ASSESSMENT OF GROUNDWATER POTENTIAL IN FRACTURED HARD ROCKS AROUND VRYBURG, NORTH WEST PROVINCE, SOUTH AFRICA

Report to the
Water Research Commission

by

Abera Tessema, Ugo Nzotta and Emmanuel Chirenje
Council for Geoscience, Pretoria

WRC Report No. 2055/1/13
ISBN 978-1-4312-0510-3

February 2014

Obtainable from

Water Research Commission
Private Bag X03
GEZINA, 0031

orders@wrc.org.za or download from www.wrc.org.za

DISCLAIMER

This report has been reviewed by the Water Research Commission (WRC) and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Executive Summary

The assessment of groundwater quality and quantity is a major step towards ensuring the sustainable use and management of one of the most basic needs of human beings. Understanding the amount of groundwater resource and its quality assists in creating awareness amongst decision makers to use, manage and to protect groundwater without adversely affecting its future demand.

In this study, investigation of groundwater potential in crystalline basement rocks of the North West Province was carried out. The area of study is located in the Naledi Local Municipality situated in the central part of the North West Province. It covers an area of ~7260 km^2 and consists of 8 Quaternary catchments. Hydrogeologically, a large part of the area falls within the Lower Vaal catchment. The average annual precipitation in the area is ~350 mm and temperature varies from very cold (below freezing point) to 35ºC during the warm seasons. Groundwater recharge in the area is low (<10 mm) and largely depends on temperature and the seasonality and intensity of rainfall. Potential evaporation rate ranges from 1960 mm to 2100 mm per annum, exceeding annual rainfall. It is typically a semi-arid to arid region, and groundwater is the main source of water supply for domestic and agricultural use.

Several groundwater studies have been carried out previously by private and government organizations aimed at development of rural and municipal water supply. Most of these studies were mainly focused around Vryburg and Stella, Dithakwaneng, Huhudi and Lima that are located within the Municipality. In the previous studies, aquifer test, recharge estimation, geophysical and geological logging were carried out at number of sites.

In general, groundwater potential depends on many hydrogeological factors including surface and bedrock lithology, structures, slope steepness and morphology, stream density, climate, soil and land use. Although the previous studies contributed significantly to the background information about the groundwater potential of the area, however, the influence of different hydrogeological features was not well addressed in light of understanding the groundwater potential of the area. The main objectives of this study were therefore:

- To delineate groundwater potential zones using multivariate statistical modelling approaches.
- To assess groundwater controlling factors.
- To investigate the groundwater quality using hydrochemical data and environmental isotopes.

Understanding the groundwater potential of the area is useful to regulate and optimize water use without adversely impacting water resource for future use. Below, a summary of the findings of the present work is presented.

Aquifer tests have been carried out at three boreholes to determine sustainable yield and aquifer parameters. Based on analysis of the data, borehole management and abstraction recommendation were provided for each of the borehole. The analysis of hydrochemical data show the presence of fresh water throughout the study area with the exception of some boreholes located around Stella and north of Vryburg. The exceptions are due to elevated concentrations of chloride, nitrate, fluoride and total hardness that exceeded the standard concentrations as prescribed by SANS241.

The water types in the study area are largely influenced by rock-water interaction and anthropogenic sources. The influence of high evaporation in the northern part of the area also partly contributed to chloride-dominated water-type. The enrichment in the stable isotopic composition ($\delta^{2}H$ and $\delta^{18}O$) of groundwater indicates the effect of evaporation which is consistent with high chloride content as shown using analysis of hydrochemical data. However, the limited isotope data used in the present study makes it difficult to understand the interaction between surface water and groundwater.

Two multivariate statistical modelling were carried out to map the most prospective areas for groundwater occurrence. These include empirical (data-driven) and conceptual (knowledge-based) approaches. The data-driven approach includes artificial neural networks (ANNs), Weights of Evidence (WofE), Logistic Regression (LG) and Principal Component Analysis (PCA). The Knowledge-based modelling involves fuzzy logic (FL) and fuzzy clustering (FC). Both statistical modelling approaches were applied to five evidential themes that include; geology, lineaments, geomorphology/slope, land use and soil texture.

For the data-driven modelling approach, 46 boreholes with yield > 10 l/s and 43 sites with no indication of groundwater occurrence (barren sites) were used as training points. The 5 evidential themes and the training data were used as an input to *feed-forward radial basis function* neural network training algorithm. Various statistical analysis were carried out that include; analysis of model sensitivity, training, validation and testing. These allowed to generate groundwater potential map of the area.

The resulting map shows a number of groundwater potential zones varying from "*very good*" to "*very poor*". The zone shown as '*very good'* and '*good'* groundwater potential covers ~17% and ~22% of the study area, respectively. The superimposed borehole yield also confirms the results derived from multivariate statistical modelling approaches, whereby high borehole yields (> 15 l/s) fall within carbonate rocks consisting of dolomite and limestone located in the southern part of the area. In addition, follow-up geophysical surveys carried out at selected sites confirmed the presence of conductive layers varying in depth from 20 m to 35 m. The high conductivity possibly indicates the presence of water-bearing formation, in particular dolomite located around Vryburg and highly fractured granite just south of Stella. The high yielding wellfields can be attributed to dissolution of carbonate rocks by water that percolates through pre-existing fractures leading to enlarged fracture apertures, and consequently resulting in the development of large cavities that can store and supply significant amount of water (e.g. Armoedsvlakte and Swartfontein wellfields located north of Vryburg).

The WofE, PCA, RBFLN and LG provided nearly similar results showing lithologic units and fracture connectivity and concentration are the main controlling factors of groundwater occurrence. The influence of lithology is significant in the southern part of the area, while fracture connectivity played important role in controlling groundwater occurrence within the hard rocks terrane consisting of the Ventersdorp volcanic rocks, Archaean granite gneisses and the greenstone belt.

The multivariate statistical approaches used for characterization of groundwater potential in crystalline basement and carbonate rocks of the study area is very effective in delineating potential areas. Differences in fracture set populations, coupled with contrast in primary permeability of various rock types significantly contributed to variability in groundwater potential throughout the study area. Although high groundwater potential zones are primarily associated with carbonate rocks, the results derived from calculation of statistical correlation between lineaments density and borehole yield suggests that high yielding wellfields are largely controlled by fracture concentration within crystalline basement rocks. Furthermore, the significant contrast in groundwater potential between the southern and northern parts of the area can be attributed to the presence of two end-members controlling the development of groundwater, i.e. dissolution channels in the south and fracture connectivity and density in the central and northern parts of the area.

Acknowledgements

The authors sincerely wish to express their gratitude to the Water Research Commission for the financial grant without which this study would not have been completed. The Council for Geoscience and the Department of Water Affairs are greatly acknowledged for allowing access to various datasets used in this study. We thank Ms Motebe and Mr Monokofala for their cooperation in allowing access to several borehole data and unpublished reports. We would like to express our gratitude to Mr Leon Pretorius Manager Strategic Planning, Support and Operations of the Naledi Local Municipality in the North West Province for allowing access to many boreholes in the study area. Many local people around the Naledi Local Municipality assisted in locating the geographical positions of the boreholes. Matome Sekiba and Melvin Sethobya are greatly acknowledged for assisting the fieldwork.

We express our gratitude to the steering committee below for providing stimulating comments during the reference group meetings.

Dr Shafick Adams	WRC (Chairman)
Mr Ernest Bertram	Department of Water Affairs
Dr Julian Conrad	GEOSS – Geohydrological and Spatial Solutions International
Dr Martin Holland	Delta-H
Ms Nancy Motebe	Department of Water Affairs
Mr Collen Monokofala	Golder Associates Africa (Pty) Ltd.
Dr M Lupanka	

Two Honours and one MSc students from the University of Venda assisted the fieldwork. We thank Daniel Sebaki for assisting access to various spatial datasets used in this project. The authors would like to thank Dr. Abraham Thomas for comments and suggestions.

“*Only one who devotes himself to a cause with his whole strength and soul can be a true master. For this reason mastery demands all of a person*”.
Albert Einstein

Table of Contents

List of Figures

List of Tables

Acronyms and Abbreviations

ANNs	Artificial neural networks
bgl	Below ground level
BP	Back propagation
C	Contrast between W^+ and W^-
CAPP	Cumulative area posteriori probability
CDT	Constant discharge test
CI ratio	Conditional Independence ratio
CIT	Conditional independence test
CP	Conditional probability
DEM	Digital elevation model
DWAF	Department of Water Affairs and Forestry
FC	Fuzzy clustering
FL	Fuzzy logic
FNN	Fuzzy neural network
GMWL	Global meteoric water line
LC	Land cover
LG	Logistic Regression
LU	Land use
MLP	Multi-layer perception
NN	Neural network
PC	Principal component
PCA	Principal component analysis
PLMWL	Pretoria local meteoric water line
PLP	Perception learning rule
PNN	Probabilistic neural network
PRC	Prediction rate curve
RBFLN	Radial basis function link neural network
SANS241	South African National Standard for drinking water quality guideline
SAR	Sodium adsorption ratio
SDT	Step drawdown test
SE	Supporting Evidence
SRC	Success rate curve
SSE	Sum of square of errors

Std. Cont	Studentized contrast. It is the contrast C divided by its standard deviation
TGA	Transvaal Griqualand Axis
TP	Training points
UCU	Unit Condition Unit
WHO	World Health Organization
WofE	Weights of Evidence
σ	Standard deviation
W^+	Weight due to the presence of high borehole yield
W^-	Weight due to the absence of high borehole yield
σw^+	Standard deviation of W^+
σw^-	Standard deviation of W^-

Glossary

***Confidence*:-** The contrast divided by the standard deviation (a student T test) for a given evidential theme. It provides a useful measure of significance of the contrast due to uncertainties arising from the weighs and areas of possible missing data.

***Contrast C*:-** The difference between $\mathbf{W^+}$ and $\mathbf{W^-}$. In binary theme the contrast **C** is an overall measure of the spatial association of an evidential theme with the training points. Contrast is the range in weighs and indicates the significance of the theme in predicting the location of training points. Large *contrast* indicates strong association between the evidential theme and the training points.

***Conditional Independence*:-** The weights of evidence model assumes that evidential themes are independent, not correlated at the location of the training points. If this assumption is true, and if the prior probability is defined as indicated below, the summation area multiplied by the predicted posteriori probability over the study area should be equal to the number of training points.

***Evidential Theme*:-** A spatial feature/theme used as evidence for the occurrence of the raining points. Theme attribute may be categorical (e.g. geological map) or ordered (e.g. proximity to fault).

***Fitting rate curve*:-** Success rate curve.

***Hard-limit-transfer function*:-** The function limits the output of the neuron to either 0 if the net input argument n is less than 0, or 1 if n is greater than or equal to 0. This function is used to create neurons that make classification decisions.

***Logit*:-**The logit function is the inverse of the sigmoidal “logistic” function used in statistics. When the function’s parameter represents a probability, the logit function returns the log-odds.

***Posteriori Probability*:-** The probability that a unit cell contains a training point, given states of information from the evidential theme. This measurement changes from location to location depending on the values of the evidence, being larger than the prior probability where the sum of the weights is positive, and the converse.

***Prediction Rate Curve*:-** A graph of cumulative testing sites versus cumulative area, from high to low posteriori probability. It is used in model validation. It quantifies how well a groundwater predictive map *delineates* the testing points.

Prior probability:- Predicted probability from the training points only, aka training points density (TP/Area). It is aka probability before considering the evidence. The probability that a unit cell contains before considering a spatial evidence.

Studentized Contrast:- A measure of the significance of the contrast.

***Success Rate Curve*:-** A plot of cumulative training sites versus cumulative area from high to low posteriori probability. It quantifies the goodness of fit between predictive map and the training points. Used in model validation.

***Training point*:-** The known locations of which is being predicted. In this study, the training points are known boreholes with yield >10 l/s.

Unit Cell:- A small unit of area for counting, and for defining the probability of occurrence of a point object. Each training point is assumed to occupy a unit cell. All area measures are transformed to counts of unit cell. The unit cell should be small, normally as small as or smaller than the minimum spatial resolution of the evidential theme. Weights values are relatively insensitive to unit cell size if unit cell is small.

1. Introduction

Basement aquifers are of particular importance in tropical and sub-tropical regions both because of their widespread extent and because there is no readily available alternative source of surface water, particularly for rural populations in Africa (e.g. Farquharson and Bullock, 1992 *in* Wright, 1992). However, the development of crystalline bedrock aquifers is highly complicated, and groundwater occurrence is spatially variable (Chilton and Foster, 1995; Mabee, 1999; Banks and Robinson, 2002). Designing an exploration approach for identification of groundwater potential zones in basement terrane is complex. One of the biggest challenges is to understand the controls on groundwater occurrence (Holland and Witthüser, 2011). According to Titus et al. (2009), tectonic and geomorphic events, together with climatic conditions resulted in the development of weathered and fractured aquifer systems (e.g. within the Namaqualand basement terrain), where structures and thick weathering horizons were described by these authors as important controlling factors for the development of hard rock aquifers. In such type of terrain, mapping fracture density and analysis of their relationship to borehole yield may provide some clue about groundwater potential zones. In addition, the use of geophysical techniques may highlight the thickness of the overlying weathered regolith. However, field verification is important to ascertain the relative importance of bedrock fractures and weathering horizons in controlling groundwater potential zones. Also, as discussed by several authors (e.g. Mabee et al., 2002; Latman and Parzek, 1964; Mabee et al., 1994), the success of lineament/structure analysis is usually measured by the magnitude of the yield produced from wells in close proximity to the lineament, i.e. wells *on* or *near* lineaments often exhibit high yields. It is therefore important to use as much data as possible combined with field studies to improve the success rate in identifying productive water wells. As described by Holland and Witthüser (2011), interrelated factors such as topography (slope), lithology, structures and surface water amongst others play important roles in controlling the occurrence of groundwater.

Fractured basement aquifers are common throughout South Africa and the North West Province in particular. Finding a sustainable source of water supply is one of a challenge in many parts of South Africa due to the arid-to-semi arid nature of the climate and the complexity of basement aquifers. These two factors have largely contributed to scarcity of water in many rural areas. In almost all parts of the North West Province, water is supplied from boreholes that are located along major streams and carbonate rocks consisting of limestone and dolomite. Characterization of existing water wells in relation to lithology, structures and other factors that control groundwater occurrence is therefore an important step towards rural community development.

1.1 Aims and Objectives

The main aim of this study is to investigate the groundwater potential of the Naledi Local Municipality which is located in the North West Province using hydrogeological, satellite imagery and geophysical data combined with field studies.

The specific objectives of the study are:

- To delineate groundwater potential zones using multivariate statistical approaches
- To assess groundwater controlling factors
- To investigate the groundwater quality and chemical characteristics based on analysis of hydrochemical data

The knowledge of the occurrence, replenishment and recovery of potable groundwater have special significance in this region.

1.2 Description of the study area

The study area, Naledi Local Municipality, is situated in the central part of the North West Province with its centre at Vryburg town (Figure 1). It covers an area of ~7260 km^2 and is subdivided into 8 Quaternary river catchments (C32A, C32B, C32C, C32D, C31E, C31F, D41B and D41C). Hydrogeologically, large part of the study area fall within the Lower Vaal catchment and a smaller portion falls within the Crocodile catchment. The upper Molopo catchment forms part of the Vryburg-Mafikeng hydrogeological series and the present study area falls within this regional-scale hydrogeological map. It is an arid-to-semi-arid region and the livestock and people greatly depend on groundwater for domestic use. The average annual precipitation in the area is ~350 mm with low and high annual precipitation of 180 mm and 788 mm, respectively. The temperature in the area is variable, ranging from very cold (below freezing point) to 35ºC during the warm season. The area is characterized by low recharge potential <10 mm as described by Vegter (1995). In general, the amount of recharge in this area is largely influenced by increased temperature and the seasonality and intensity of rainfall. Potential evaporation rate ranges from 1960 mm to 2100 mm per annum, thus exceeding annual rainfall. It is characterized by flat topographic relief with isolated small hills made up of resistant rocks such as banded ironstone and quartzite of Gold Ridge Formation. Elevation varies from 1100 m in the south to 1400 m a.m.s.l. in the northeastern part of the area. The landscape is generally flat and increases in altitude in the eastern part of the study area. Surface drainage is sparse due to strong seasonal characteristics of rainfall as well as high evapotranspiration losses. Most of the streams drain southward and northward forming surface-water-divide in the northern part of the study area (Figure 1).

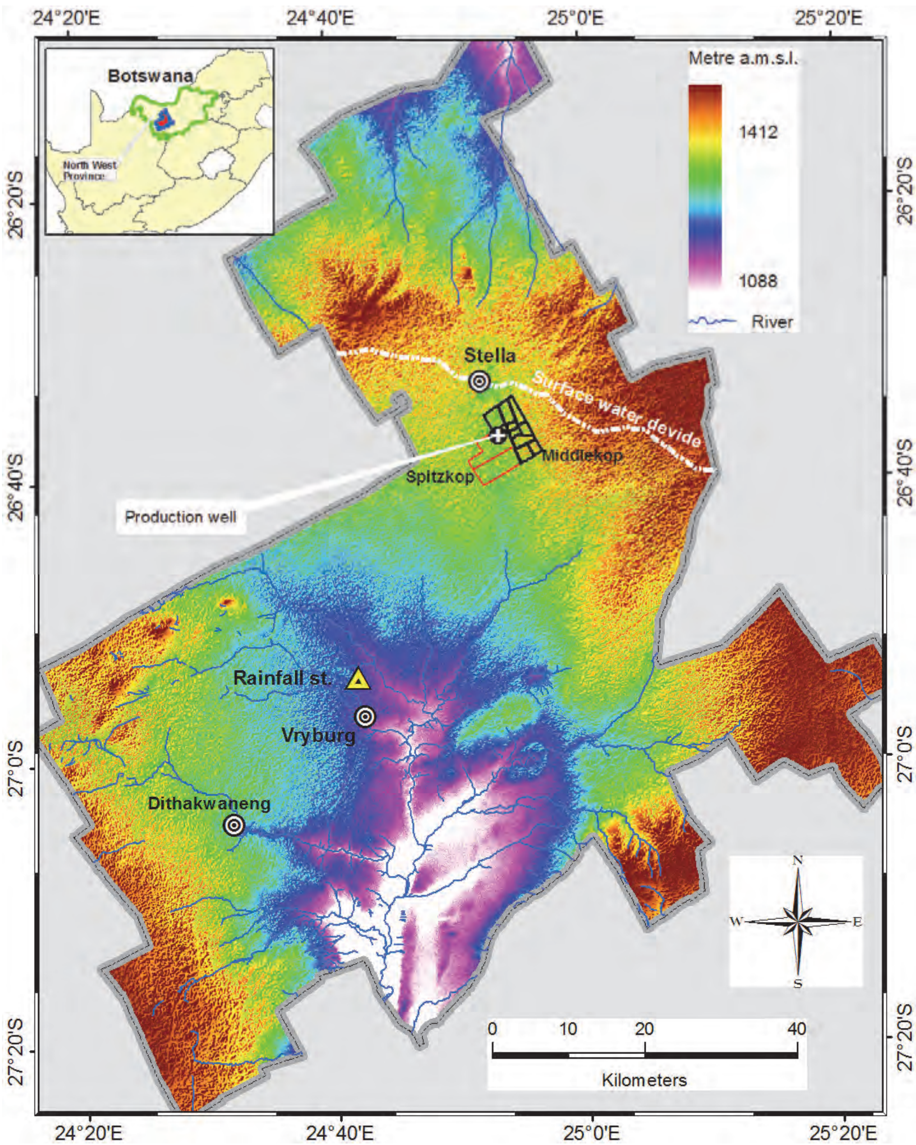


Figure 1: Location map of the study area superimposed on topographic relief highlighting hilly and flat areas.

The main urban centres and villages in the Naledi Local Municipality include Vryburg, Stella, Huhudi, Dithakwaneng, Devondale, Kameel and Broedersput. The total population is around 70,000 including urban and rural settlements. Stella is like a typical rural village with approximately 1500 residents with no large-scale industries. The residents depend on over-exploited wellfields located in the nearby aquifers which are barely able to supply the demand of water for the town.

The main land-use includes settlement (urban and rural), agricultural and some minor mining activities. Livestock farming dominates most of the land use in the study area. As there is no precise figure for the total number of cattle on the communal land within the study area, the water consumption for cattle cannot be quantified. However, based on the present field observation the supply of water for livestock and domestic use entirely depend on groundwater.

1.3 Surface Water

Useable surface water is very limited throughout the study area with the exception of some water that accumulates and form pans after heavy precipitation events. Small pans are scattered in the central and northern parts of the study area and are visible from satellite images and digital elevation (DEM) data. Seasonal rivers can be observed in the northern and southern parts of the area. In the northern part, there is a surface water divide between C32 and D41 Quaternary catchments, which is shown by a curvilinear transect passing through Stella (Figure 1). The lack of surface water distribution appears to be due to high evaporation rate that varies from 220 mm to 440 mm per month in summer and 145 mm to 70 mm per month in winter (South African Bureau of Weather Service, 2011). In general, an increase in evaporation rates largely contributed to the lack of surface water in the region, and as a result of this the only source of water supply for domestic and agricultural use is groundwater.

1.4 Previous study

Several groundwater studies have been carried out previously by private and government organizations aimed at development of rural water supply within District Municipality. Van der Westhuizen and Hodgson (1981) extensively explored different aquifer potential to the north and northwest of Vryburg. In their report, the two authors recommended detailed exploitation of the Vryburg Quartzite intersected at a depth ranging from 100 m to 150 m, on the basis of its aquifer parameters as well as groundwater quality. Nel and van Wyk (1999) also conducted aquifer testing at Spitzkop farm located south of Stella town (Figure 1). They used a production well and four monitoring boreholes that are surrounding the production well. The pumping was carried out for 72 hours at constant discharge rate of 11.6 l/s. The interpretation of the aquifer test indicates that the transmissivity of fractured porosity is 2000 m^2/d while the matrix is 3 m^2/d. The storativity ranges from 0.02 to 0.002 corresponding to granite gneiss of the Kraaipan greenstone belt. Nel and van Wyk (1999) also used water level, precipitation and aquifer test and the resulting integration of these data suggests that groundwater recharge is lower than abstraction south of Stella. They recommended groundwater management and protection in order to ensure sustainable use.

Geophysical logging using neutron, gamma ray and resistivity logging were carried out at Middlekop farm located south of Stella by Nel (2001). In addition, borehole camera logging was also conducted in some of the wells in the Middlekop wellfield to characterize the permeability of the formations and aquifer thickness. The results show that the thickness of the Middlekop granitic-gneiss aquifer varies from 40 m to 45 m with a borehole yield of 25 l/s. According to Nel's (2001) result, the fractures that form the secondary porosity occur at a depth ranging from 20 m to 45 m.

Previous studies (e.g. Nel, 2001) indicated that high blow yields around Stella are associated with dense fracture networks within the Middlekop granitic rock. In addition, the Gold Ridge Formation is an important area of focus due to cross-cutting structures allowing groundwater storage. According to this interpretation, lineament length and density are important parameters for optimization of borehole yield within the hard rock terrane of the northern part of the study area.

