rekolainen, S., Kämäri, J., Hiltunen, M. and Saloranta, T.M., 2003. A conceptual framework for identifying the need and role of models in the implementation of the Water Framework Directive. *International Journal of River Basin Management*, *1*(4), 347-352.

Repetto, M.P., Burlando, M., Solari, G., De Gaetano, P., Pizzo, M., Tizzi, M., 2018. A web-based GIS platform for the safe management and risk assessment of complex structural and infrastructural systems exposed to wind. *Advances in Engineering Software, Special Section* - CIVIL-COMP 2017 117, 29–45. https://doi.org/10.1016/j.advengsoft.2017.03.002.

Russo, D., 1984. Design of an optimal sampling network for estimating the variogram, *Soil Sci. Soc. Am. $.*, 48, 708-716.

Saaty, T.L., Vargas, L.G., 2006. *Decision Making with the Analytic Network Process.* Springer

Sanders, T.G., Ward, R.C., Loftis, J.C ., Steele, T.D., Adrian, D.D., Yevjevich, V. 1983. *Design of Networks for Monitoring Water Quality*. Water Resources Publications LLC, Highlands Ranch, CO.

Sharp, W.E., 1971. A topologically optimum water-sampling plan for rivers and streams. *Water Resources Research*, *7*(6), 1641-1646.

Shreve, R.L. 1996. Statistical Law of Stream Numbers. *The Journal of Geology*, 74(1), 19-37

Shreve, R.L., 1967. Infinite topologically random channel networks. *The Journal of Geology*, *75*(2), 178-186.

Sivertun, Å, Prange, L., 2003. Non-point source critical area analysis in the Gisselö watershed using GIS. *Environmental Modelling & Software 18 (*10), 887-898.

Steele, T.D., 1987. Water quality monitoring strategies. *Hydrological sciences journal*, *32*(2), 207-213.

Strobl, R., Robillard, P., Shannon, R., Day, R., McDonnell, A., 2006. A water quality monitoring network design methodology for the selection of critical sampling points: Part I. *Environmental Monitoring and Assessment* 112 (1), 137-158.

Strobl, R.O. and Robillard, P.D., 2008. Network design for water quality monitoring of surface freshwaters: A review. *Journal of environmental management*, *87*(4), 639-648.

Telci, I.T., Nam, K., Guan, J. and Aral, M.M., 2009. Optimal water quality monitoring network design for river systems. *Journal of environmental management*, *90*(10), 2987-2998.

Templin, W.E., 1984. *Ground-water-quality monitoring network design for the San Joaquin Valley ground-water basin, California.* USGS Water-Resources Investigations Report 83-4080, USGS, Denver, CO.

ThoughtCo. (2017). *How Are the World's Waterways Classified?* Available at: https://www.thoughtco.com/what-is-stream-order-1435354 [Accessed 11 Oct. 2018].

Tong, S.T. and Chen, W., 2002. Modeling the relationship between land use and surface water quality. *Journal of environmental management*, 66(4), 377-393.

Vieux, B.E., 1991. Geographic information systems and non-point source water quality and quantity modeling. *Hydrological Processes*, *5*(1), 101-113.

Wang, X. and Yin, Z.Y., 1997. Using GIS to assess the relationship between land use and water quality at a watershed level. *Environment International*, *23*(1), 103-114.

Ward, R.C., 1973. *Data acquisition systems in water quality management*. Socioeconomic Environmental Series Studies Report No. EPA-R5-73-014, US EPA, Washington, DC.

Weepener, H.L., le Roux, J.J., van den Berg, E.C., Tswai, D.R., and Nell, J.P., 2015. Developing methods for converting digitised rivers into a hydrological drainage network. Pretoria, South Africa: *Water Research Commission*, No. 2164/1/15

Woolfenden, L.R., 1984. *Geohydrology and ground-water quality monitoring network for the Lower Mojave River Valley, California*. USGS Water-Resources Investigations Report 83-4148, USGS, Denver, CO.

World Health Organization, 2004. *Guidelines for Drinking-water Quality: Recommendations*. World Health Organization.

Yi, Z., 2018. Chapter Three - Techniques for Identifying Users' Needs and Wants, in Marketing Services and Resources in Information Organizations, Chandos Information Professional Series. Chandos Publishing, pp. 29–37. https://doi.org/10.1016/B978-0-08-100798-3.00003-9

Zwane, M, 2018. Personal Communication, 25 June 2018.

Appendix A

Questionnaire

Jukskei Water Quality Data Capturing Questionnaire

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

A Masters Project on creating a GIS platform for monitoring Water Quality with the benefits of Data Management, Data Accessibility, Data Visualization, Data Distribution and Data Simulations for broader evaluation and interpretations.

Note: Personal information on interviewee is required and will be used in the reporting of the research

What is the sampling Strategy?

____

____

When were the sampling points established? ☐

How were the sampling points established?

____

____

Has there been a change in sampling points since established? ☐

If yes (mention change): ____________________

Main challenges with Capturing Data (X)

- Writing (Scribbling) ☐
- Geo referencing ☐
- Missing Data ☐
- Misplaced sheets ☐

Backup Database (X)

- No ☐
- Yes
  - Online ☐
  - External Drive ☐
  - Other (Specify) __________

Ethical Clearance: GAES 5-2018-0013

Jukskei Water Quality Data Capturing Questionnaire

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

What kind of platform would you prefer for the GIS platform (X)

- Public ☐
- Non-public ☐
- Web based ☐
- Non- Web based ☐

What main function will the GIS platform perform? (X)

- Record Data ☐
- Share Information ☐
- Statistical Analysis ☐
- Visualize and Model Data ☐

Ethical Clearance: GAES 5-2018-0013