Aquifer test, borehole inventory and recharge estimation have been carried out by Motebe (2004) and Monokofala (2011) around Vryburg town, Dithakwaneng and Lima villages (Figure 4). Motebe (2004) used the chloride mass balance technique for recharge estimation. The results show that mean groundwater recharge around Vryburg is 8 mm/a during storm events. Constant and step-drawdown tests were carried out at 3 boreholes in

Dithakwaneng village and 2 boreholes in Lima village. In addition, similar tests were carried out at 6 boreholes located within Armoedsvlakte farm. The step-drawdown tests were conducted using 4 to 5 steps each step lasting for 60 minutes prior to conducting constant discharge tests. The constant discharge tests were analyzed using FC method, and the results show sustainable yield vary from 2 l/s to 3.5 l/s at Armoedsvlakte wellfield, 0.6 l/s to 6 l/s at Lima village and 1.5 to 3 l/s at Dithakwaneng village. According to her results, the borehole which was pump-tested for 72 hrs gave 353 m^2/day early time transmissivity which was reduced to 173 m^2/day at late time. In general, the results of the previous pumping tests around Dithakwaneng village suggest that the boreholes are poor-to-moderate in performance due to lack of fracture connectivity resulting in no flow boundary.

Monokofala (2011) carried out constant discharge tests at several wellfields that include Biesjiesvlakte, Armoedsvlakte, Vegter, Swartfontein and Vryburg. The results show, sustainable yield varying from 0.3 l/s to 12.5 l/s. The least sustainable yield corresponds to a borehole in Vryburg wellfield, while the maximum corresponds to a borehole located at Armoedsvlakte farm.

Terrain conductivity mapping, magnetic and gravity surveys were carried out by Motebe (2004) to identify drilling sites. The location of the geophysical survey lines is shown on Figure 4. Six sites were recommended for drilling based on combined interpretation of EM, gravity and magnetic data. Drilling was carried out at the six sites, and five became successful in striking groundwater at depth varying from 17 m to 79 m.

The Department of Water Affairs carried out geohydrological studies around Vryburg town to identify productive wells that supply water for domestic and industrial use (van Dyk, 1993). The results show that boreholes that were pump tested around Vryburg town exhibit poor or slow recovery rate which may be attributed to an increase in distance between the drilled boreholes and fracture concentration.

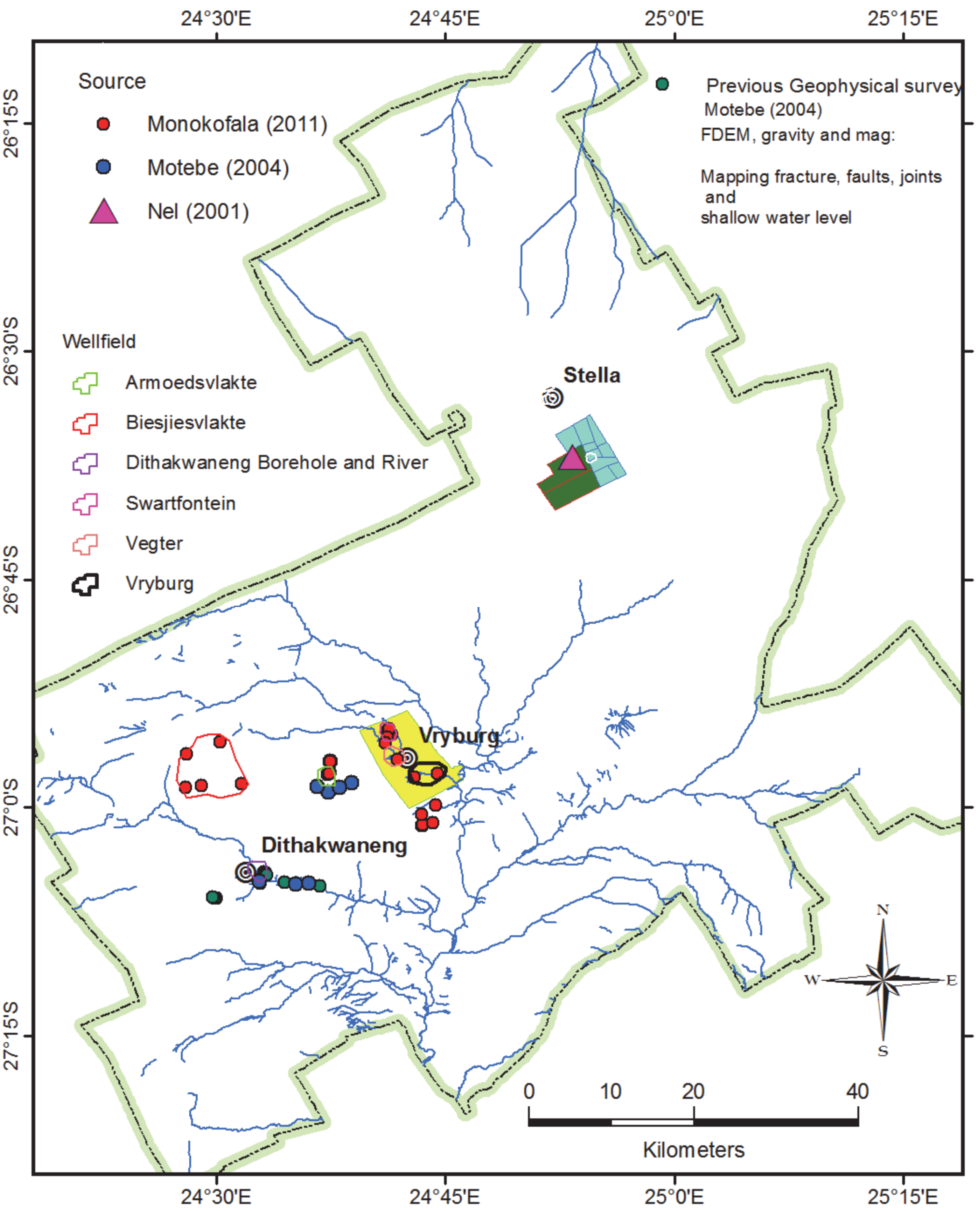


Figure 2: Location map of previous pump tested boreholes and geophysical studies.

2. Geological and Hydrogeological Setting

2.1 Geological Formations

The study area consists of 4 major lithological domains, *viz.* from the oldest to the youngest: 1) Archaean Kraaipan granite-greenstone terrane consisting of metavolcanic and metasedimentary rocks and the associated intrusive granite-gneiss 2) the Neo-Archaean Ventersdorp Supergroup comprised of mafic and felsic lavas as well as sedimentary rocks, 3) the Transvaal Supergroup consisting of a succession of dolomite, limestone, sandstone, shale, banded iron formation, mudstone and minor conglomerates, 4) the youngest cover sequence comprises of the Karoo Supergroup and the Tertiary-to-Quaternary Kalahari Group.

The outcrops of the oldest rocks are limited due to the recent cover sequence. However, the full extent of the Kraaipan greenstone belt in the study area can be traced from the aeromagnetic data showing a strong magnetic signature due to the presence of banded iron formation. It comprises of a linear belt of Archaean metasedimentary and metavolcanic rocks which are separated by granitoid units. Fractured banded ironstones with interbedded schist are also observed within the north-south striking greenstone belt due to their strong magnetic contrast with respect to the host rocks. According to Anhaeuser and Walraven (1999), the Kraaipan greenstone belt was subjected to lateral compressional forces during episodic granitoid emplacement in the western Kaapvaal craton resulting in localized shearing and faulting which produced cross-cutting structures that develop along N-S and NE-SW strike.

The Ventersdorp Supergroup is preserved over an area of approximately 200,000 sq. km, extending from the Northern Cape to the North West Province including the present study area (van der Westhuizen and de Bruiyn, 2000). It consists of alternating units of volcanic and 60% sedimentary rocks that are dominated by alluvial-fan and lacustrine deposits. There are 3 Groups that form the Ventersdorp Supergroup, *viz*. Klipriviersberg Group at the base, followed by the Platberg Group and the Pniel Group on top (van der Westhuizen, 1991). The Klipriviersberg Group exclusively consists of mafic volcanic rocks and has been subdivided into 6 formations (Winter, 1976). The Platberg Group comprises of various sedimentary rocks and mafic-to-intermediate lava flows and is divided into 3 formations. The Pniel Group consists of the sedimentary Bothaville Formation and the volcanic Allanridge Formation that constitutes the youngest of the Ventersdorp Supergroup.

According to Burke et al. (1985), the Ventersdorp Supergroup was variably deformed during or after the collision of the Kaapvaal and Zimbabwe Cratons. Post collision events were followed by deposition of sedimentary rocks along fault bounded grabens. The presence of intra-basinal fault-bounded coarse sedimentary deposits together with faults and folded structures within the Ventersdorp Supergroup suggest possible development of fractured basement and inter-granular aquifers that are favourable for groundwater occurrence.

Hydrogeologically, the study area can be subdivided into 3 potential aquifer types, *viz*. basement aquifers consisting of the Kraaipan, Ventersdorp volcanics and Archaean intrusive rocks. Faulted and weathered volcanic rocks also fall under this category of aquifer types. The inter-granular aquifers correspond to the Platberg and Kalahari Group sediments consisting of sandstone and alluvial gravel. The Malmani Subgroup of the Transvaal basin forms a karst aquifer located in the southern part and the northern tip of the study area (Figure 2). It consists of limestone, dolomite and calcareous sedimentary rocks that largely cover south of Vryburg. The Tertiary-to-Quaternary Kalahari Group predominantly consists of sandstone, calcareous-grit and conglomerates. They form inter-granular aquifers which are located in the northern and northeastern parts of the study area (Figure 2).

The borehole yield corresponding to the carbonate rocks (limestone and dolomite) varies from 2 to 25 l/s with the maximum yield occurring south of Vryburg (Figure 2). Because of the high productivity of the wells (yield > 5 l/s), the number of drilled boreholes is high within the carbonate rocks compared to elsewhere in the study area. The basement rocks extensively cover the central parts of the study area. The high borehole yield within these rocks appears to be controlled by the availability of cross-cutting structures and the thickness of weathered horizons. However, in this area the influence of cross-cutting structures on borehole yield is not clear. Mapping the lateral continuity of structures within the crystalline basement rocks is very important in order to understand their significance to borehole productivity.

.

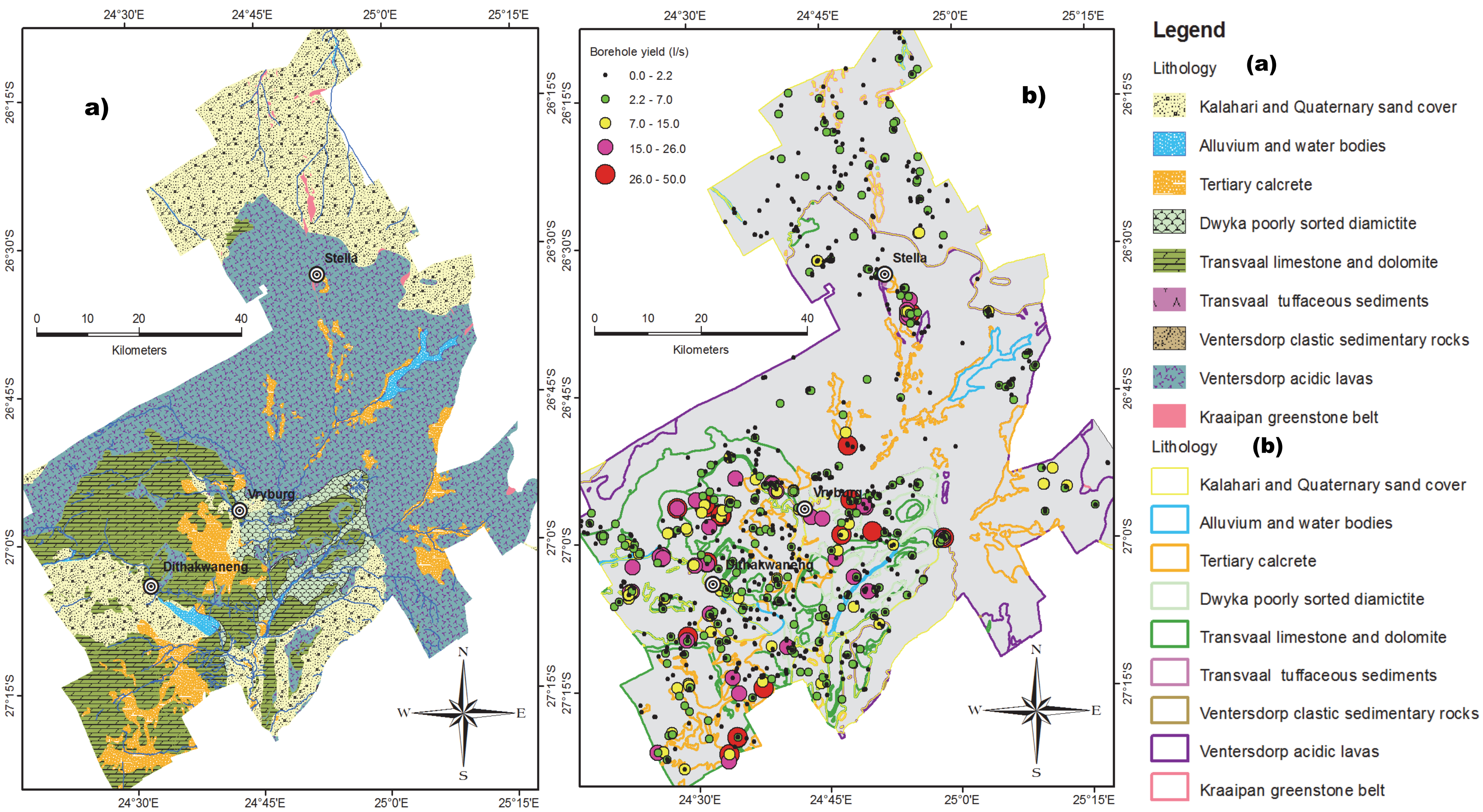


Figure 3: **a)** Simplified geological map of the study area (1:250000 CGS Map). **b)** Borehole yield superimposed on lithological boundaries.

2.2 Structures

Fractures are one of the important linear features controlling groundwater occurrence. They can be subdivided into joints, fissures and faults and are formed by brittle fracturing of rocks. Folds are produced by ductile deformation, and the extent of this deformation reflects the magnitude of the features formed, i.e. synclines and anticlines. The identification of these effects of deformation is useful for selection of targets for detailed groundwater exploitation.

The area of study forms part of the western edge of the Kaapvaal craton where structures and lithological boundaries are poorly exposed. Previous studies (e.g. Burke et al., 1985), noted that extensional faulting accompanied by extrusion of large volumes of Ventersdorp lavas covered parts of the North West and Northern Cape Provinces. According to these authors, the thick sedimentary deposits and large volume of lava flows suggest faulting was accompanied by crustal extension that took place at ~2643±80 Ma. The maximum thickness of the sediments deposited is confined by structural grabens that are bounded by N-S striking series of fault systems. Evidence for faulting can also be supported by aeromagnetic and satellite images that show several NE-SW and N-S striking lineaments (Figures 3a, 3b and 3c). The NE-SW striking lineaments are the most prominent though NNW-SSE and N-S are also observed in the northwestern and southern parts of the area (Figure 3d).

According to Roering et al. (1990), the western part of the Kaapvaal craton including the present study area experienced compressional and extensional tectonic activities which resulted in faulting and folding during the emplacement of Ventersdorp lavas. Winter (1976) also suggested that block faulting was active during post magmatic events. The enormous outpouring of lava flows together with thick sedimentary deposits may provide supporting evidence for the area being subjected to block faulting accompanied by basin development. These processes were continued during the emplacement of the Bushveld Complex and Pilanesberg alkaline extrusion, which produced a series of normal and strikes-slip fault systems with NE-SW and N-S strike. In order to understand the influence of these structures on the development of fractured basement aquifers, it is important to map lineaments based on geological, geophysical and remote sensing data.

Several lineaments with two dominant strikes (ENE-WSW and WNW-ESE) were identified based on combined interpretation of aeromagnetic map and satellite images shown in Figures 3b and 3c. The lineament map of the area (Figure 3d) was constructed based on on-screen digitizing of persistent linear features of regional extent. In the central part of the study area just north of Vryburg, the lineaments are less than 1 km wide and relatively closely spaced than elsewhere in the study area. This suggests that areas that sustained brittle deformation possibly exhibited by high lineament density are important area of focus for detailed groundwater exploration.

A rose diagram of all mapped lineaments was plotted which shows three discrete trends (ENE-WSW, N-S and WNW-ESE) (bottom-left on Figure 3d). The ENE-WSW lineaments cross-cut large terrane boundaries, which appear to be part of the westward continuation of the late Archaean Thabazimbi-Murchison-Lineament (TML). According to Good and De Wit (1997), these lineaments exhibit all three senses of faulting, viz. normal, strike-slip and reverse faulting. They are widespread north of Vryburg and Stella towns extending over a distance of >100 km (Figure 3d).

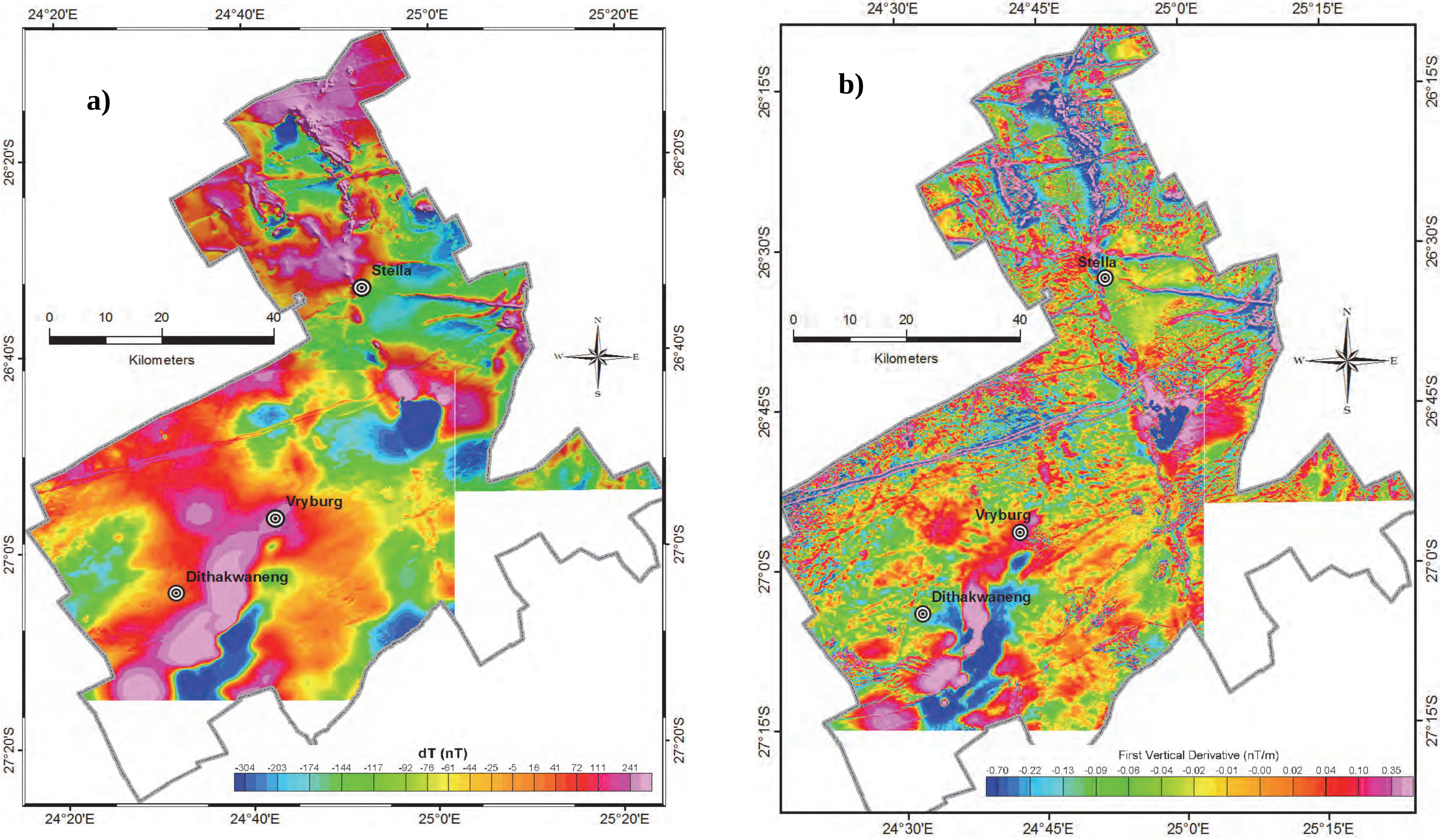


Figure 4: **a)** Total magnetic intensity and **b)** Edge-enhanced magnetic map of the study area.

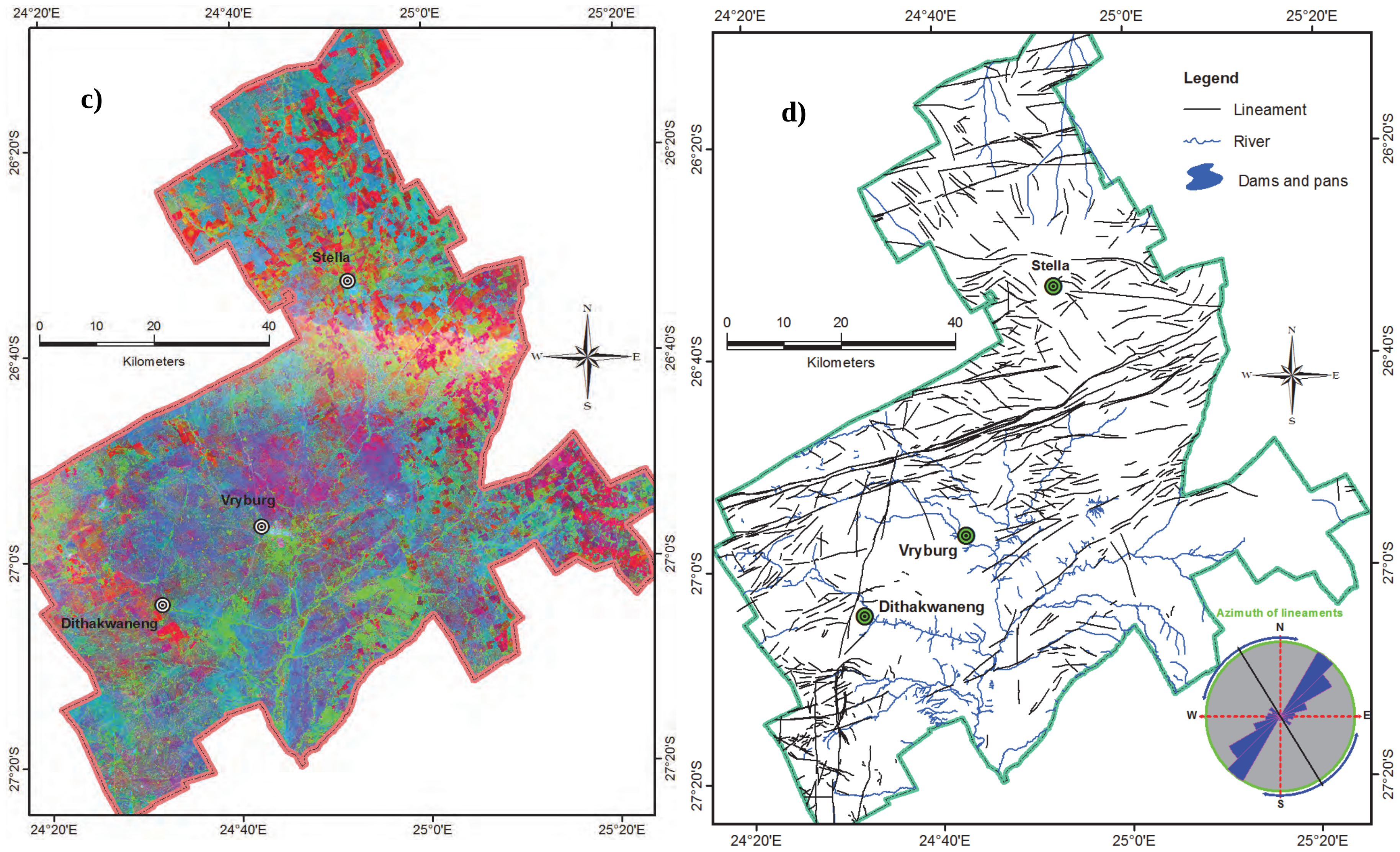


Figure 4: **c)** Landsat 7 ETM composite image derived from principal components 1, 2 and **d)** Lineaments derived from magnetic and satellite images.

2.3 Response of Water level to Rainfall

Monitoring of groundwater levels is an important indicator of groundwater potential, because a change in water level is directly related to a change in the volume of water available in an aquifer. The temporal change in groundwater level also indicates the response of the aquifer to replenishment, abstraction, climate change and water use amongst other factors.

The water level data for the present study area was obtained from the National Groundwater Archive, though many boreholes were not continuously monitored. The data was converted from depth to height above datum by subtracting the water level below surface from the topographic height at each borehole. The water level ranges from 1 m to 140 m depth below ground level. The contour map of the water level provides regional-scale patterns of the water table in the area (Figure 5).

The map shows areas of groundwater recharge, discharge and direction of flow. The water level generally mimics topographic relief (Figures 1 and 5), with groundwater flowing from higher recharge areas to lower discharge areas though this is not consistent with what we observe in the southern part of the study area. Figure 5 shows that water level decreases southward suggesting groundwater flows from north to south similar to the surface water that forms a curvilinear boundary passing through Stella town in the north (Figure 5).

The water level varies widely across the study area reflecting the influence of lithology, topography, land use and land cover as well as water use. In the southern part of the area, water level is largely influenced by topography and lithology. The low water level shown in areas of settlement (e.g. Vryburg and Dithakwaneng) may be attributed to over abstraction although this is not evident around Stella because of the influence of the topographic high.

The slope of the potentiometric surface (imaginary water level) is gentle in the central part of the study area, and then becomes steeper as we move southward showing a recharge zone in the north and a discharge sector in the south. Vryburg town and Dithakwaneng village are located on a steep potentiometric surface relative to Stella town indicating variability in groundwater distribution controlled by both morphology and geology.

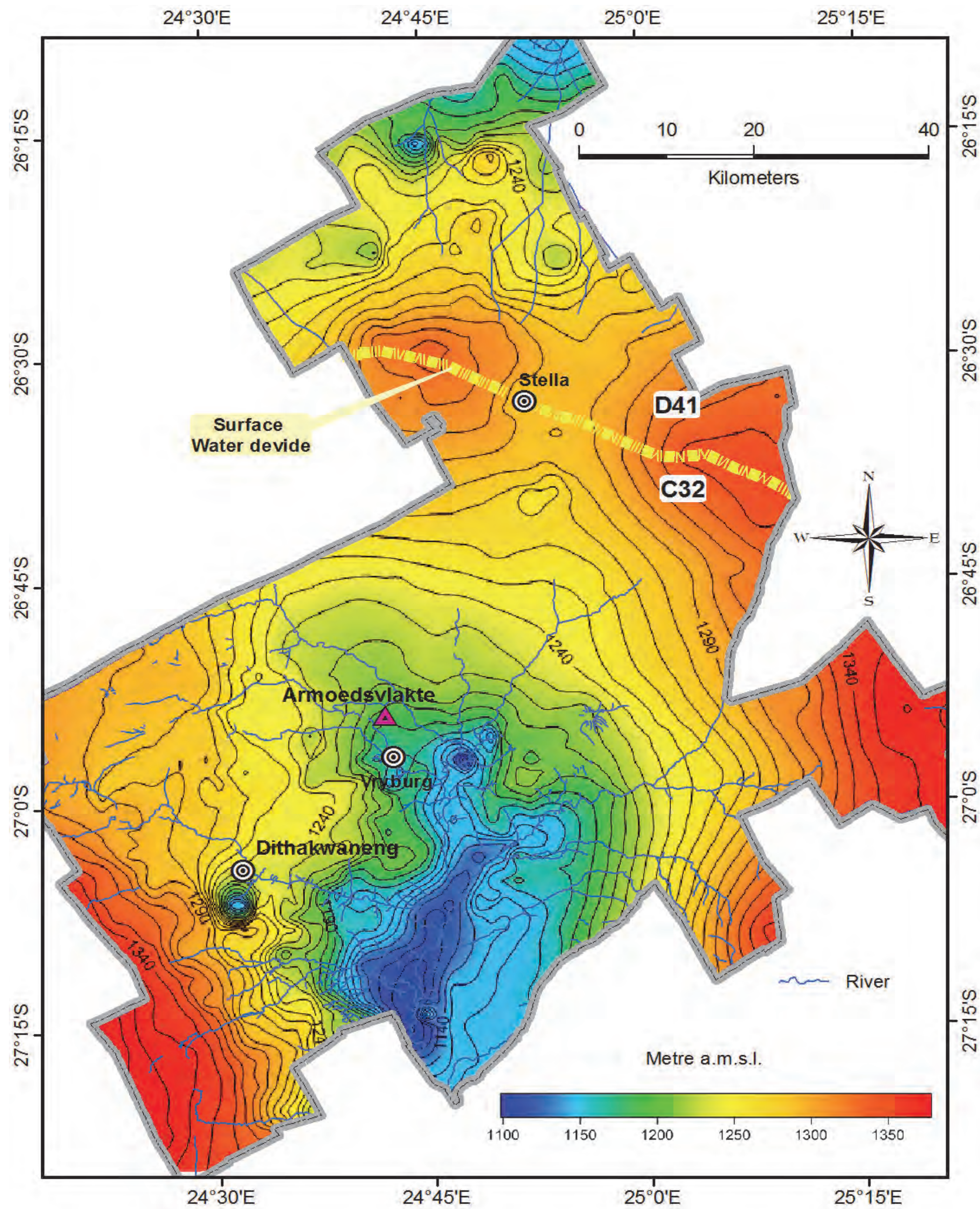


Figure 5: Potentiometric surface map constructed from interpolation of water level recorded at several boreholes. The contour lines are in metre a.m.s.l. The yellow thick line demarcates the boundary between Quaternary river systems.

There are few boreholes where continuous monitoring of water level was recorded over long period. One of the few boreholes is located at Armoedsvlakte meteorological station (Figure 5). At this station, rainfall and water level were continuously monitored from 1985 to 2004. The monthly mean rainfall for each year was plotted to characterize its long-term impact on water level (Figure 6). The graph shows that from 1985 to 1988, the water level rapidly increased from 1215 m a.m.s.l. to 1235 m, which a 20 m rise in water level over 3 years. From July 1998 to January 2000, the water level declined steadily from 1240 m to

1205 m, i.e. a decline of 35 m. The fluctuations of water level during the two periods can be described as short-term which may be attributed to seasonal climate conditions such as rainfall (flood events) and fluctuation of temperature that decreases or increases evaporation and evapotranspiration over a localized region (e.g. Armoedsvlakte wellfield). Similar fluctuation of water level resulting from seasonal changes in climatic conditions and land use and land cover has been suggested by Gong et al. (2012). According to these authors, an increase in water level as high as 11 m can occur after rain events if the top soil is silt loam and covered by trees with less dense leafs. However, if the loam soil is covered by grass, the water level increases only by 6 m, suggesting that the type of soil as well as land use and land cover play an important role in fluctuation of water level.

Previous studies (e.g., Motebe, 2004) also noted the short-term and long-term response of water level to fluctuation of rainfall intensity. According to this author, at Armoedsvlakte wellfield, periods of above average rainfall are normally succeeded by attainments of the highest level of water table. The maximum response of water level commonly lags the period of the maximum rainfall event by approximately two to three months. The length of lag depends on the hydraulic properties of the overlying material and it may vary depending on site characteristics such as land cover, geology, soil and surface temperature.

The monitoring of water level at wellfields Biesjiesvlakte, Swartfontein and Stella show slow recovery during the rainy season. In general, 90% of the boreholes in the study area show incomplete recovery in water level owing to imbalance in pumping, influence of evapotranspiration and recharge. The variability in water level may also be attributed to lithology, i.e. the recharge of groundwater occurs where the rocks are sufficiently permeable to allow water to move downward through the unsaturated layer until it reaches the water table (Gong et al., 2012). For example, the southern part of the study area is dominated by dolomite and limestone where seepage of surface water can take place and recharge groundwater. This process possibly contributed to shallow water level compared to the central and northern part of the study area where rocks are characterized by poor permeability resulting in slow recovery of water level after rain events.

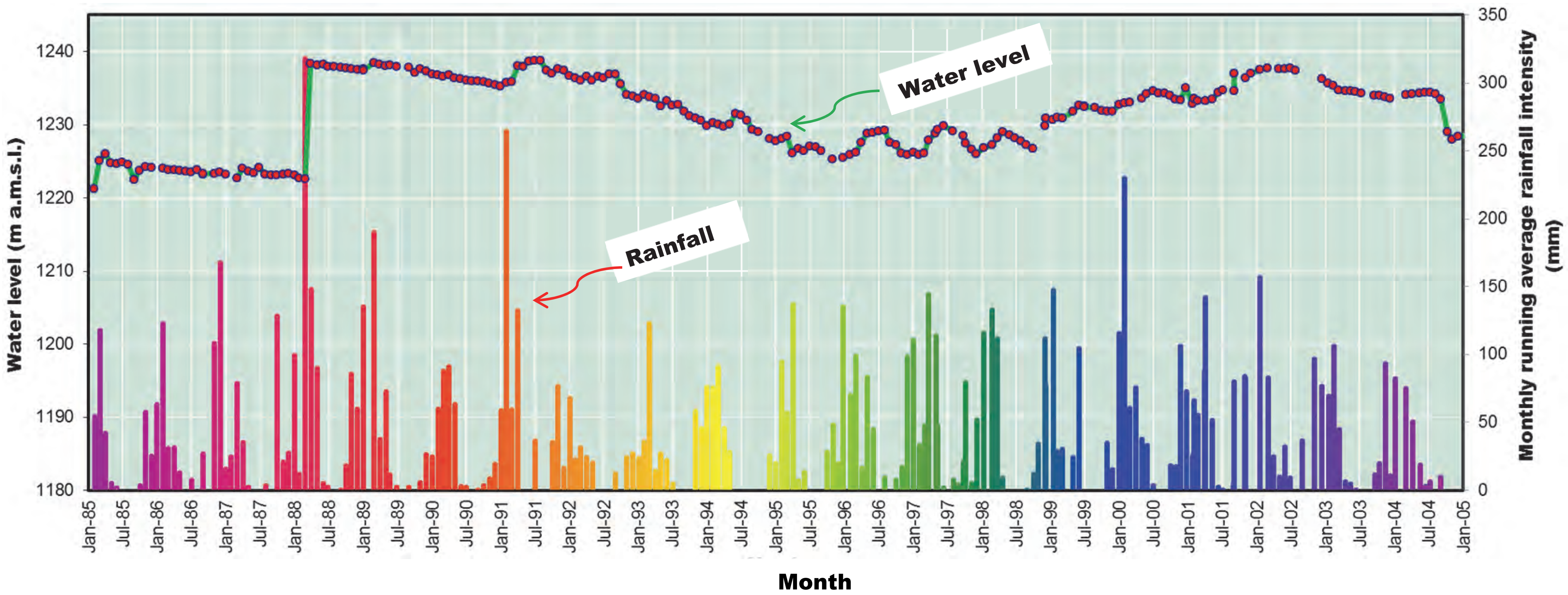


Figure 6: Water level response to monthly running average rainfall intensity at Armoedsvlakte wellfield.

2.4 Depth of Water Strike

The water strike data for the present study area was obtained from the Department of Water Affairs' National Groundwater Archive. At each water strike encountered during drilling, the depth and yield were recorded. Figure 7a shows the depth of water strike in the study area interpolated from the existing boreholes at a grid spacing of 200 m. The depth varies from 4.5 m to 160 m and the statistical distribution shows skewed patterns towards shallow depth (Figure 7b). The number of boreholes drilled within the carbonate rocks is greater than elsewhere in the study area. In addition, the depth of water strike within the carbonate rocks is relatively shallow and this accounts for the skewed statistical distribution of depth of water strike. The maximum depth corresponds to the Ventersdorp volcanic and sedimentary rocks (Figure 7).

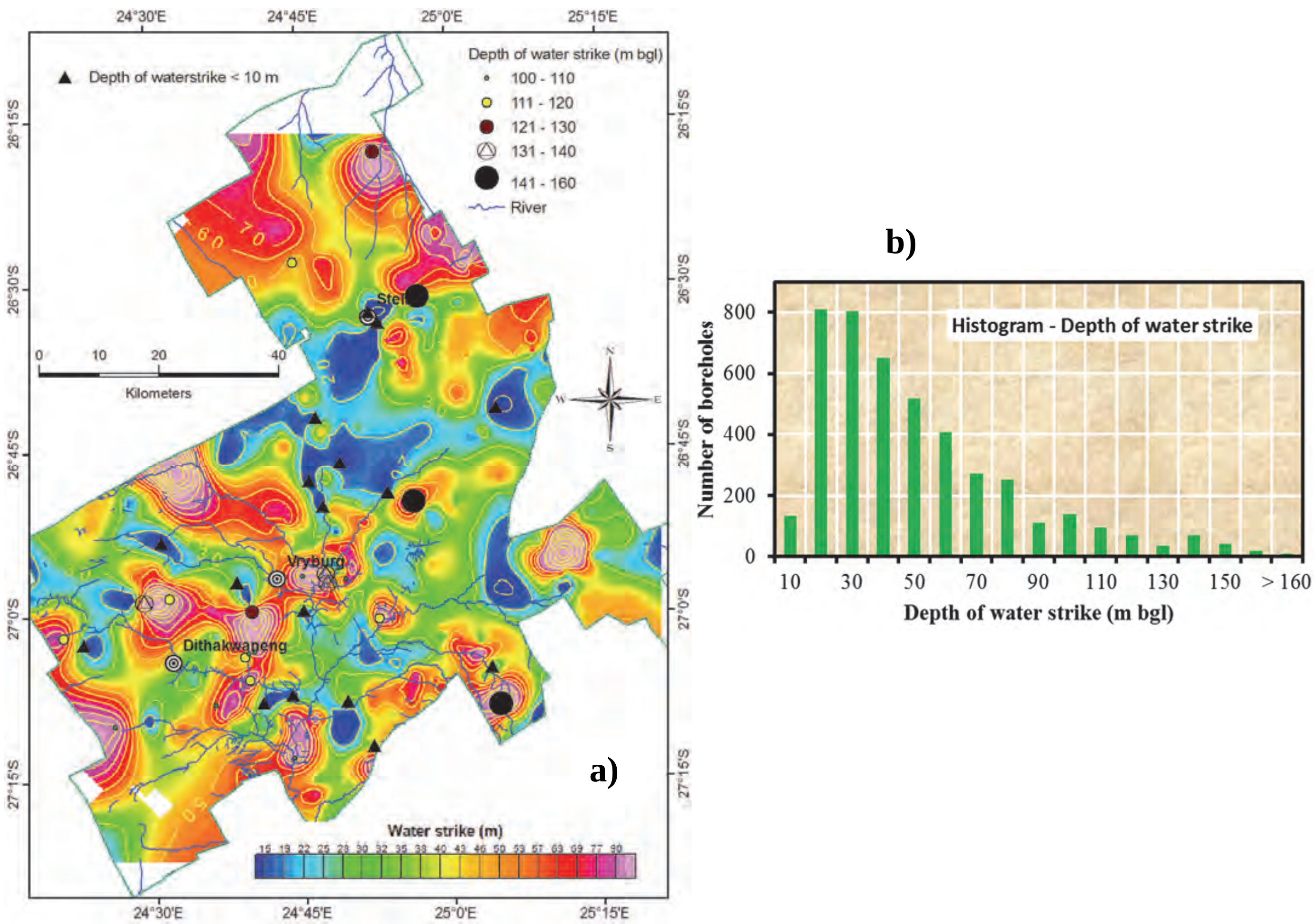


Figure 7: **a)** Depth of water strike interpolated from the existing borehole data with a grid spacing of 200 m. **b)** Histogram of depth of water strike.

Many factors control the variability in the depth of water strike, which include; rock type, fracture connectivity and density, topography, site characteristics, amongst others. In general, the depth of water strike within the carbonate rocks is relatively shallow varying from 4.5 m to 120 m. In addition, boreholes drilled closed to river channels are relatively shallow owing to the presence of near-surface gravel deposits.

3. Groundwater Quality

3.1 Background

Protection and management of groundwater resources includes the establishment and implementation of water quality standards that are designed to minimize impacts to ground water-dependent ecosystems. The quality of groundwater is as important as quantity and investigation of groundwater potential must also include characterization of its quality in order to determine whether water resource meets the minimum standard required for domestic and agricultural use. This ensures the needs of groundwater for domestic and agricultural use as well as its protection from anthropogenic and natural sources of contaminants are met.

The quality of groundwater is a function of physical and chemical parameters that are greatly influenced by geological formations and anthropogenic activities (Muqtada et al., 2012). When analyzing groundwater quality, it is common practice to characterize physical parameters, *viz.* temperature, pH, alkalinity, TDS, hardness and electrical conductivity (EC). In addition, it is important to analyze the concentration of major constituents and minor or trace elements as both cations and anions that can be derived from dissolution of minerals during interaction between groundwater and the host rocks.

In South Africa, the quality of public drinking water is regulated by guidelines set by the Department of Water Affairs. The guidelines are expressed as numerical constants indicating the maximum concentration of certain dissolved or suspended material which can be harmful to human health if consumed beyond the prescribed limit. In this study, selected physical properties and the concentration of certain dissolved and suspended materials were used to assess the quality of water with respect to the SANS241 guide line presented on Table 1. Below summary of the water quality is presented.

Table 1: Guidelines for groundwater quality as set out by the SANS241 and WHO

Parameter	WHO Guideline (mg/l)	SANS241 Guideline (mg/l)		Possible Impact
		Guideline value (mg/l)	Maximum acceptable value (mg/l)	
Total Hardness as $CaCO_3$	500	0	450	Scale and scum
Chloride	250	<300	600	Taste and corrosion
Fluoride	1.5	1.5		Dental/skeletal fluorosis
Nitrate (as NO3) (as NO3-N)	50 11.3	<11	<11	Blue baby syndrome (Acute health)
Nitrite as NO2	3	<0.9		Similar impact as NO_3
TDS	1500	<1200		Taste, corrosion/encrustation
pH	-	5 < pH > 9.7	10	High:- Taste; low corrosion
EC	-	170 mS/m		Gastrointestinal disease
Sodium as Na		<200		

3.2 Electrical conductivity

Electrical conductivity (EC) is a measure of the ability of water to conduct electricity. This physical property of water is owing to the presence of ions in water such as carbonates, bicarbonates, sulfates, nitrate, sodium, potassium, magnesium and calcium, all of which carry electrical charges. Most organic compounds which dissolve in water do not dissociate into ions, consequently they do not affect the electrical conductivity of water. The EC is therefore the proxy to TDS, since electrical conductivity of water is a function of the amount of dissolved solids in water.

The EC varies widely and the maximum limit for drinking water according to DWA guide line is 170 mS/m at 30^0C (SANS241, 2011). In the present study area, EC varies from 8 to 900 mS/cm with high values located in the southern part of the area (Figure 8a). The high electrical conductivity located north of Vryburg and around Stella towns possibly attributed to evapotranspiration. However these values may not be an indicative of the presence of harmful dissolved substances in groundwater as locally elevated EC is common in any hydrogeologic settings.

3.3 Total Dissolved Solids (TDS)

The most important indicator of water quality is the Total Dissolved Solids (TDS). It is the measure of the amount of various salts dissolved in water and it quantitatively provides information about the amount of major ions in water. Virtually, all natural waters contain varying concentrations of TDS as a consequence of dissolution of minerals in rocks, soils and decomposed plant materials. When water slowly flows through various types of rocks, it chemically reacts and spatially and temporally varies in TDS. The TDS of natural water therefore depends on the characteristics of the geological formations through which groundwater circulates or in contact with. As stated in Section 5.1, the TDS concentration is directly proportional to the electrical conductivity of water. Since EC is much easier to measure than TDS, it is routinely used as a proxy to TDS concentration.

The gridded TDS data for the present study area is shown on Figure 8b. The values > 1000 mg/l are shown using graduated circle and superimposed on the gridded data to show areas where groundwater quality appears to be deteriorated. Generally, 75% of the total number of boreholes in the study area is good quality (Figure 8c). The remaining 20% shows elevated TDS, i.e. > 1000 mg/l, and the wellfields which fall under this category are located around Stella and southeast of Vryburg. There are seven boreholes where the concentration of TDS exceeds 2000 mg/l (Figure 8b) indicating poor groundwater quality. Possible causes of high TDS concentration may include runoff, irrigation return flow, high evaporation and discharge of underlying bedrock aquifers. In general, the TDS values presented in Figure 8b shows good quality with the exception of a few boreholes. The use of boreholes with high TDS values (e.g. TDS > 2000 mg/l) can be discontinued until verified based on analysis of samples.

3.4 pH

The pH is one of the essential parameters that are used for assessment of the quality of water and it can be described in terms of the inverse logarithmic concentration of hydrogen atom in water. Conditions that favour production of hydrogen ions result in the lowering of pH, i.e. an acidification process. Alternatively, conditions that favour neutralizing hydrogen ions result in an increase in pH, i.e. an alkalinisation process. The pH of water does not have direct health threat except at extremes (e.g. pH >11.0). The adverse effects of pH result from the solubility of toxic heavy metals. According to DWAF guidelines for drinking water quality (SANS241, 2011), the pH for freshwater should vary from 7.0 to 8.5. In the study area, it varies from 6.5 to 10.2, with the lower values (6.5) corresponding to the Ventersdorp acidic lavas (Figure 9b). High pH values ranging from 7.46-10.20 are observed in areas underlain by Dwyka tillite, dolomite, limestone, felsic lavas and aeolian sand (*conf*. Figures 2 and 9b). An increase in Ca and Mg content within the carbonate rocks consisting of limestone and dolomite may have contributed to the elevated pH of groundwater observed in the southern part of the study area. However, the percentage of boreholes with high pH (> 8.5) is very low (2%) compared to those with normal pH of water ranging from 6.5 to 8.5 (Figure 9c).

3.5 Chloride

Chloride is present in all natural water at greatly varying concentrations depending on the geochemical conditions. The chlorides of sodium, potassium, calcium and magnesium are highly soluble in water. The major sources of chloride in groundwater are igneous and metamorphic rocks such as granite-gneiss. Chloride inputs to groundwater can also arise from irrigation return flows, sewage effluent discharges and various anthropogenic sources. Because of sewerage disposal and leaching of saline residues in the soil, anomalous chloride concentrations may occur.

Chloride is of concern in domestic water supplies, because elevated concentrations impart a salty taste to water and accelerates the corrosion rate of metals. However, the taste threshold and the corrosion acceleration threshold of chloride are dependent on the action of other water quality constituents such as associated cations, the pH and the calcium carbonate concentration. High concentrations of chloride can also be detrimental to chloride-sensitive plants.

According to DWAF guidelines set out for drinking water, chloride concentration exceeding 600 mg/l is not recommended for domestic use (DWAF, 1996). The threshold for an increased corrosion rate is approximately 50 mg/l. At chloride concentrations greater than 200 mg/l, there is likely to be a significant shortening of the lifetime of domestic appliances as a result of corrosion.

Figure 9a shows the concentration of chloride in groundwater. The high values > 500 mg/l are shown using graduated symbol superimposed on the gridded image. The boreholes with high chloride concentrations (>500 mg/l) are located north and south of Stella, and north, southwest and west of Vryburg. The elevated chloride content north of Stella is because of the Kalahari sand consisting mainly of fine sand, silt and clayey deposits which have accumulated mineral salts. This was enhanced by the low rainfall and runoff as well as high evaporation. The quality of groundwater at boreholes showing high chloride content is not suitable for domestic use because the water can significantly affect the lifespan of household appliances such as materials made up of iron and copper. Domestic water use at boreholes that are located west of Vryburg and north of Stella should be discontinued, because the high chloride content may cause dehydration in infants.

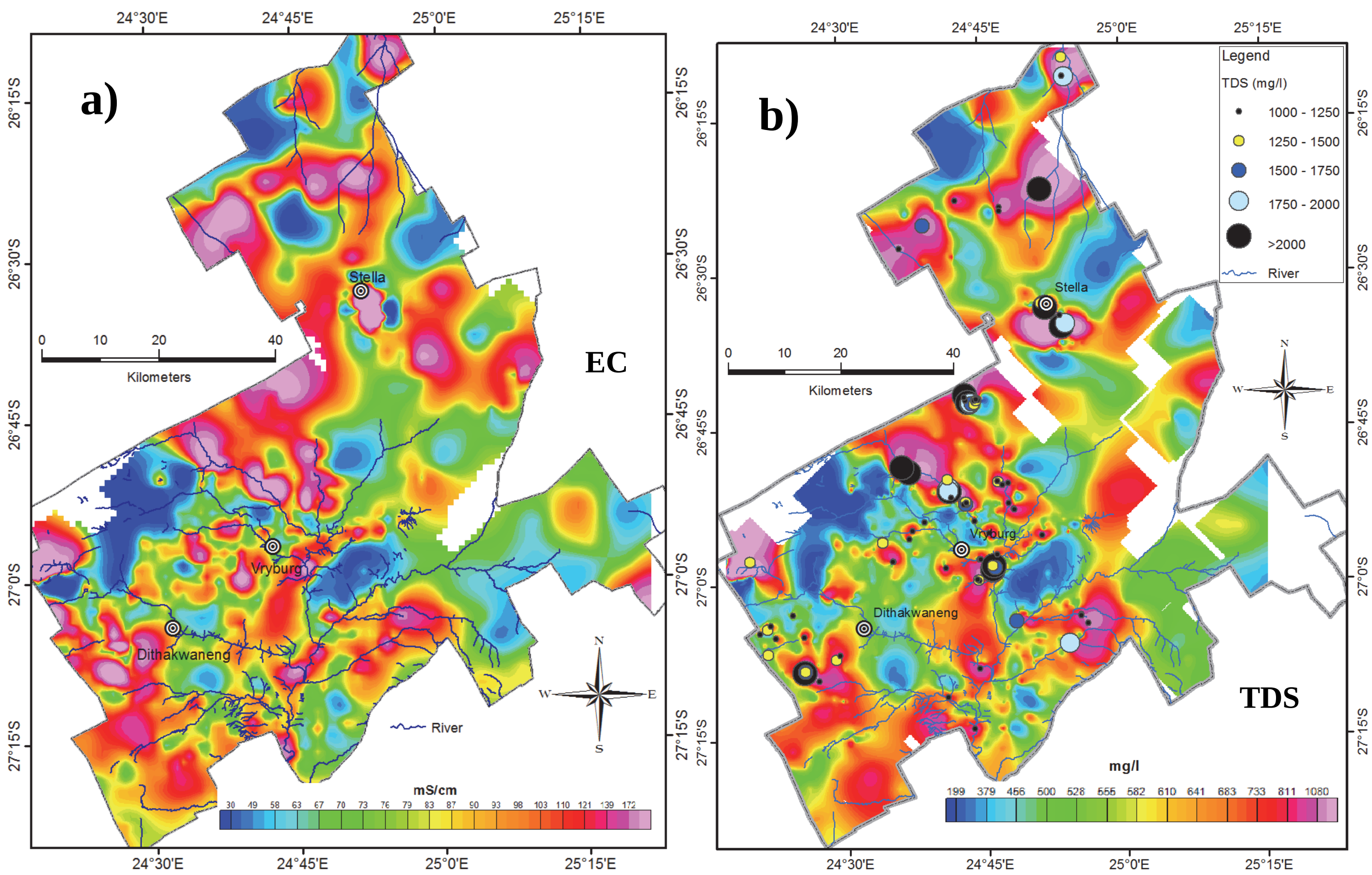


Figure 8: **a)** Electrical conductivity and **b)** TDS map of the study area.

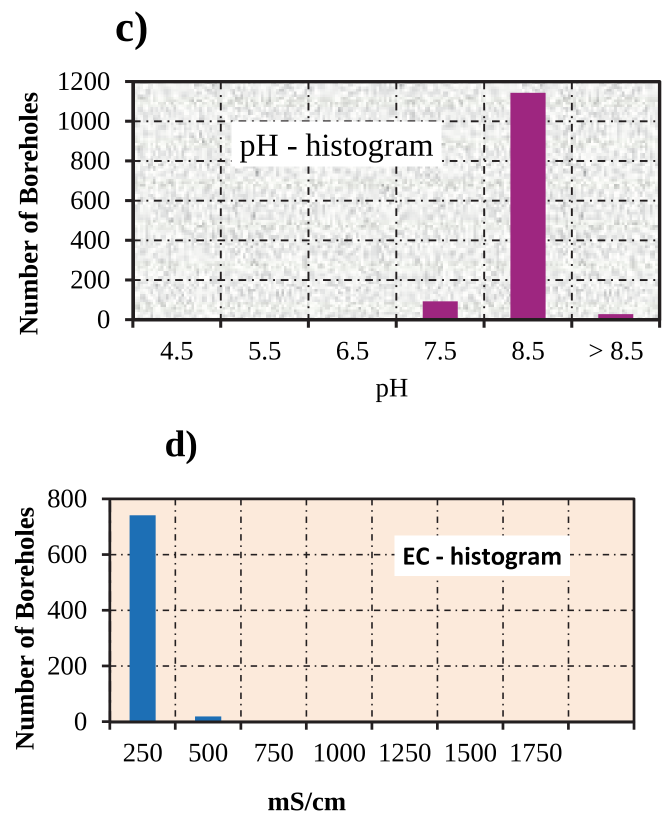
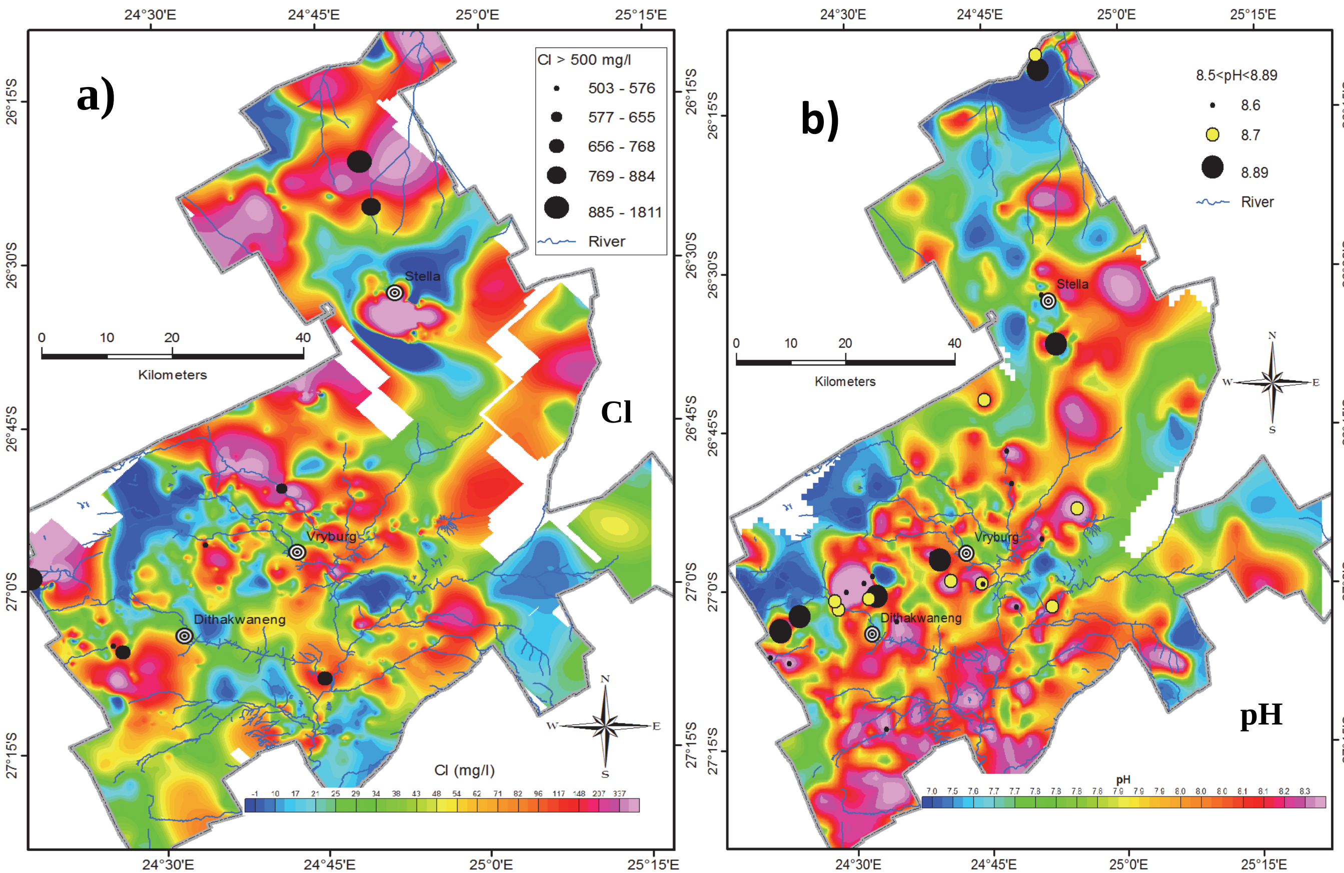


Figure 8: **c)** and **d)** Histogram of TDS and EC respectively.



Figures 9: **a)** and **b)** Chloride concentration in groundwater and pH, respectively

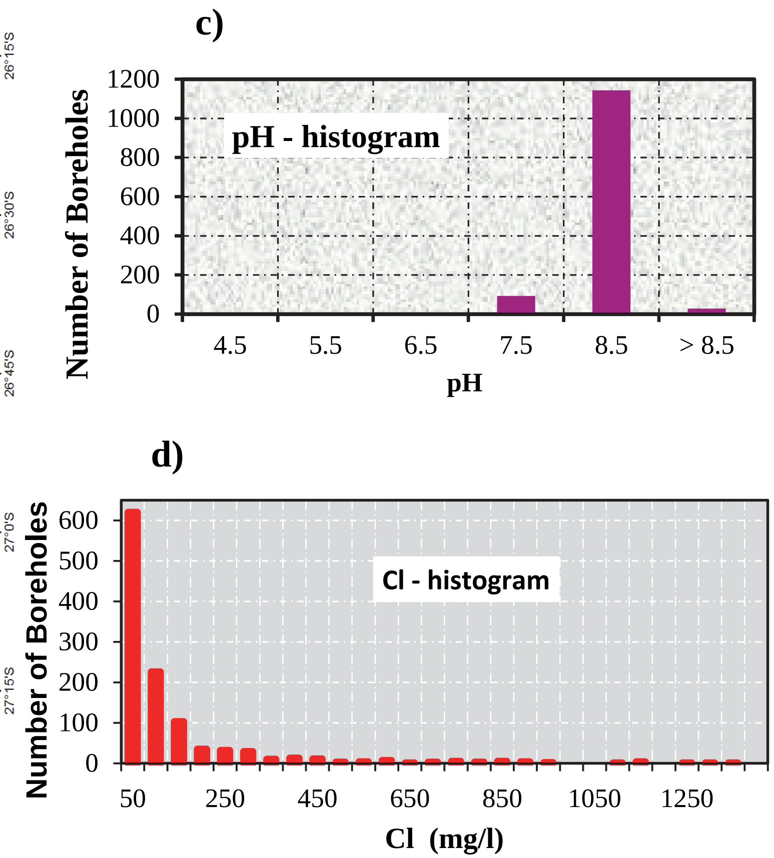


Figure 9: **c)** and **d)** Histogram of pH and chloride, respectively

3.6 Total Hardness

Water hardness is a measure of the capacity of water to precipitate or deteriorate soap and it is mainly controlled by the concentration of calcium and magnesium and other polyvalent cations. The total hardness of water is the sum of Ca and Mg concentrations expressed as calcium carbonate in mg/l. Water hardness is mainly due to elevated carbonates, bicarbonates, sulphates and chlorides of Ca and Mg. It is not caused by one substance, but by a variety of dissolved polyvalent metallic ions, predominantly calcium and magnesium although other cations contribute (e.g. iron, manganese, zinc, etc.). Excessive softness of water can also be a problem by causing corrosion which is of concern where copper is used for plumbing installations.

According to the DWAF guidelines (SANS242, 2011), excessive hardness in water used for domestic purposes causes two main problems:

- It forms scale on heat exchange surfaces such as cooking utensils, hot water pipes, kettles and geysers.
- It results in an increase in soap required to produce lather when bathing and in household cleaning.

According to DWAF (1996) guide line for quality of drinking water, the total hardness should be limited to between 50-100 mg/l where possible.

The total hardness gridded image of the study area is shown on Figure 10. The wide range of variation of total hardness is due to the influence of the catchment geology, e.g. areas underlain by carbonate rocks in the southern part of the area exhibit corresponding elevated levels of total hardness. In addition, high Ca content within the felsic lavas of the Ventersdorp and granitic intrusive rocks of the Kraaipan greenstone belt in the north are the reason for the very hard water observed in those areas. The high total hardness of water in the southern part of the area is due to the presence of carbonate rocks (limestone and dolomite) consisting of Ca and Mg. In addition, elevated total hardness around Vryburg and Stella towns may also be attributed to non-bicarbonate and carbonate hardness such as chloride hardness. The chloride concentration exceeding the regulatory standards north of Vryburg and Stella, possibly contributed to the observed elevated levels of total hardness.

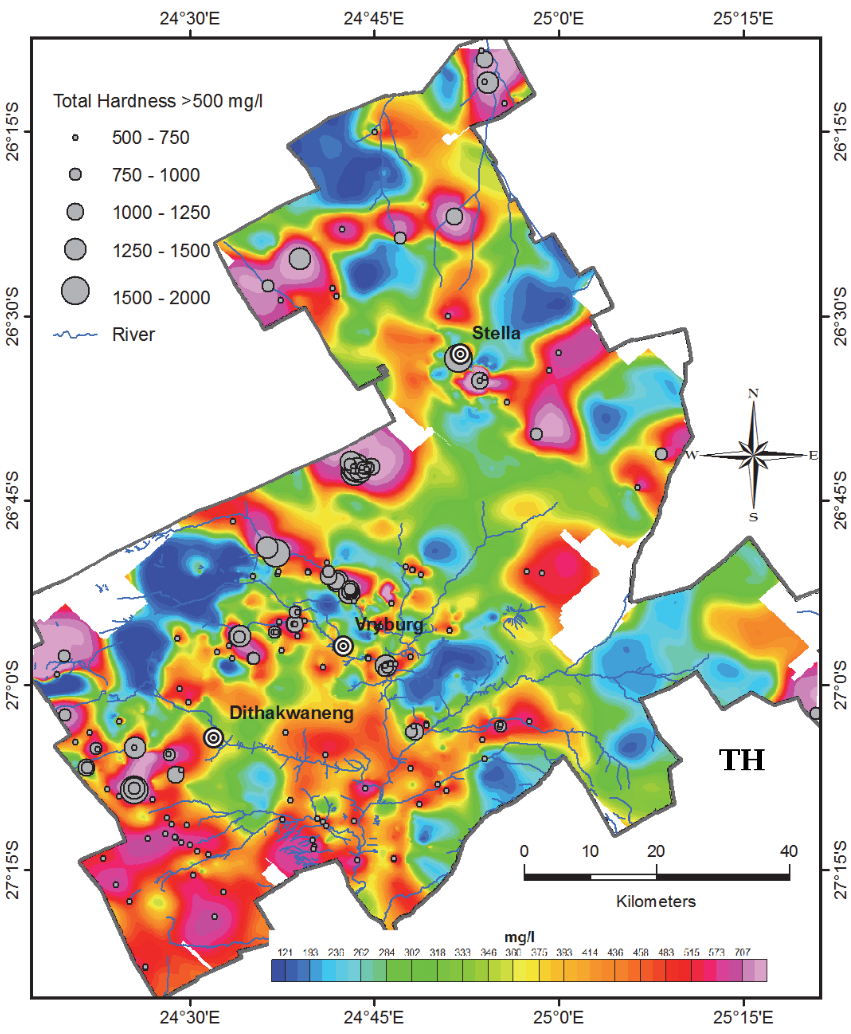


Figure 10: Total hardness of groundwater. The superimposed graduated symbols show total hardness exceeding 500 mg/l.

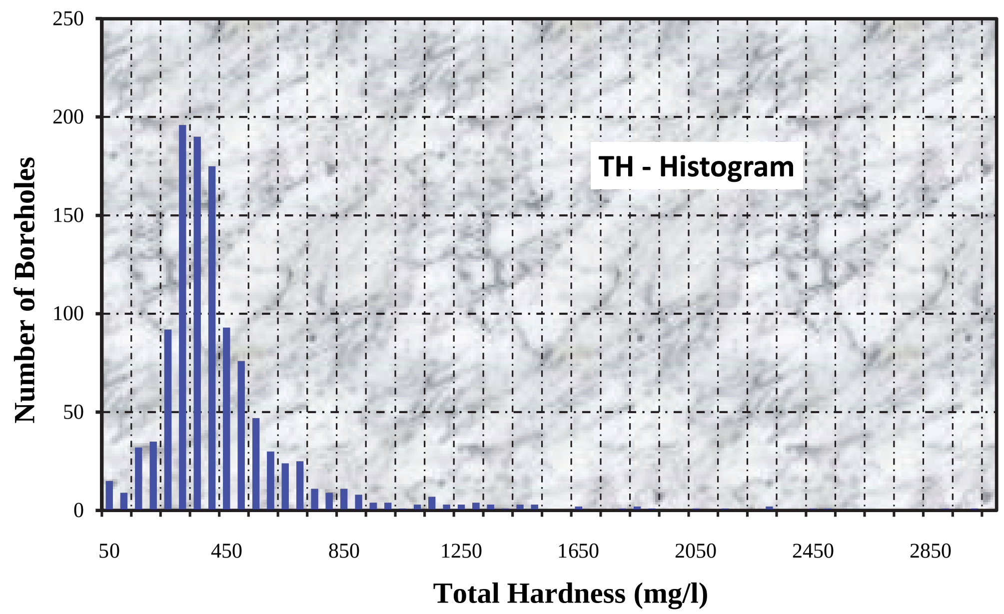


Figure 11: Histogram of total hardness of groundwater.

3.7 Nitrate as NO_3+NO_2-N

Nitrate and nitrite are naturally occurring ions that are part of the nitrogen cycle. The nitrate ion (NO_3) is the stable form of combined nitrogen for oxygenated systems. Although chemically inert, it can be reduced by microbial action. The nitrite ion (NO_2) contains nitrogen in a relatively unstable oxidation state. Chemical and biological processes can further reduce nitrite to various compounds or oxidize it to nitrate. Under oxidising conditions nitrite can be converted to nitrate, which is the most stable positive oxidation state of nitrogen and far more common in the aquatic environment and soil than nitrite.

An elevated nitrate concentration in groundwater can be due to high nitrogen input into the environment as a result of human activities. Intensive agricultural production, domestic and industrial wastes, sewage and atmospheric nitrogen pollution are considered to be the main sources of nitrate contamination in water. Nitrate is also formed where wastes and sewage rich in nitrogen are biologically decomposed under aerobic conditions. Private households and livestock are examples of the origin of nitrate. According to the DWAF guideline set out for drinking water, nitrate concentration exceeding 10 mg/l-N results in adverse health effects in infants of less than 2 years old (DWAF, 1996).

Nitrate concentration in groundwater is presented on Figure 12. It varies from 1 to 40 mg/l –N, with high values occurring around Stella. Areas that are located 12 km southeast of Stella town encompassing Middlekop and Spitzkop wellfields are characterized by elevated nitrate content, exceeding the accepted range of drinking water standard as set out by the Department of Water Affairs. Generally, areas around Stella (west, east and north of the town) show elevated nitrate content, probably resulting from cattle farming and poor sewerage systems associated with informal settlements. Similar elevated nitrate content south and southwest of Vryburg are owing to cattle farming and human waste within the rural areas. In general, the main sources of nitrate in the study area are animal excrement at the watering points and nitrogenous fertilizers applied on irrigated areas (e.g. Spitzkop and Middlekop wellfields). Areas that pose potential health risks to infants and susceptible areas of chronic disease to adults resulting in from high nitrate concentration in drinking water are shown using graduated symbols (circles shown by blue and pink colours) (Figure 12). These boreholes can be abandoned from domestic water use until the nitrate levels are verified based on standard laboratory analysis.

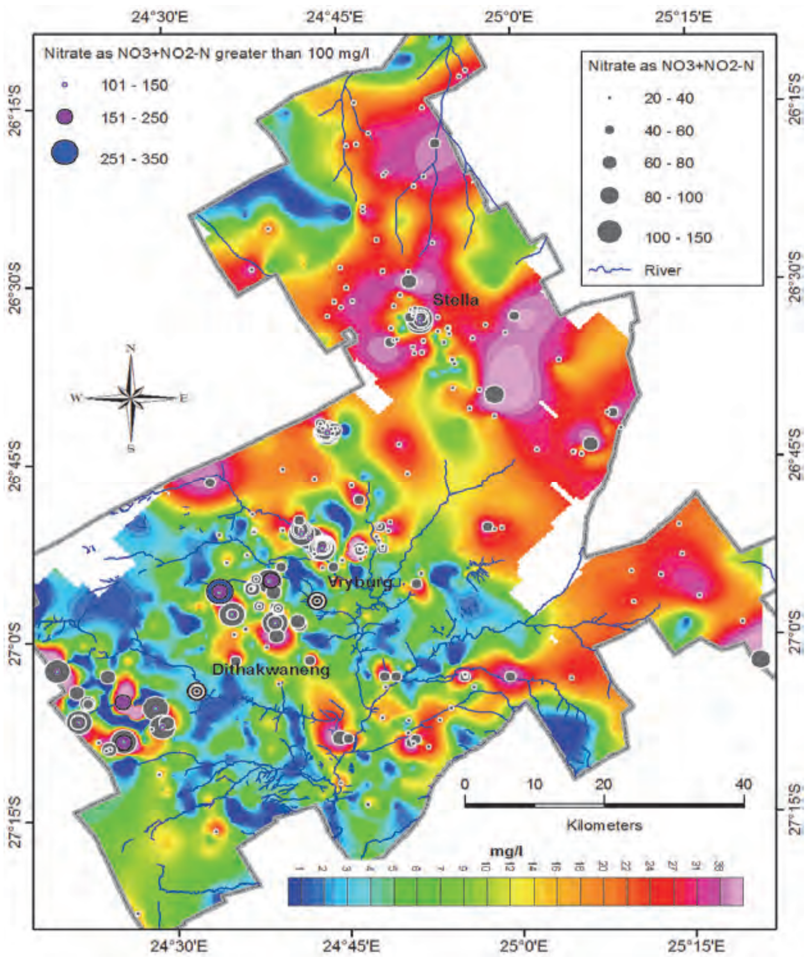


Figure 12: Concentration of nitrate in groundwater.

3.8 Fluoride

Fluoride minerals are fluor-spar (CaF_2) and fluor-apatite, a calcium fluoro-phosphate are mainly associated with volcanic rocks. Hydrogen fluorine is one of the most soluble gases in magmas and comes out partially during eruptive activity. Like volcanic ash, fly ash from the combustion of fossil fuels also account for high fluoride. In addition, phosphate containing fertilizers add up to the fluoride content in soil and groundwater. Fluoride can also be the weathering product of micaceous rock such as granite and gneisses.

The concentration of fluoride in groundwater increases owing to rock-water interaction, long residence time and evapotranspiration. According to DWA guidelines set out for drinking water (DWAF, 1996), the recommended concentration of fluoride in potable water is ≤ 1.5 mg/l (Table 1). As the concentration increases above the prescribed limit, it can cause the likelihood of skeletal fluorosis with crippling, as well as the loss of teeth.

In the present study area, there are spots of relatively high concentration of fluoride as shown on Figure 13. The high values are located north of Stella and west of Vryburg. The one located north of Stella may be attributed to granite-gneiss and the weathering product of felsic intrusive rocks that are associated with the Kraaipan greenstone belt. In the far western part of Vryburg town the high fluoride values may be due to localized granitic intrusion though not possible to verify based on field geological observation. In general, the limited fluoride data covering the present study area makes it difficult to assess the quality of groundwater.

3.9 Water Quality for Agriculture

The relative concentration of Na with respect to Ca and Mg is an important indicator as to whether the water is suitable for agricultural use. If water used for irrigation is characterized by elevated sodium and low Ca and Mg content, the cation exchange complex may become saturated with sodium. This can destroy the soil structure owing to dispersion of the clay particles and this in turn decreases permeability, and aeration of the soil which will affect plant growth by limiting the uptake of water and air. The Sodium Adsorption Ratio (SAR) can be used to assess the risk resulting from anomalous sodium content in groundwater (Fetter, 1994):

$$SAR\ (\text{meq/l}) = \frac{\mathrm{N_a}}{\dfrac{\sqrt{Ca + Mg}}{2}}$$

As shown on Figure 14, 98% of the samples have good water quality suitable for irrigation, while the remaining 2% are unsafe for agricultural use. Areas shown by graduated symbols >10 indicate elevated SAR and Na content which follows major rivers that are located northeast and southwest of Vryburg town. In addition, some boreholes located around Stella (Middlekop-8 well-field) are characterized by anomalous concentration of Na. Another boreholes located in Stella and north of Vegter wellfields are characterized by SAR > 30 and Na concentration > 50 meq/L. These two boreholes may pose a risk if used for irrigation.

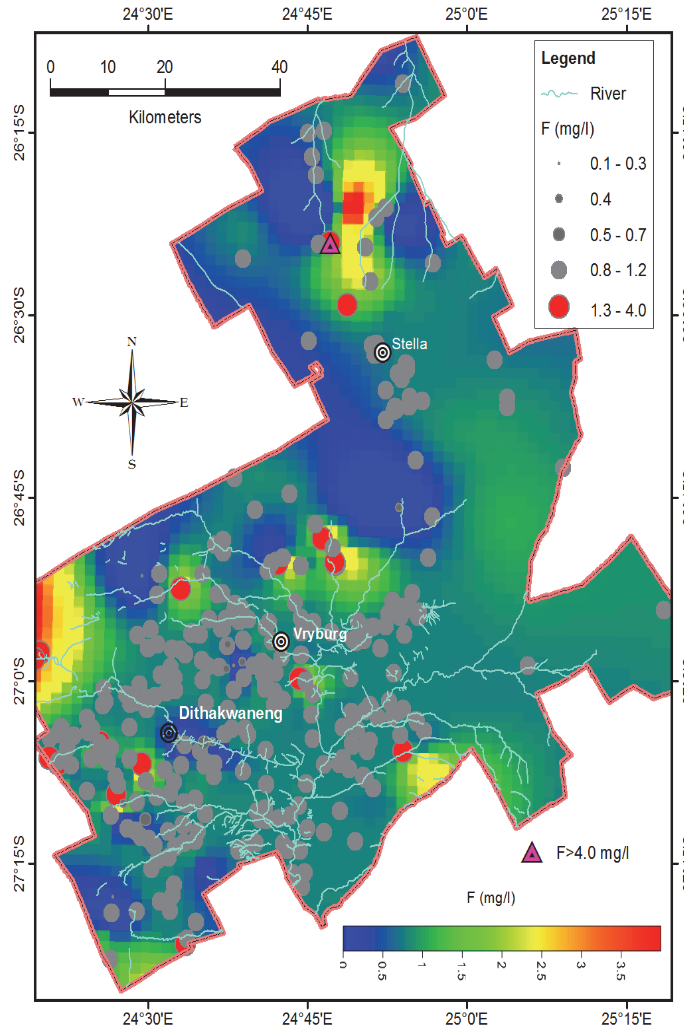


Figure 13: Concentration of fluoride in groundwater.

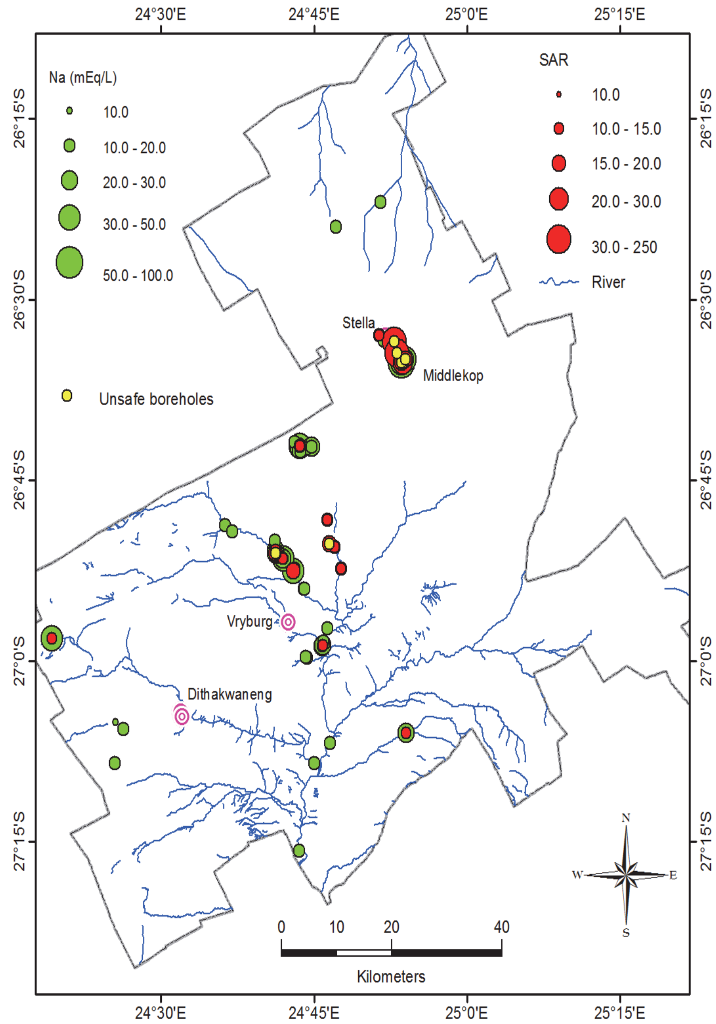


Figure 14: High concentration of Na and SAR in the study area. The yellow circles show boreholes with SAR> 15.

4. Environmental Isotopes Studies

4.1 Overview of Application

Stable and radiogenic isotopes are commonly used for understanding the origin, age and chemical evolution of groundwater. Tracing groundwater by means of environmental isotopes offers unique and supplementary information about the origin and movement of groundwater and its dissolved constituents. It also allows a quantitative evaluation of mixing and other physical processes such as evaporation and isotopic exchange processes.

Variations in stable isotope compositions of oxygen and hydrogen constituting the water molecule have been used in several studies to investigate the source of water and to understand post-precipitation processes during groundwater recharge. The stable isotopes of water are influenced directly by the atmospheric processes (e.g. water vapour, condensation or evaporation) and during groundwater recharge (Clark and Fritz, 1997). All these processes result in variation in the isotopic composition of water.

Stable isotope compositions commonly reported as the atomic ratio *R* of the less frequent to the abundant isotope changes of a sample. It is usually determined by mass spectrometry and expressed as delta (δ) values referring to a certain standard sample used for isotope studies. The δ values are calculated using: δ (in ‰) = $(R_x/R_s\text{-}1)1000$. Here R denotes the ratio of heavy to light isotope (e.g. $^{18}O/^{16}O$ or $D/^{1}H$) and R_x and R_s are the ratios in the sample and standard, respectively. δD and $\delta^{18}O$ values are reported relative to SMOW (standard mean ocean water) or the equivalent of Vienna SMOW.

In spite of the great complexity in different components of the hydrological cycle, $\delta^{18}O$ and δD in meteoric waters behave in a predictable manner and correlate and defines the global meteoric water line (GMWL). According to Craig (1961), this correlation can be expressed as:

$$\delta D = 8\delta^{18}O+10$$

The above relationship can be presented in more generalized expression:

$$\delta D = s\delta^{18}O + d$$

Where '*s*' is slope and '*d*' is deuterium excess (*d*-excess). The slope *s* depends on many factors, but most importantly the relative humidity of vapour source, evaporation and mixing. As the humidity decreases, total evaporation increases which give rise to gentler slope relative to the GMWL. The process of mixing of water with varying isotopic composition gives a slope that is gentler than the evaporation line. On the other hand, as the humidity increases, the change of $\delta\ ^{2}H$ and $\delta\ ^{18}O$ values becomes insignificant which give rise to a line that plots along a slope of GMWL. However, as the relative humidity decreases, total evaporation fractionates leading to enrichment of $\delta^{2}H$ and $\delta^{18}O$ in the residual water relative to the initial isotopic composition. The resulting $\delta^{2}H$ versus $\delta^{18}O$ gives a line that falls below or to the right of the GMWL with a slope <8.0. The decrease in slope is due to the fact that more enrichment in $\delta^{18}O$ than $\delta^{2}H$ occurs during rainfall. The intersection between the GMWL and the line derived from the observation points indicates the original isotopic composition of water prior to kinetic and equilibrium fractionation processes (Clark and Fritz, 1997).

The '*d-excess*' is expressed as $d=\delta D\text{-}8*\delta^{18}O$ and represents the relative magnitude of kinetic fractionation in different water masses (Clark and Fritz, 1997). It is the measure of the deviation of given data points from a line with a slope of 8. The *d*-excess provides valuable information about the moisture source of precipitation. It is mainly related to the meteorological condition at the vapour source region and kinetic fractionation during evaporation of falling rain drops (Clark and Fritz, 1997). The *d*-excess also indicates the distance the moist air-mass travelled from vapour source. For example, as latitude increases, the *d*-excess decreases, i.e. the moist air-mass becomes dry as distance from the source of moisture increases. In addition, if excessive secondary evaporation of raindrops occurs during its fall beneath the cloud-base (Clark and Fritz, 1997) this effect will enrich the isotopic content of precipitation, change its actual isotopic signature and lower its *d*-excess. Therefore the *d*-excess can be an indication of the climatic condition at the vapour source and the relative humidity as well as secondary evaporation that takes place under the cloud.

The length of δ^2H versus $\delta^{18}O$ line is important in indicating isotopic variation resulting in from total evaporation under different relative humidity and temperature conditions. According to Clark and Fitz (1997), if the plotted line is short for a given relative humidity (e.g. 50%), it shows that the water changes little during the entire or total evaporation processes.

Deviations of regional or local meteoric water lines (LMWL) from GMWL help to identify the climatic and geographical dimensions of hydrological processes and provenance of different water masses. According to Gupta and Deshpande (2005), during exchange process between reservoirs, the process of isotope fractionation can be equilibrium type (e.g. condensation of water drops in the cloud) or non-equilibrium or kinetic type (e.g. diffusion or evaporation from a water body). In general, heavier isotopes preferentially tend to stay with the more condense phase in any fractionation process. The equilibrium fractionation is strongly dependent on temperature, whereas the kinetic fractionation is strongly dependent on relative humidity. According to Gupta and Deshpande (2005), kinetic fractionation for ^{18}O is more than that for deuterium. Therefore, during the process of evaporation, the relative enrichment of residual water is more for ^{18}O than deuterium.

Within the cloud, equilibrium fractionation between vapour and the condensing phases preferentially partition ^{18}O and ^2H into rain or vapour. As a result, along the trajectory of the air mass, each rainout preferentially removes the heavy isotopes from the vapour. The remaining vapour then becomes progressively depleted both in ^{18}O and ^2H. Whereas each rainout gives isotopically enriched rain with respect to the remaining vapour, it will be depleted with respect to earlier rainout.

As evaporation proceeds, the residual water not only enriches in heavier isotopes, but also shows progressively low *d*-excess values due to relatively more enrichment of ^{18}O. Due to the involvement of evaporation, most meteoric and subsurface processes shift the $\delta\ ^{18}O$ and δ^2H signatures of water to a position below the LMWL. According to Kendall and McDonnel (1998), it is rare to find precipitation or groundwater that plots above the LMWL, i.e. showing higher *d*-excess. But, re-evaporation of precipitated water under low-humidity regions creates vapour masses with isotopic content that plots above the LMWL.

To understand the processes that modify the isotopic composition of groundwater in the study area, stable isotopes of water molecule (^{2}H and ^{18}O) and radiogenic isotope (tritium) were used. In the following sections, sampling, analysis and interpretation are presented.

4.2 Water sampling and analysis

In November 2012, groundwater samples were collected from 18 wells that are currently supplying water to the Naledi Local Municipality (Figure 15). The samples were collected using clean HDPE (high density polyethylene) bottles for isotope and hydrochemical analysis. The bottles were rinsed with the same water prior to sampling. They were filled completely to the top for major ions analysis and capped tightly to avoid evaporation and exchange of sample water with atmospheric materials such as oxygen and CO_2. Two sample bottles, one each for cations and anions were required for sampling. The third bottle was used for sampling water for isotope analysis. The samples for the cations were filtered on site using 0.45 μm filter and acidified with concentrated nitric acid (HNO_3) to give a pH <2 that preserves the metals dissolved and prevent precipitation. Immediately after sampling, they were stored in a cooler box till the samples were being delivered to the Laboratory. Physiochemical parameters pH, EC, TDS, redox potential (E_h) and alkalinity (bicarbonate concentration) were measured on site using an Aquaread GPS Aquameter 200 multi parameter water quality meter.

The water samples were analyzed at the Waterlab in Pretoria using inductively coupled plasma optical emission spectrometry (ICP-OES). The equipment used was a Perkin Elmer Optima 2100 DV ICP-OES spectrometer. The major anions composition (Cl^-, SO_4^{2-}, NO_3^-, NO_2^-, PO_4^{3-} and NH_3) were analysed using a Thermo Scientific Aquakem 250 photometric analyzer. Fluoride (F^-) was analyzed using the ion selective electrode (ISE) methodology with a Radiometer Analytical IONcheck 45 ion analyser. The concentration of HCO_3^- anions in each sample was determined by titrating hydrochloric acid (HCl) with each sample.

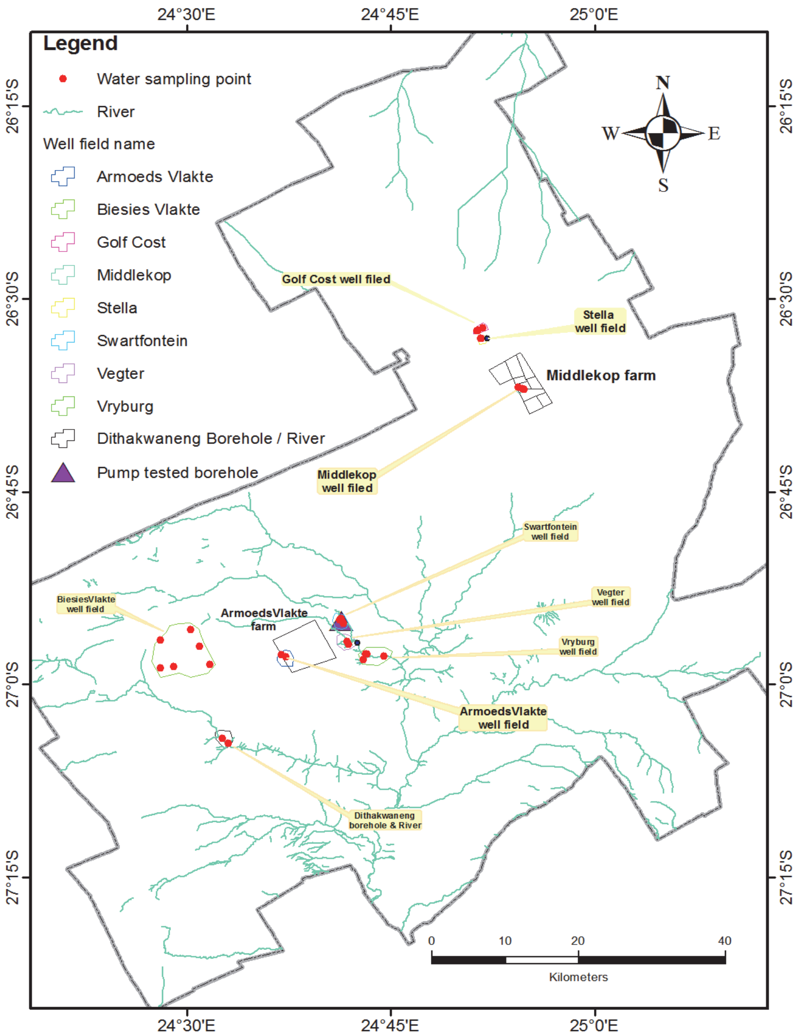


Figure 15: Location map of well-fields at which groundwater sampling was carried out.

4.3 Analysis of Environmental Isotopes

Stable isotopes ratios $^2H/^1H$ and $^{18}O/^{16}O$ were analysed at iThemba LABS in Gauteng Province. The $^{18}O/^{16}O$ ratio was analysed by equilibrating a small aliquot of water with CO_2 gas, which is then introduced into an isotope ratio mass spectrometer (IRMS). The $^2H/^1H$ ratio was analysed by equilibrating water with H_2 gas in the presence of platinum catalyst. The analytical results were reported as δ^2H and $\delta^{18}O$ relative to Vienna Standard Mean Ocean Water (SMOW) as described by Gonfiniatini (1981) and Paternoster (2007).The analytical results of $\delta^{18}O$ and δ^2H values were calculated using the equation below:

$$\delta^{18}O\ (‰) = \left[\frac{(^{18}O_f/^{16}O_f}{(^{18}O_S/^{16}O_S}-1\right]X1000 \quad \text{and} \quad \delta^2H\ (‰) = \left[\frac{(^2H_f/^1H_f}{(^2H_S/^1H_S}-1\right]X1000$$

Where O_f, O_s, H_f and H_s are oxygen and hydrogen isotopes sample from the field and standard, respectively. The results for $\delta^{18}O$ and δ^2H is reported in parts per thousand and presented on Table 2. The precision of measurements are ±0.1‰ and ±0.3‰ for ^{18}O and ^2H, respectively.

Table 2: Results of analysis of environmental isotopes samples

Field Name	δD‰ SMOW	$\delta^{18}O$‰ SMOW	*d*-excess	TU	±	Depth (m)
ARV 01	-27.8	-5.26	14.28	1.1	0.3	8
BV 01	-28.5	-5.66	19.34	1.0	0.2	5
BV 04	-22.0	-4.66	13.68	1.2	0.3	12
DTK 01	-27.4	-5.25	14.60	0.5	0.2	10
DTKR 01	-25.7	-4.39	9.42	0.9	0.2	11
MDK 01	-28.6	-5.08	12.04	1.0	0.2	4
ST 02	-29.0	-5.23	12.84	0.6	0.2	3
ST 03	-27.9	-5.23	13.94	0.9	0.2	7
STGC 01	-28.3	-4.60	8.5	0.4	0.2	6
STGC 04	-29.6	-4.74	8.32	0.7	0.2	2
SWF 01	-30.1	-5.23	11.74	1.1	0.3	1
SWF 06	-27.5	-4.16	5.78	1.3	0.3	9

4.4 Interpretation of Isotopes Data

The Global Meteoric Water Line (GMWL) and Local Meteoric Water Line (LMWL) are commonly used as a reference for interpretation of the isotopic composition of surface water and groundwater. As the study area is situated in similar climatic and physiographic setting with Pretoria, the long-term monitoring of climatic condition recorded at Pretoria meteorological station was used as a Local Meteoric Water

Line (PLMWL). In this study, the δ^2H versus $\delta^{18}O$ was plotted to show the isotopic composition of groundwater and surface water with respect to the GMWL and PLMWL as shown on Figure 16. The regression line fitting the observed isotope data nearly coincides with the PLMWL suggesting that precipitation in Pretoria and in the study area have the same vapour source confirming the validity of PLMWL as a reference for interpretation of the stable isotopic variations of groundwater and surface water in the study area.

According to Craig (1961), if the observed isotope data fall to the right of the PLMWL, the precipitation that recharges groundwater has undergone evaporation prior to infiltration thereby enrichment in the stable isotopes ($\delta^{18}O$ and δ^2H). Groundwater and surface water that have not undergone extensive evaporation fall along the Pretoria LMWL. Those which fall on the left of the PLMWL may indicate fast recharge through vertical fracture or through highly permeable media before being undergone kinetic fractionation.

As shown on Figure 16a, the PLMWL offset to the right with respect to the GMWL indicating precipitation in the study area has undergone moderate-to-high fractionation with respect to the global average relationship between δ^2H and $\delta^{18}O$. The three samples collected from Armoedsvlakte, Biesjiesvlakte wellfields and Dithakwaneng borehole fall along the GMWL signifying direct recharge from precipitation with minimal evaporative fractionation impact. These may be attributed to availability of fractures or primary porosity, allowing fast percolation of rainwater prior to evaporation. The two samples collected from the Stella Golf course plot slightly to the right of the PLMWL indicating deuterium is more depleted and ^{18}O is more enriched, suggesting the impact of evaporation on the isotopic composition of groundwater. The remaining samples plot close to the PLMWL suggesting that the rainwater is isotopically in equilibrium with the standard seawater, i.e. the degree of isotopic fractionation resulting in from evaporation and other processes is not significant. This suggests that precipitation has taken place under an ambient atmospheric condition. These samples were collected from the wellfields located 12 km northwest of Vryburg town where the boreholes penetrate through the dolomite of the Transvaal Supergroup. Faster recharge through existing dissolution channels might be responsible for little modification of the isotopic composition of groundwater.

The $\delta^{18}O$ versus $\delta^{2}H$ plot cluster around $\delta^{2}H$ ~-28 ‰ and $\delta^{18}O$ between -4.5‰ and -5.0‰ showing very limited length or variability in $\delta^{18}O$ versus $\delta^{2}H$ patterns. This shows variability in isotopic fractionation of rainwater or groundwater is not extensive during total evaporation. The linear regression line for the cluster has a slope of ~6.6 which is less than PLMWL and GMWL, suggesting that evaporative losses of rainwater before infiltration which is consistent with the arid-to-semi-arid nature of the climate. The samples exhibit nearly similar isotopic composition varying in $\delta\ ^{2}H$ from -22.0‰ to 30.1‰, while $\delta\ ^{18}O$ varies from -4.16‰ to -5.98‰.

The isotopic composition for the two samples collected at Dithakwaneng River and the borehole located along the river bank slightly differ from each other. The borehole is characterized by $\delta^{2}H$ and $\delta^{18}O$ values of -25.7‰ and -4.31‰, respectively, while the river-water exhibits -5.25‰ and -27.4‰ for $\delta^{18}O$ and $\delta^{2}H$, respectively. The borehole is relatively enriched in both heavy isotopes which can be due to a number of possible scenarios: i) groundwater is not necessarily being recharged by the river, and ii) isotopic exchange is taking place between groundwater and aquifer material, thus enrichment in heavy isotopes of the groundwater.

The *d*-excess is an important parameter in describing the relative magnitude of kinetic fractionation, and has been used extensively for investigation of recharge pathways, flow regimes and correlating groundwater with palaeo-climate (Gupta and Dashpande, 2005). In the present study, the *d*-excess ranges from 5.78‰ to 19.34‰, where the minimum and maximum correspond to Swartfontein and Biesjiesvlakte wellfields, respectively (Table 2). The *d*-excess in the present study area is greater than PLMWL and GMWL, i.e. 7.2 and 10‰, respectively with the exception of Stella Golf course and Swartfontein wellfields (Table 2).

For atmospheric moisture not influenced by secondary evaporation processes, the *d*-excess approximates the *y*-intercept of the GMWL of 10‰, and by analogy to the *y*-intercept of the PLMWL of 7.2‰ (Figure 16). High deuterium excess in groundwater arises from evaporation of rainwater before it percolates to underground.

The elevated *d*-excess observed in almost all boreholes suggests enrichment in heavy isotopes resulting from extensive total evaporation. This is consistent with the arid nature of the climate discussed in Section 1.1. If water from precipitation is evaporated on its way before it percolates to the subsurface, the water molecules with lighter isotopes will contribute preferentially to the isotopic composition of the water vapour and this leads to an enhanced deuterium excess in the percolating water thereby the groundwater (Gupta and Dashpande, 2005). Most of the groundwater from Armoedsvlakte, Biesjiesvlakte, Middlekop and Stella well-fields are characterized by similar *d*-excess that probably indicate a common origin. The boreholes located at Armoedsvlakte and Biesjiesvlakte wellfields are relatively depleted in $\delta^{2}H$ and $\delta^{18}O$ values, however the *d*-excess is elevated contrary to the stable isotopic composition. In addition, the boreholes at Armoedsvlakte, Biesjiesvlakte, Swartfontein wellfields are characterized by nearly similar $\delta^{2}H$ and $\delta^{18}O$ values, but variable *d*-excess, thus questioning the hypothesis that all groundwater around Vryburg have a common origin. This problem may be explained using additional isotope data collected from both surface water and groundwater.

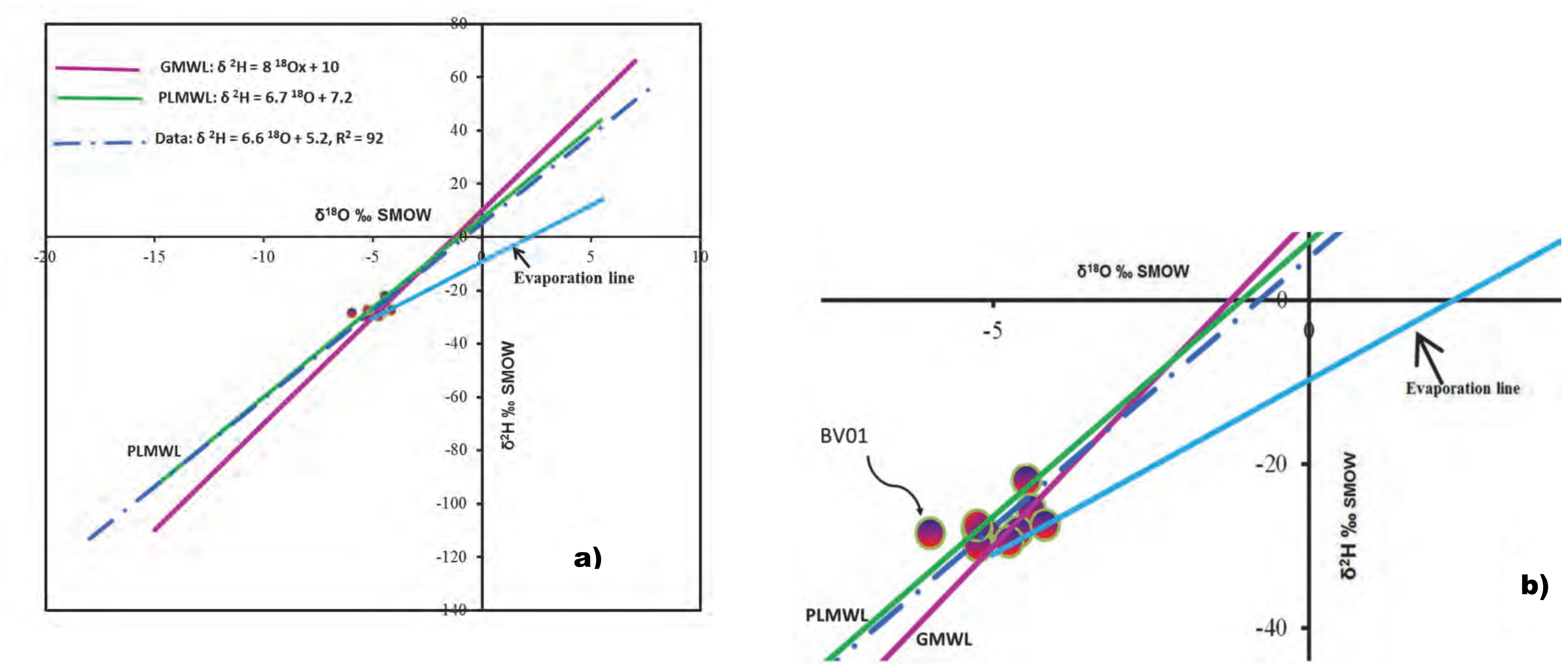


Figure 16: δ^2H versus $\delta^{18}O$ plot showing variation of isotopic composition of groundwater with respect to local and global meteoric water line. The lower graph shows magnification of the position points with respect to PLMWL and GMWL.

It has been suggested that low *d*-excess values around 10‰ indicate palaeo-recharge, while higher *d*-excess values in the range of 15‰ to 25‰ indicate modern-day recharge (Gat and Galai, 1982; Vengosh et al., 2007; Gupta and Dashpande, 2005). Therefore, one could argue that the high *d*-excess values observed at Armoedsvlakte and Biesjiesvlakte wellfields can be recent recharge zone, while Swartfontein and Stella Golf course wellfields are palaeo-recharge zones. However, the low δ^2H and $\delta^{18}O$ values corresponding to Armoedsvlakte and Biesjiesvlakte are not consistent with the elevated *d*-excess. In addition, groundwater collected from Armoedsvlakte and Biesjiesvlakte wellfields show very low tritium values as described in the section below and also shown on Table 2. This may be attributed to mixing of deep groundwater depleted in δ^2H and $\delta^{18}O$ with shallow and recently recharged aquifer system. The hydrochemical data covering the southern part of the study area also provides supporting evidence about mixing of various types of groundwater.

4.5 Interpretation of Tritium

Assuming that piston flow conditions (no dispersion or mixing) are applicable, Clark and Fritz (1997) noted the following guidelines for interpretation of the radiogenic tritium for groundwater investigation:

1. Groundwater that contains tritium less than 0.8TU may correspond to recharge zone prior to 1952.
2. Waters with a tritium content ranging from 0.8 to 4TU may represent a mixture of water that contains components of recharge before and after 1952.
3. Tritium concentration from about 5-15TU may indicate recharge after about 1987.
4. Tritium concentration from about 16-30TU are indicative of recharge since 1953 but cannot be used to provide a more specific time of recharge.
5. Water with more than 30TU probably is from recharge in the1960s or 1970s.
6. Water with more than 50TU predominately is from recharge in the 1960s.

In the present study area, the tritium content in groundwater ranges from 0.4 to 1.3 TU. The low tritium content shown on Table 2 suggests that groundwater in the study area has little or no recently recharged water. The minimum and maximum values correspond to Stella Golf Course and Swartfontein wellfields, respectively. This result is consistent with the stable isotope composition, in that Stella wellfields were recharged under evaporative fractionation process. In addition, all well-fields close to Vryburg town (Armoedsvlakte, Biesjiesvlakte and Swartfontein) are nearly similar in tritium values, while wellfields located around Stella town, including Middlekop and Spitzkop fall under the same category. This suggests that all the wellfields around Vryburg town were recharged by similar source and under the same climatic condition where mixing of old and recently recharged water overprinted the original isotopic composition of groundwater. However, the lack of sufficient tritium data hinders evaluation of the relative timing of groundwater recharge, mixing and rock-water interaction.

5. Hydrochemistry

5.1 The Use of Hydrochemical Data

The concentration of inorganic constituents in groundwater can be controlled by the availability of the elements in the soil or rock through which water flows, solubility and adsorption of constituents and the sequences in which the water has come into contact with various minerals occurring in the geological materials along the flow path (Freeze and Cherry, 1979). As water flows through geological formations, the composition will start to change gradually from dilute rainwater to concentrated ionic composition at depth. This is as a result of progressive interaction between slowly moving water and the minerals that are in contact with water along the flow path. According to Chebotarev (1955) and Back (1960, 1966), variation in water composition evolves along groundwater flow direction even in a homogeneous aquifers due to chemical reactions and the length of time water being in contact with rocks. Different water facies along the flow lines from recharge to discharge area and/or from shallow to deep groundwater circulation can tell us the distinct feature of water evolved from source to sink. Characterization of the composition of water therefore has a great importance to investigate groundwater source, evolution and quality. The use of major ions as natural tracers has become a very common approach to delineate flow paths in aquifers. Generally, the approach is to divide the samples into hydrochemical facies, i.e. groups of samples with similar chemical characteristics that can then be correlated with location, geology, topography or aquifer types among other parameters. The samples with similar chemical characteristics often have similar hydrologic histories, similar recharge areas, infiltration pathways and flow paths in terms of climate, mineralogy and residence time.

The spatial variability observed in the composition of these natural tracers can therefore provide an insight into aquifer heterogeneity and connectivity, as well as the physical and chemical processes controlling water chemistry. In addition, the classification of groundwater into water types on the basis of major ion composition provides a framework for interpretation of groundwater flow and groundwater quality. Thus, a robust classification scheme for partitioning water chemistry samples into nearly homogeneous groups can be an important tool for successful characterization of hydrogeologic systems.

A variety of graphical techniques have been devised since the early 1920s in order to facilitate the classification of water with ultimate goal of dividing samples into similar homogeneous groups (each representing a hydrochemical facies). The piper diagram is one of the most widely used approaches in representing and comparing water quality. It conveniently reveals similarities and differences among various water samples.

In the present study area, hydrochemical data obtained from the Department of Water Affairs National Groundwater Archive was used to characterize the water types. The data was edited prior to further processing and these include calculation of ion charge balance between cations and anions. Comparison of measured and calculated TDS values was also used to assess the level of accuracy of the data. Hydrochemical data with charge balance >20% and TDS values that have significant contrast with calculated values (measured/calculated TDS ratios out of the range from 1 to 1.2) were excluded from subsequent graphical display of the results. The edited data was subdivided into 4 regions depending on geology, hydrogeology and aquifer types. Below interpretation of the hydrochemical data covering the 4 sub-regions is presented.

5.2 Kalahari sand cover

The hydrochemical data covering the Kalahari sand located in the northern part of the study area is shown on Figure 17. The groundwater in this area can be grouped into four main water types based on major anions and cations composition. These are: Ca-Mg-HCO_3, Ca-Mg-HCO_3-Cl, Ca-Cl-HCO_3 and Na+K-HCO_3-Cl. The dominance in Ca/Mg with strong HCO_3 content suggests an interaction between groundwater and aquifer material at shallow depth presumably carbonate rocks consisting soluble minerals such as Ca and Mg. The piper diagram shows variability in anions while the cations show nearly similar patterns (Figure 17). The clustering of the samples around the same point in the cation triangle suggests similarity in cationic composition.

The water type around Stella is Ca-Mg-$HCO_{3;}$ relatively enriched in Ca, Mg and HCO_3 ions typical of groundwater circulating at shallow depth with normal TDS content (<800 mg/l), and is suitable for domestic use. This water type is consistent with shallow groundwater strike varying from 10 m to 30 m depth around Stella town.

The dominance in Ca-Mg-HCO_3 water type within the Kalahari is contrary to the hydrochemical facies commonly observed in typical semi-arid-to arid environments, where the rate of evaporation is rapid which creates suitable condition for generating elevated salt content (Na+K-Cl water type). This is further evidenced by number of saltpans observed around Stella town showing high evaporation rate. Although the elevated Ca/Mg and HCO_3 content is unclear, the hole-to-hole section constructed from existing borehole logging located within the Kalahari sand shows that the overlying Kalahari sediments are thin (<15 m thick) and most of the water strike correspond to deeper rocks such as dolomite and limestone (Figure 18). The high concentration of HCO_3, Ca and Mg suggests interaction between groundwater and the Transvaal carbonate rocks at depth greater than the thickness of the Kalahari sedimentary layers.

Most of the hydrochemical samples in the Kalahari are characterized by $10\%<Na+K<40\%$, $20\%<Mg<45\%$, $35\%<Ca<95\%$ and $10\%<HCO_3<95\%$. In the anion triangle, the chloride content varies over wide range ($5\%<Cl<85\%$) suggesting the presence of different flow path or variability in source of groundwater.

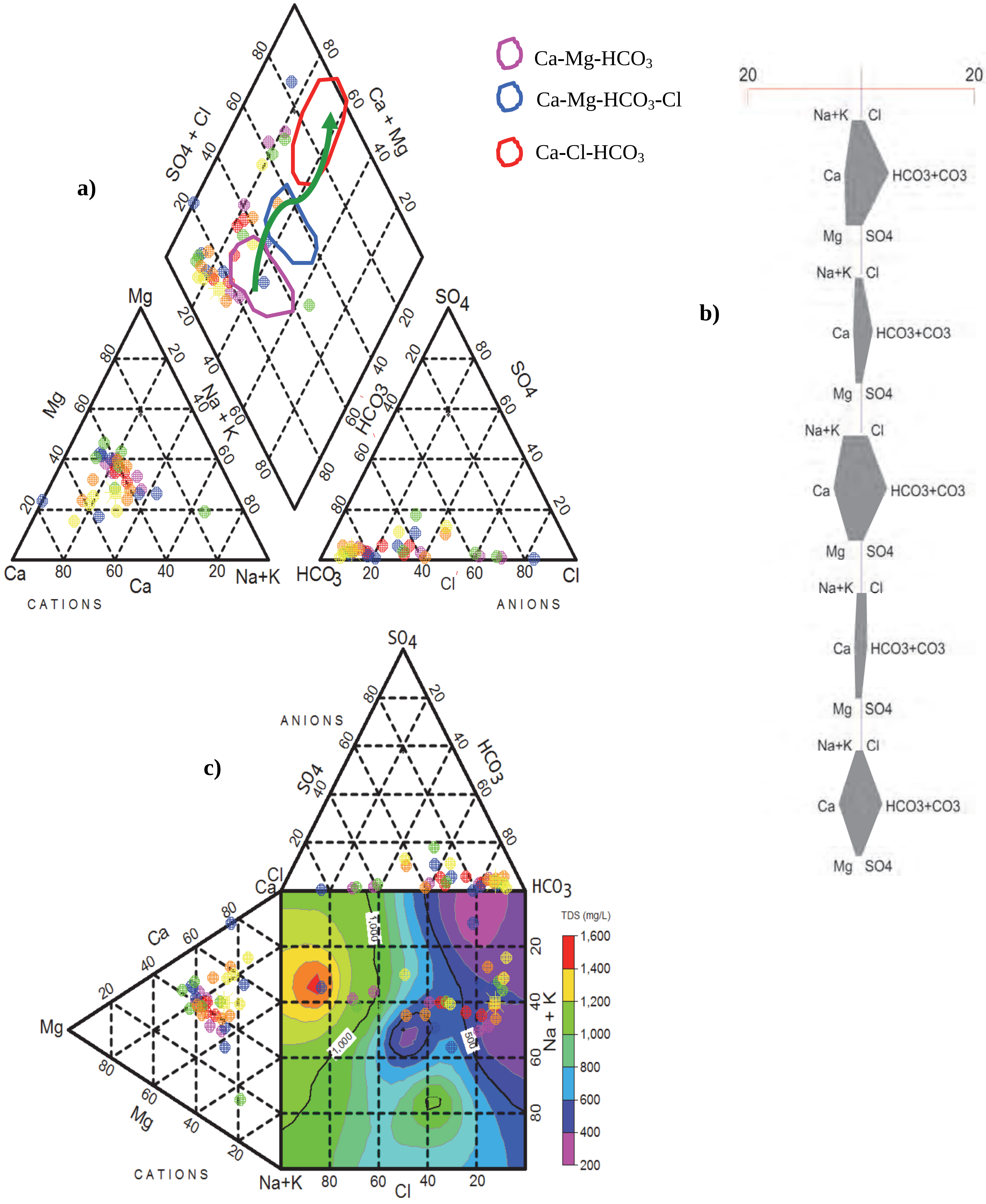


Figure 17: Graphical presentation of hydrochemical data covering Kalahari sand. **a)**, **b)** and **c)** are Piper, Stiff and Durov diagrams, respectively.

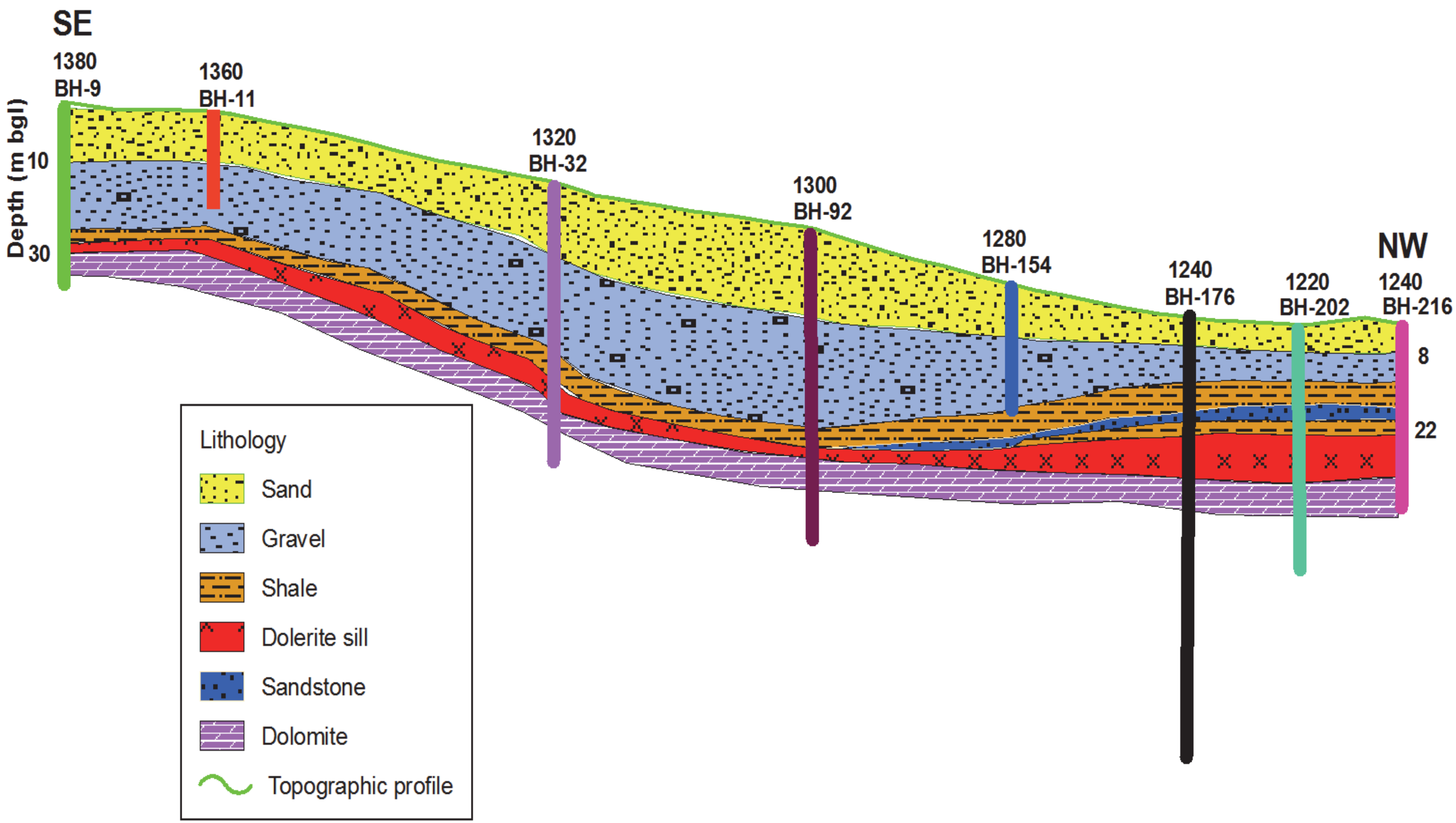


Figure 18: Hole-to-hole section drawn from 8 boreholes located within the Kalahari sedimentary rocks. The numbers written above the boreholes indicate topographic height a.m.s.l.

The high chloride content can also be attributed to excessive groundwater withdrawal around Stella, where pumping possibly altered the direction of local groundwater flow resulting in change in water-rock interaction leading to local highs in certain anion species such as chloride.

There are three boreholes that are characterized by relatively elevated chloride content (801.2, 458.3 and 331.3 mg/L) as shown on Figure 19. The high chloride content may be attributed to high evaporation which is consistent with enrichment in ^{18}O and deuterium isotopic composition of groundwater. The elevated chloride content in these boreholes could also be due to salt being accumulated on the surface during short rainfall events which was episodically re-dissolved and leached into the groundwater during heavy rainfall events.

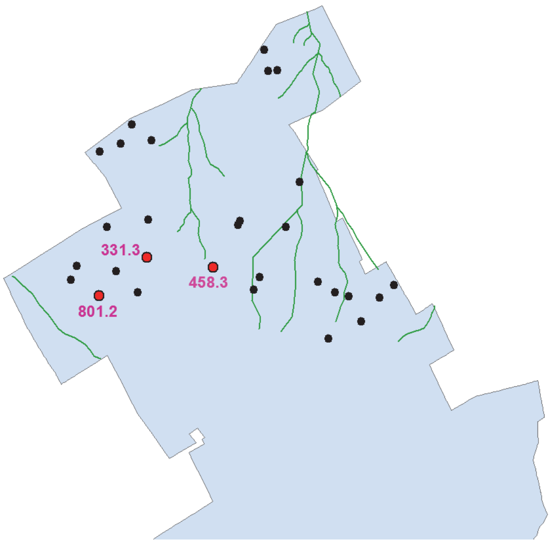


Figure 19: Location map of elevated chloride content in the Kalahari sand. The red circles show boreholes with very high chloride and TDS values.

Most of the boreholes around Stella are <40 m deep below the ground and the re-dissolved salts can easily reach the water level thereby elevated TDS values. This is consistent with evaporative enrichment in stable isotopic composition of groundwater. The range of TDS values within the Kalahari sand falls within normal fresh water quality (<700 mg/L) with the exception of few boreholes that are characterized by elevated chloride content (Figure 17c). The boreholes that are located in the south western part of the Kalahari sand are very high in TDS (>1000 mg/L) which fall under the category of brackish water type. Considering the guidelines for groundwater quality for domestic use, these boreholes can be abandoned until verified through analysis of samples.

5.3 Transvaal carbonate rocks

The southern part of the study area is underlain by dolomite and limestone of the Transvaal Supergroup. There are several boreholes supplying water for municipal and agricultural use and the groundwater chemistry is a factor of both anthropogenic and natural processes. The graphical presentation of the hydrochemical data covering this area is shown on Figure 20.

The classification of groundwater on the basis of major ions composition shows the presence of three main water types which are listed according to their order of significance: Mg-HCO_3-Ca, Mg-Ca-HCO_3 and Mg-Cl-Na+K water types. The variability in groundwater chemistry may be attributed to various factors. One of the main factors is natural chemical evolution of groundwater resulting from water-rock interaction. The second reason is most boreholes were drilled close to settlements (e.g. Vryburg town and surrounding villages), as a result of which anthropogenic sources may have contributed to change in the normal hydrochemical facies.

According to Bartolino and Cole (2003), Mg-Ca-HCO_3 water type shows early stage of water-rock interaction which often takes place at shallow depth where rock mineralogy do not significantly influence the composition of groundwater. This is consistent with the shallow depth of water strike (<50 m) we observed at wellfields located 10 to 15 km northwest and west of Vryburg town (e.g. Armoedsvlakte, Biesjiesvlakte and Swartfontein wellfields). The water type corresponding to these boreholes is Mg-Ca-HCO_3 with the exception of deep wells > 80 m bgl at Armoedsvlakte wellfield.

The Mg-Ca-HCO_3-Cl water type is intermediate or slightly mineralized relative to the first group. The water strike corresponding to this group ranges from 50 to 70 m below ground where dolomites and limestone are intercalated with quartzite, shale, diabase sills and

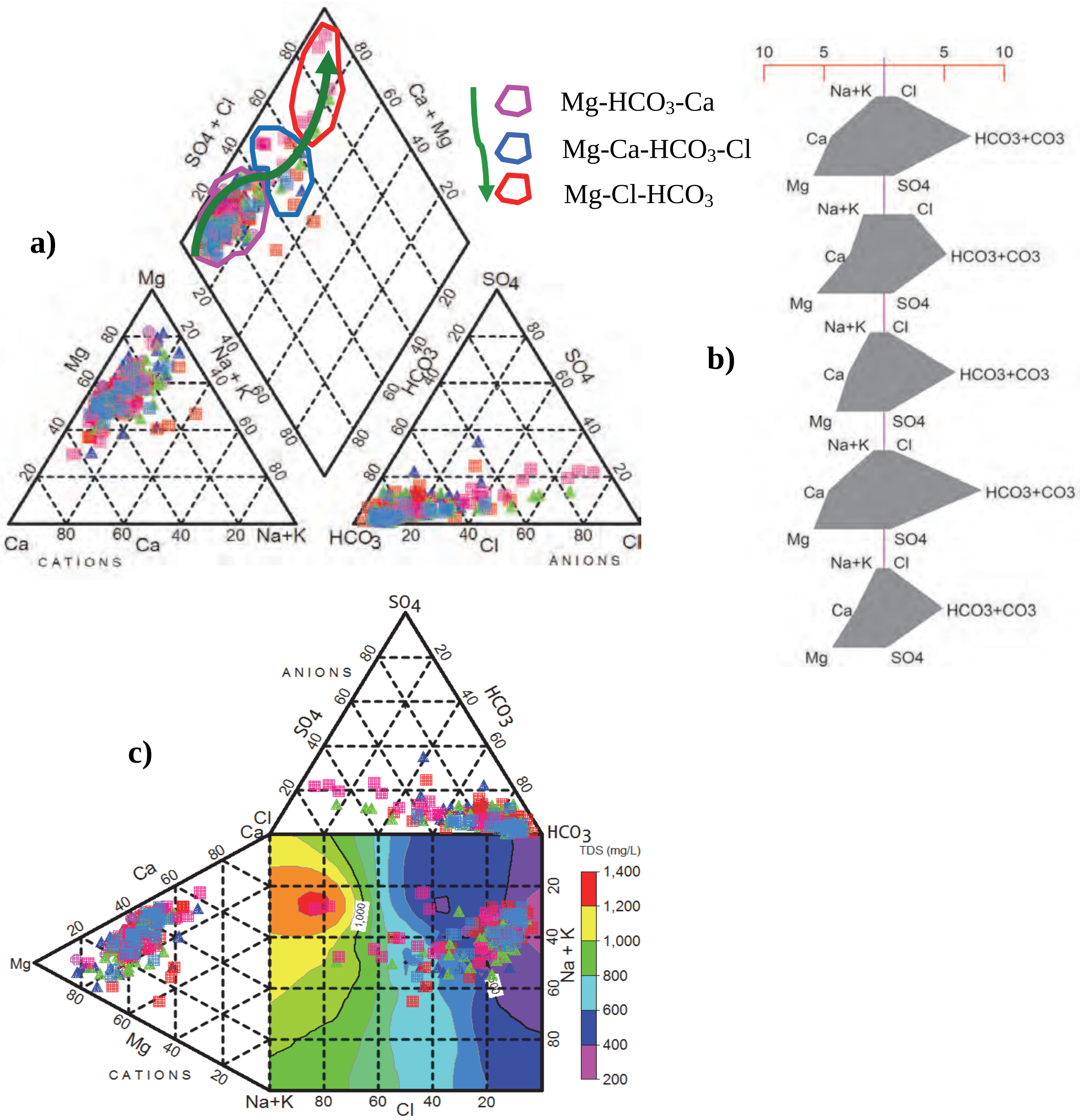


Figure 20: Graphical presentation of hydrochemical data covering carbonate rocks in the southern part of the area. **a)**, Piper diagram **b)** Stiff and **c)** Durov diagrams.

dolerite sheets (Figure 21). This suggests that the boreholes tapping water are not only from the carbonates, but fractured rocks that are rich in Ca and Mg. The Na and K content is insignificant compared to Ca and Mg in the cations triangle, which possibly suggests the silicate minerals (e.g. plagioclase feldspar:- albite or hornblende) are not as important dissolved species in groundwater as compared to carbonate minerals. This possibly suggests that water moving through different layers react with varying degrees with the surrounding minerals (and other constituents). The cations triangle shows this phenomenon, i.e. the silicate minerals do not react quite readily with the groundwater compared to the carbonate minerals such as Ca and dissolved CO_2. The low Na+K content in groundwater may be attributed to an ion exchange process or a weak chemical reaction between the silicate minerals and groundwater. The progressive trend from Mg-HCO_3-Ca to Mg-Cl via Mg-Ca-HCO_3-Cl suggests an evolving groundwater chemistry as it percolates from the shallower layers to deeper; exhibited by fresh and relatively juvenile at shallow depth to more evolved and mineralized matured groundwater at depth resulting in from chemical reaction along the groundwater flow path as shown by the dark-green arrow on Figure 20.

The average value of TDS is 600 mg/l (Figure 20c) with the exception of the two samples located 10 km east of Vryburg (along Vryburg-Stella road) that are characterized by elevated TDS (>1000 mg/l). In general, the groundwater quality in the southern part of the area is fresh with the exception elevated total hardness as discussed in Section 2.2.4.

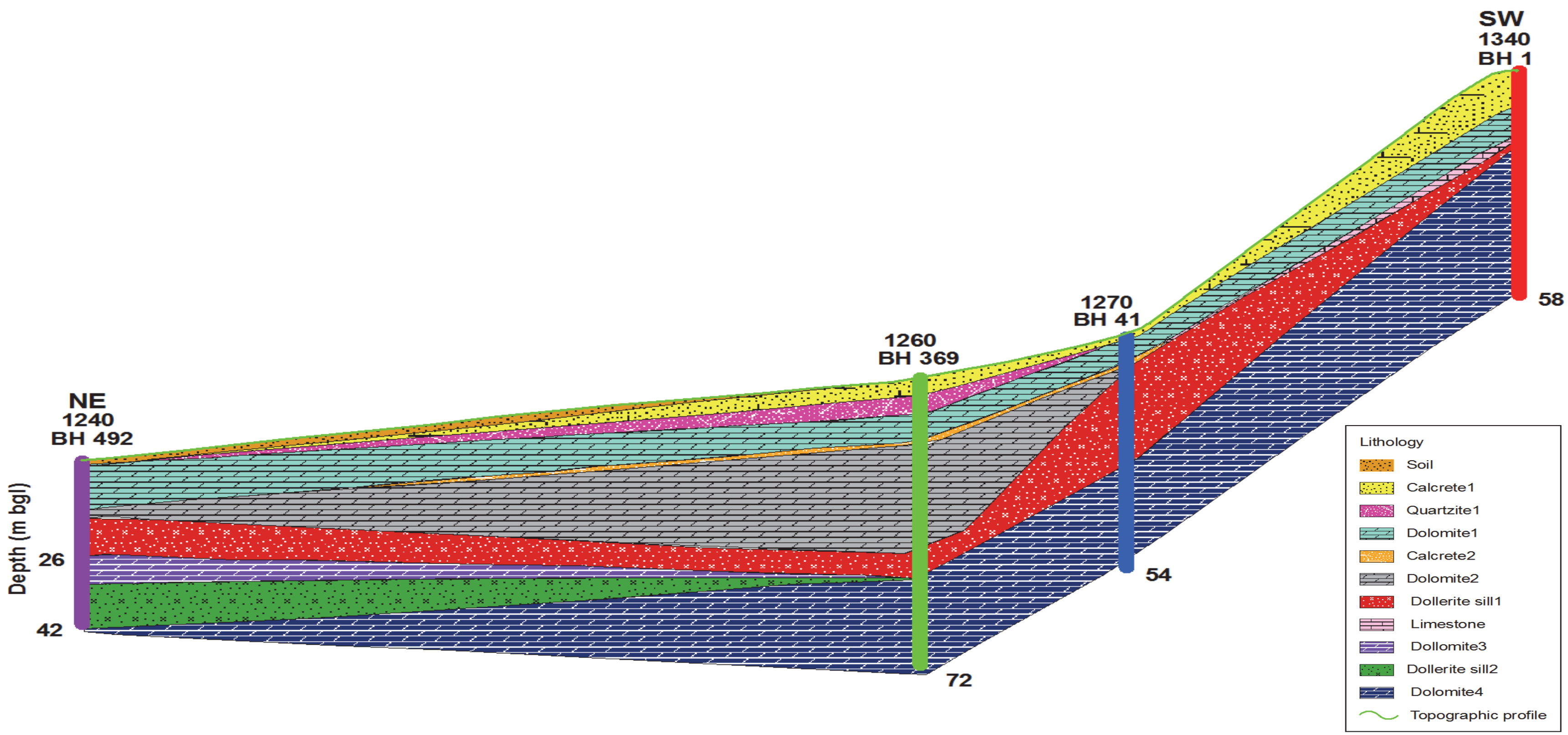


Figure 21: Hole-to-hole section showing the intercalation of dolerite sills and quartzite within the dolomite sequence.

5.4 Ventersdorp

The hydrochemical data covering the Ventersdorp volcanic and sedimentary rocks located in the central part of the study area is shown on Figure 22.The chemical characteristics of groundwater varies very widely compared to Kalahari and the southern part of the study area. The samples show four major types of water *viz.*: Mg-Ca-HCO_3, Mg-Na-HCO_3, Mg-Na+K-Cl and Mg-Cl-Na+K. The piper diagram shows scattering of the data indicating mixing of various types of water.

The chloride content varies over wide range probably owing to anthropogenic influence and extensive evaporation. The anthropogenic contribution arises from contamination through irrigation return flow and the use of fertilizers. The stiff map shown on Figure 22c also indicates elevated chloride content at several boreholes. The majority of the boreholes within the Ventersdorp are mainly found in relatively flat topography (*conf*. Figure 1), low-lying areas and used waste water in the surrounding upland can easily get into groundwater, which possibly elevates chloride content.

The TDS shown on Figure 22b indicates normal quality of water suitable for domestic use with the exception of elevated TDS (>2000 mg/l) located in the northern part of Vryburg and south of Stella towns, similar to high TDS values shown on Figure 8b.

During the present work, 18 water samples were collected from the wellfields that are located around Vryburg, Stella towns and Dithakwaneng village. The piper diagram showing the types of water is presented on Figure 23. The wellfields that are located close to Vryburg (Biesjiesvlakte, Armoedsvlakte and Swartfontein) show Ca-Mg-HCO_3 water type with relatively high alkalinity.

In summary, compilation of existing and new hydrochemical data collected during the present study do not show any major problem that may be harmful to human health. However, the boreholes recommended for domestic use need to be evaluated in order to avoid any potential health risk.

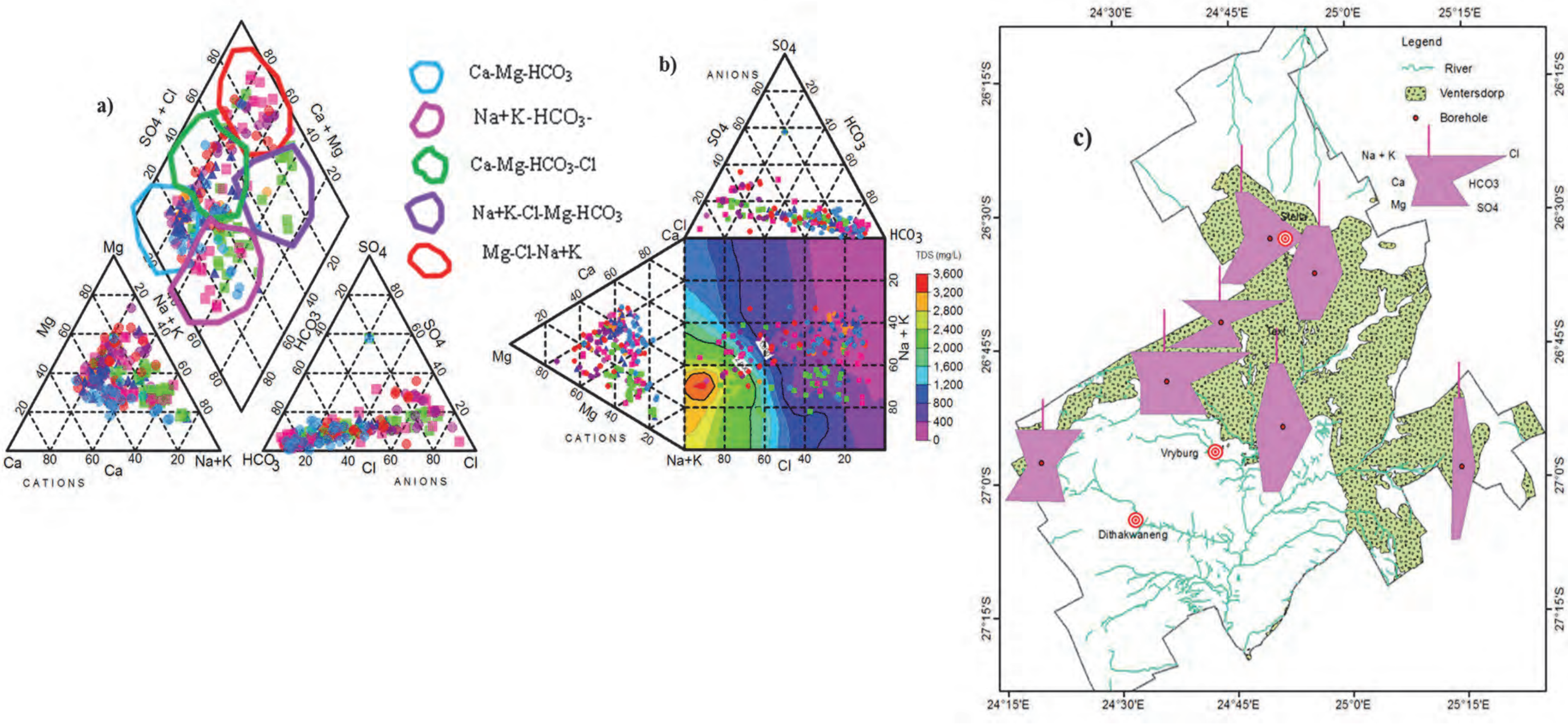


Figure 22: Graphical presentation of hydrochemical data covering the Ventersdorp volcano-sedimentary rocks. a) Piper diagram b) Durov diagram and c) Stiff map superimposed on Ventersdorp volcanic and sedimentary rocks.

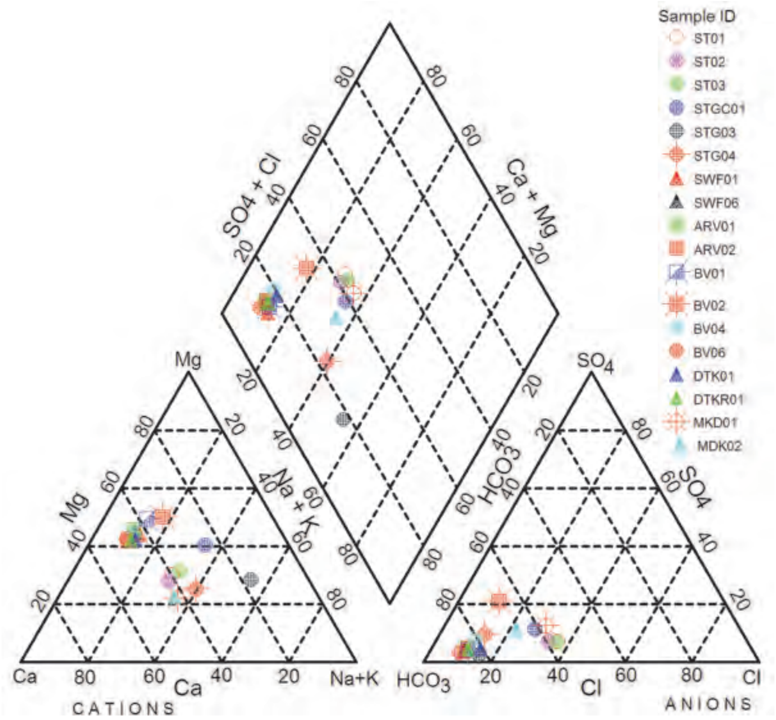


Figure 23: Piper diagram showing the types of water at the wellfields located around Vryburg, Stella towns and Dithakwaneng village.

6. Aquifer Test

Aquifer tests can give information about how much groundwater can be extracted from a well. It gives an initial indication about the success of a well and whether the well should be abandoned due to low yield or poor aquifer performance. Aquifer test helps to characterize the efficiency of the well and determine the aquifer properties, i.e. how the aquifer responds during recharge and discharge. It should be routinely done in order to prove the sustainable yield of a borehole before commissioning for water supply. Aquifer test is a classical approach and perhaps the only way to provide important information about in-situ aquifer hydraulic properties (hydraulic conductivity, transmissivity and storativity).

Despite many advantages of conducting aquifer tests on completed boreholes, it is not routinely tested in many well fields. If a borehole supplies any quantity of water, it is generally assumed productive. There are several reasons for this: lack of appropriate equipment, cost of pumping test and time, among others. However, the high failure rates and poor sustainability of rural-water supply such as the present study area makes the results of aquifer tests increasingly important. Knowing the performance of a borehole and its sustainability when it is drilled makes it easier to work out reasons for failure should it occur at a later stage.

Aquifer tests can be performed for 12, 24, 48 or 72 hrs depending on safe determination of aquifer parameters. An aquifer tests commonly begins with a “step test”, i.e. it involves pumping a borehole at different rates each for 60 minutes, until the maximum rate the borehole can deliver. The water level is constantly monitored and noted during each step. This gives an indication of the possible yield the borehole can sustain for a constant discharge test (CDT). The CDT consists of pumping a borehole at a specific rate for a duration of relatively longer hours, with a sudden switch off of the pump after the pump cycle, with a recovery test following immediately afterwards. The constant discharge curve is then analysed by curve fitting to give an indication of transmissivity and storativity values which characterize aquifer properties

In the present study, aquifer tests were carried out at 3 boreholes and the results were combined with previous aquifer tests conducted at the wellfields located around Vryburg, Dithakwaneng and Lima wellfields. Figure 24 shows the location of boreholes that were tested during the present work and the previous studies.

In the previous aquifer tests, prior to a constant discharge test, a step-drawdown test was carried out with steps varying from 3 to 5. Each step was 60 min in duration and was performed until drawdown began to stabilize before proceeding to the next step. This was done to examine well efficiency and non-linear behavior. The result was used to determine the optimum discharge rate required for the constant pumping test. At Dithakwaneng, the step-drawdown tests were followed by constant discharge tests which were carried out at two boreholes for 48 hrs and one borehole for 72 hrs (Motebe, 2004). The results show that sustainable yield at the three boreholes varies from 1.5 l/s to 3.0 l/s. A complete recovery (99.8%) was detected at the two boreholes pump-tested for 48 hrs, which suggests good well performance. Analysis of the recovery test independently confirmed the results derived from the interpretation of constant discharge test.

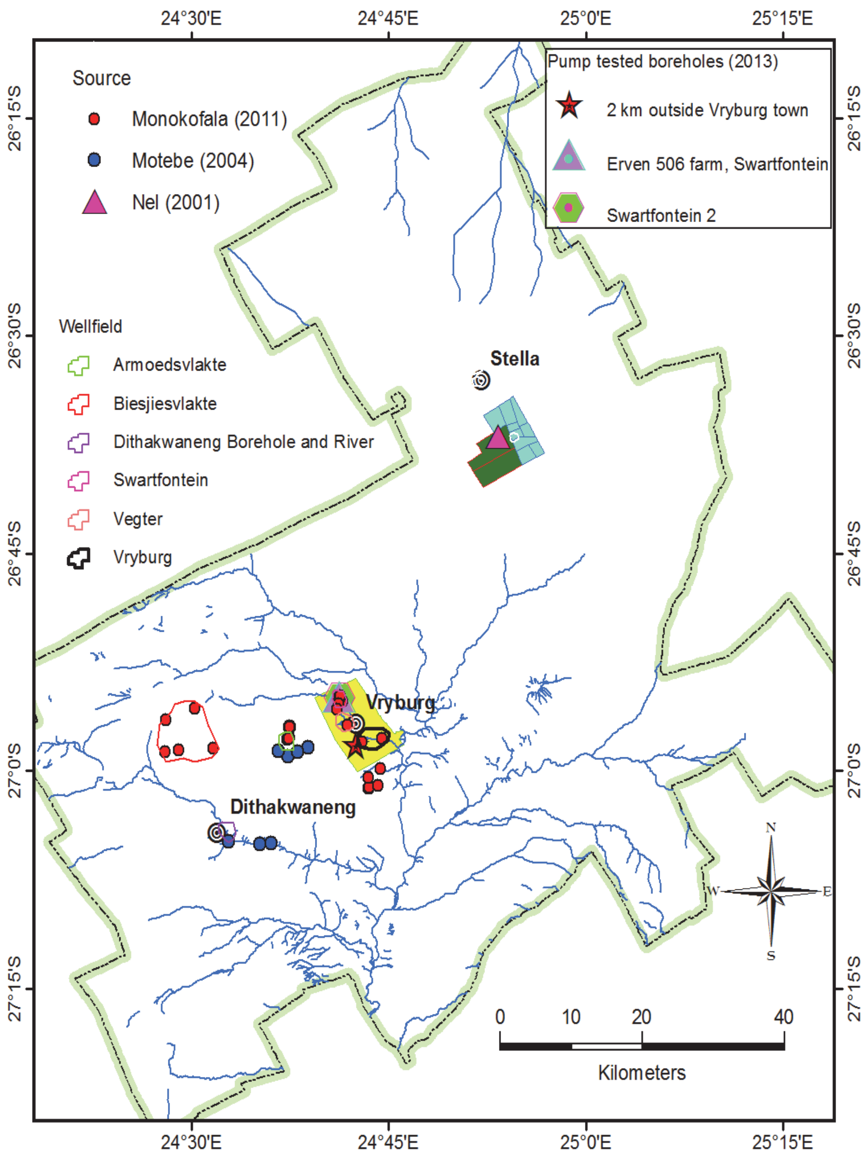


Figure 24: Location map of boreholes that were aquifer tested by the previous and the present studies.

The three boreholes drilled at Dithakwaneng village are relatively deep, varying in depth from 69 m to 88 m bgl. The groundwater quality in this area is characterized by elevated total alkalinity (TAL) with hardness varying from 327 mg/l to 421 mg/l (very hard water). The high alkalinity possibly suggests that most boreholes around Dithakwaneng village penetrate the Malmani dolomites Subgroup of the western Transvaal basin. The rainfall recharging groundwater chemically reacts with the rocks that contain Ca and Mg. In the presence of CO_2 gas, the chemical reaction gives rise to elevated HCO_3 and CO_3 ions. The pH level, temperature and the oxidation-reduction condition during water-rock interaction may also have facilitated enrichment in alkalinity, which consequently give rise to very hard water.

Aquifer Test at Borehole 1

In the present study, aquifer tests were carried out at three boreholes located at Swartfontein and Vryburg wellfields (Figures 24 and 26). The main objective of the aquifer test was to determine sustainable yield and aquifer parameters. Pumping test began with the borehole located just 2 km outside Vryburg town. The borehole is 130 m deep bgl with static water level at a depth of 64.96 m. The aquifer test started with a step drawdown test (SDT), and the first step lasted for 60 minutes with a discharge rate of 3.04 l/s. The second step with a discharge rate of 5.05 l/s lasted only for 10 minutes before the water reached the pump inlet. The drawdown corresponding to the step test was 29.03 m. Following the step test, a CDT was carried out for 24 hrs at a discharge rate of 3.5 l/s. The drawdown detected at the end of the CDT was 7.29 m. The drawdown versus time graph shows a slow and steady decline in drawdown during the first 7 minutes and this was followed by steeply increasing drawdown thereafter, suggesting the presence of no flow boundary at a depth of about 67 m bgl. The well performance diagnostic is shown on Figure 25 plotted on a logarithmic scale for the time axis.

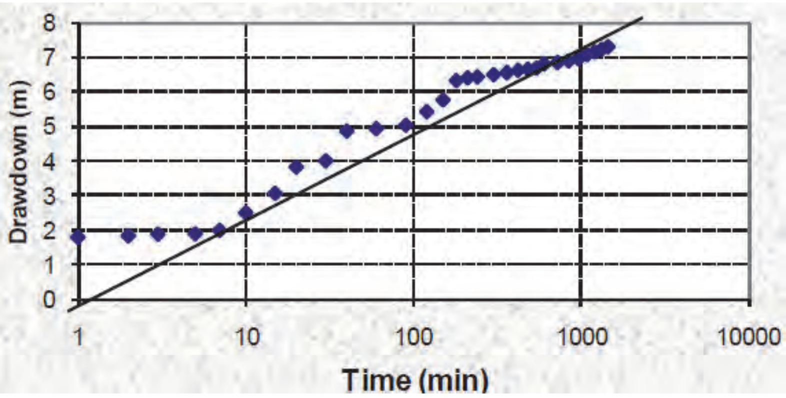


Figure 25: Drawdown versus time showing the performance of the borehole and aquifer.

The CDT was analyzed using FC and the results show transmissivity (T) values of 26 m^2/day for early time and 15.2 m^2/day for the late time using basic FC method. Using

Cooper-Jacob method, T value of 53.4 m^2/day was determined for the late time. The storativity (S) is 1.00E-03 using FC method and 1.43E-05 using Barker. The sustainable yield ranges from 2 l/s to 3.5 l/s using different methods as shown on Table 3. Although the CDT results show good fracture connectivity and radial flow regime, the transmissivity is very low, suggesting the availability of poor fracture connectivity at a distance away from the pumping point indicative of little horizontal flow of water towards the borehole.

The optimum recommendation for the borehole can be to equip with a motorized pumping system installed at a depth of 90 m bgl in order to supply water at a sustainable discharge rate of 1.5 l/s and 2.12 l/s for 24 hrs and 12 hrs, respectively. The dynamic water level should not exceed 68.1 m and 69.1 m bgl during 24 hrs and 12 hrs, respectively.

Aquifer Test at Borehole 2

The second borehole is located on Erven 506 farm at Swartfontein wellfield (Figures 24 and 26). It is located at about 5.2 km northwest of Vryburg town. The borehole is 83.5 m deep with a static water level at 50.47 m bgl. The borehole was pumped for three steps each 60 minutes at rates of 0.95, 1.84 and 2.77 l/s.

The water level draw down after each step test indicated 4.06, 9.43 and 22.95 m below the original static water level. At the end of the SDT, the water level did not reach pump inlet. The borehole was then subjected to a CDT for total duration of 1440 minutes or 24 hours at a rate of 2.44 l/s. Following the CDT session, the pump was switched off. The water level drawdown was 27.96 metres below the original static water level at the end of the CDT. It recovered to the original static water level in 25 minutes indicating fast recovery rate. The well performance is shown on Figure 27 plotted on a log scale for the time axis.

The drawdown versus time shows a steady and gentle decline in drawdown during the first 20 minutes and this was followed by a sudden rise just after ~20 minutes or at a depth of ~60 m bgl, indicating the response of fracture networks to pumping. This was followed by a steady and gentle decline in drawdown from 20 minutes to 450 minutes. Thereafter the drawdown stabilized during the remaining CDT session, indicating the cone of depression possibly intersected a recharge zone, where the amount of recharge is nearly equal to discharge.

The CDT was analysed using the basic FC, FC inflection point, Cooper-Jacob and Barker/Bangoy methods, to determine T and S values. The summary of the results are shown on Table 4. The average recommended abstraction rate (based on 24 hour duty cycle) corresponding to these methods was taken to calculate the recommended yield for 12 hours per day. The recommended abstraction rate can be 0.90 l/s and 1.27 l/s for 24 and 12 hours, respectively. The optimum amount of water that can be abstracted per month is 2332.2 m^3.

The performance of the borehole is moderate-to-good, however the transmissivity is very low, possibly due to lack of fracture connectivity at a distance away from the borehole which is similar to the borehole located just outside Vryburg town.

Summary of Borehole Abstraction Recommendation

- Routine monitoring of water level (monthly), abstraction volumes (monthly), rainfall figures (daily) and water quality (bi-annual) is strongly recommended and should strictly be adhered to. This data will form the basis for which any changes in the groundwater regime can be recognised.

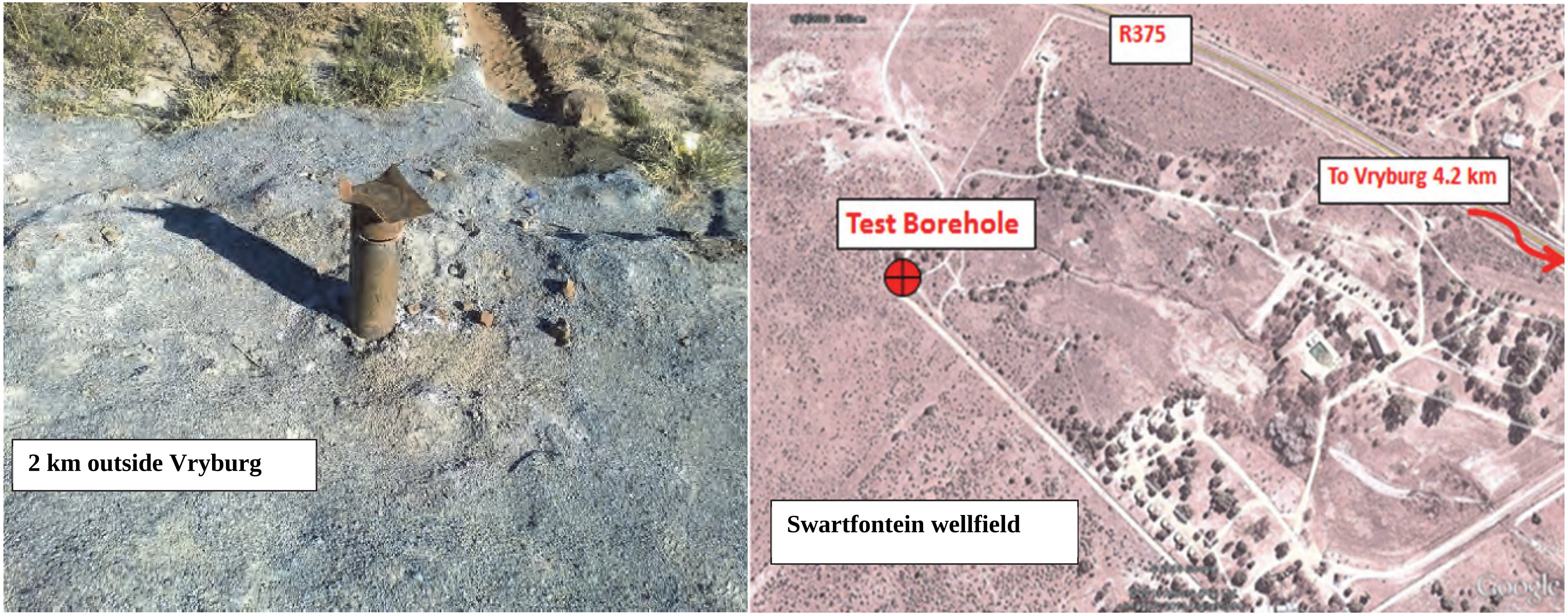


Figure 26: Location map of the boreholes aquifer-tested at Swartfontein wellfield and 2 km outside Vryburg town.

Table 3: Summary of results of aquifer test (2 km outside Vryburg).

Method	Sustainable yield (l/s)	St. Dev	Early T (m^2/day)	Late T (m^2/day)	S
Basic FC	0.75		26	15.2	1.10E-03
Advanced FC			26	15.2	1.10E-03
FC inflection point	0.61				
Cooper-Jacob	1.51			53.4	4.96E-09
FC-Non linear	0.69		15		1.10E-05
Barker	1		K_r = 108		1.43E-05
Average Q (sus l/s)	0.91		b = 0.11		

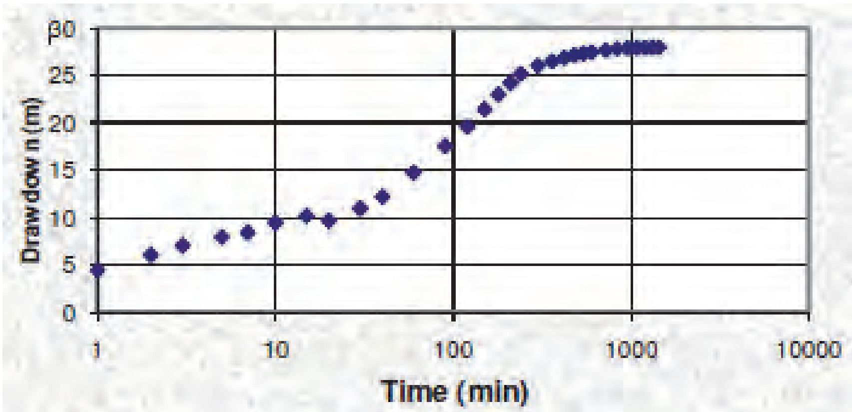


Figure 27: Drawdown versus time showing performance of borehole and the aquifer.

- Hydrogeological monitoring data (described above) should be evaluated bi-annually by a qualified hydrogeologist.

- A Groundwater Management Plan, with relevant Groundwater Monitoring and Reporting Protocol, should be established and calibrated annually.

- Groundwater level monitoring in all boreholes on the farm on a monthly basis is strongly advised, once water abstraction for the development start. The information gathered during such a monitoring programme can be used to manage future abstraction figures. The implementation of well managed groundwater monitoring system can therefore not be over emphasized.

Table 4: Borehole abstraction recommendation

Recommendation				Dynamic water level (mbcl)	
Sustainable yield l/s for 12 hrs /d	Volume m^3/d	Sustainable yield l/s for 24 hrs/d	Volume m^3/d		
1.3	56.2	0.9	77.8	58	

Aquifer Test at Borehole 3

The third borehole is located at Swartfontein Recreation Centre situated at about 5 km northwest of Vryburg town, along R378 road (Figure 24). The static water level is at a depth of 38.59 m bgl. The borehole was pumped for three steps each 48 minutes at a rate of 1.81 l/s, 3.57 l/s and 5.71 l/s. Water level drawdown was constantly monitored throughout the SDT and a total drawdown of 33.21 m was measured below the original static water level. Following the SDT, the borehole achieved 99.5% recovery after 14 minutes. This was followed by a CDT at a rate of 4.3 l/s for a period of 1010 minutes. A drawdown of 15.77 m was achieved during the CDT with a recovery of 97.2% after 11 minutes.

Table 5: Summary of aquifer test results and recommendations (Swartfontein wellfield).

Method	Sustainable yield (l/s)	St. Dev	Early T (m^2/day)	Late T (m^2/day)	S
Basic FC	0.45	0.3	5	2.1	1.82E-03
Advanced FC			5	2.1	1.00E-03
FC inflection point	1.41	0.41			
Cooper-Jacob	1.16	0.75		21.9	2.91E-18
FC-Non linear					
Barker	0.6	0.5	K_f = 1425		S_s = 1.57E-04
Average Q (sus l/s)	0.91	0.45	b= 0.04	Fractal dimen. = 1.8	

The drawdown versus time graph plotted on a log scale for the time axis shows the effect of wellbore storage at the early stage of pumping (Figure 28). This is followed by a recovery in water level, indicating that the water supplied by the matrix exceeded the water being pumped out of the borehole. This is thought to be as a result of the presence of fracture connectivity. After ~90 minutes of pumping, the drawdown reached a depth of about 14.13 m below the static water level, then it is observed that the borehole was slowly being dewatered (Figure 28).

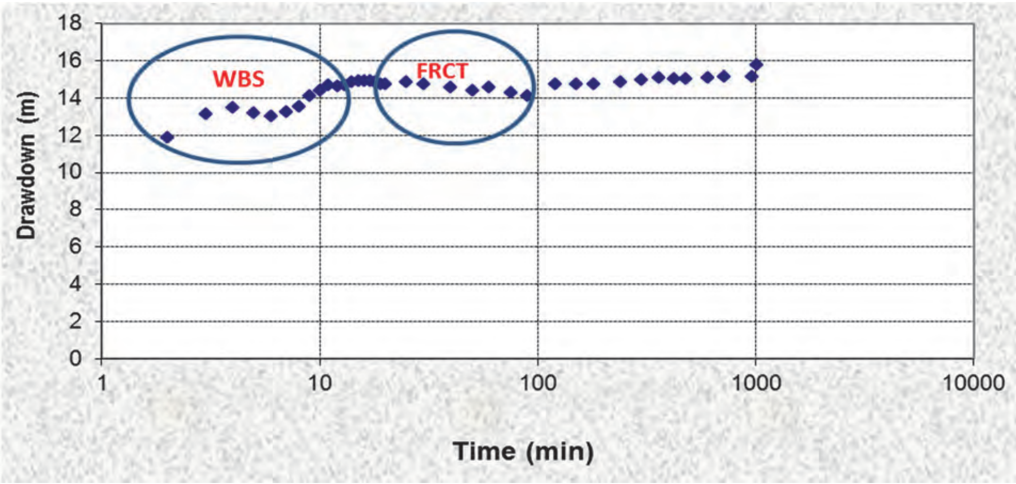


Figure 28: Drawdown versus time for the borehole located at Swartfontein recreation centre.

The CDT was analysed using the basic FC, FC inflection point, Cooper-Jacob and Barker/Bangoy methods, to determine sustainable yield and aquifer parameters T and S values. The summary of the results are shown on Table 5. The FC method provided T values of 24.1 m^2/day for early time and 11.3 m^2/day for the late time. Using Cooper-Jacob method, T value of 25.3 m^2/day was determined for the late time. The storativity (S) value of 1.70E-03 was obtained using FC method (Table 5). Based on analysis and interpretation of the yield test using the FC programme, the recommended abstraction rate is 1.31 l/s over a production period of 24 hours. The recommended abstraction rate was attained by averaging the sustainable yields obtained using basic FC, FC Inflection Point, Copper-Jacob and Barker/Bangoy methods. The result shows that the borehole can supply a total volume of ~3395.52 m^3 per month.

Table 6: Summary of aquifer test results and recommendations (Swartfontein recreation centre).

Method	Sustainable yield (l/s)	St. Dev	Early T (m^2/day)	Late T (m^2/day)	S
Basic FC	1.03	0.58	30	11.3	1.70E-O3
Advanced FC			30	11.3	1.00E-03
FC inflection point	1.48	0.71			
Cooper-Jacob	1.73	1.12		53.4	5.20E-11
FC-Non linear	0.04	0.04	18.0		4.00E-05
Barker	1.01	0.56	K_f = 100	S_s =	1.00E-07
Average Q (sus l/s)	1.31	0.35	b= 0.20	Fractal dimen n =	2.0

Borehole Abstraction Recommendation

The maximum abstraction rate recommended for BH Swartfontein 2 is 1.31 l/s for 24 hours, with allowable drawdown of less than 14 m below the static water level. However it is strongly emphasised that boreholes in close proximity also ought to be taken into consideration during production. Monitoring of water level on monthly basis is therefore very important.

The geology of the site comprised of the Malmani Subgroup dolomite of the Transvaal basin. The interaction between groundwater and the carbonate rocks may produce sinkholes which results sudden fluctuation of water level. It is therefore essential to monitor groundwater level periodically in order to monitor water-rock interaction.

7. Groundwater Potential Mapping

7.1 Data and Approach

Conventional approaches for groundwater exploration using geological, hydrogeological and geophysical field-mapping involve increased cost, are time consuming and not practical for large region such as the present study area. The use of existing data (e.g. geology, airborne geophysics, topography, etc.) covering large area has become increasingly useful for groundwater potential mapping (Teeuw, 1995). In addition, for effective groundwater exploration, it is important to analyze the different parameters in an integrated approach. The integration of multiple dataset with various indications of groundwater availability can decrease the uncertainty and lead to successful results (Sander, 1996).

The area of study is largely barren with little vegetation cover which makes the use of remote satellite images ideal for mapping lineaments, morphological features, landforms, folded structures and lithological boundaries that are essential for characterization of groundwater potential (Drury, 1987). The airborne magnetic data and Landsat 7 ETM image shown in Figures 3a, 3b and 3c, respectively were used to extract lineaments. SPOT5 imagery with spatial resolution of 2.5 m by 2.5 m was used to capture additional lineaments which were not possible to retrieve from magnetic and Landsat 7 ETM images. Additional data such as land cover/land use, soil type, lithology and lineaments were combined in a GIS environment in order to assess the groundwater potential of the study area.

7.2 Preparation of Thematic Layers

A groundwater potential map of the area was produced based on integration of six thematic layers. Each layer was processed in terms of a set of physical properties that characterize groundwater flow and storage. The layers include lithology, lineament, drainage length density, geomorphology, land use and soil type. Below, is a summary of multivariate statistical approaches used for integration of thematic layers that allow mapping groundwater potential zones.

7.2.1 Geological Formations

The bedrock geology determines the regional hydrogeologic setting upon which water table configuration and groundwater flow systems (local, intermediate or regional) are manifested (Pascal et al., 2013). According to these authors, the bedrock geology provides the conduit system for water movement, which controls the amount, rates, patterns and distribution of groundwater.

As described in Section 2, hydrogeologically, the study area can be broadly subdivided into 3 subgroups of rocks, *viz.* 1) crystalline basement rocks that are older than 550 million years and cover ~40% of the study area, 2) Transvaal carbonate rocks and 3) The Kalahari sedimentary rocks.

The crystalline basement rocks comprise granite, metamorphosed sedimentary and lava flows that are generally described as having low porosity and permeability. However, over the course of geological time, various tectonic forces and release of confining pressure caused the rocks to break along horizontal and vertical sets of fractures, which can then serve as water-bearing openings. The classification of these rocks into six categories is therefore based on the availability of high fracture length density and lineament intersection, i.e. the availability of tectonically controlled interconnected fracture fabric is an essential parameter.

Carbonate rocks (limestone and dolomite) are soluble in weak acidic conditions such as rainwater that percolates through the soil. As the rocks dissolve along existing fractures or bedding planes, these openings enlarge due to continuous dissolution of $CaCo_3$ to form groundwater flow-paths, sinkholes, and caverns which act as suitable environments for groundwater storage. The carbonate rocks cover ~25% of the study area and the classification of these rocks is based on proximity to surface water, density of cross-cutting structures, chert content, rainfall intensity and topographic relief which collectively determine the development of karstic groundwater aquifer.

The Kalahari sedimentary rocks consist of poorly sorted unconsolidated-to-semi consolidated aeolian sand and of fluvial origin that cover ~25% of the study area

(Figure 2). The overlying Kalahari Group comprises unconsolidated clay-dominated sand, but as depth increases the sand grades into sandy-clay, duricrust, silcrete and calcrete. Where more consolidated, the sandstone contains either a siliceous or calcareous matrix which makes the rock to have low permeability. The occurrence of groundwater within the Kalahari Group therefore depends on proximity to river channel (e.g. southern and northern parts of the area), low cementing material (siliceous or clay) and rainfall intensity. Based on lithologic and hydrogeologic attributes, the geology of the study area is classified into six categories das shown in Figure 29.

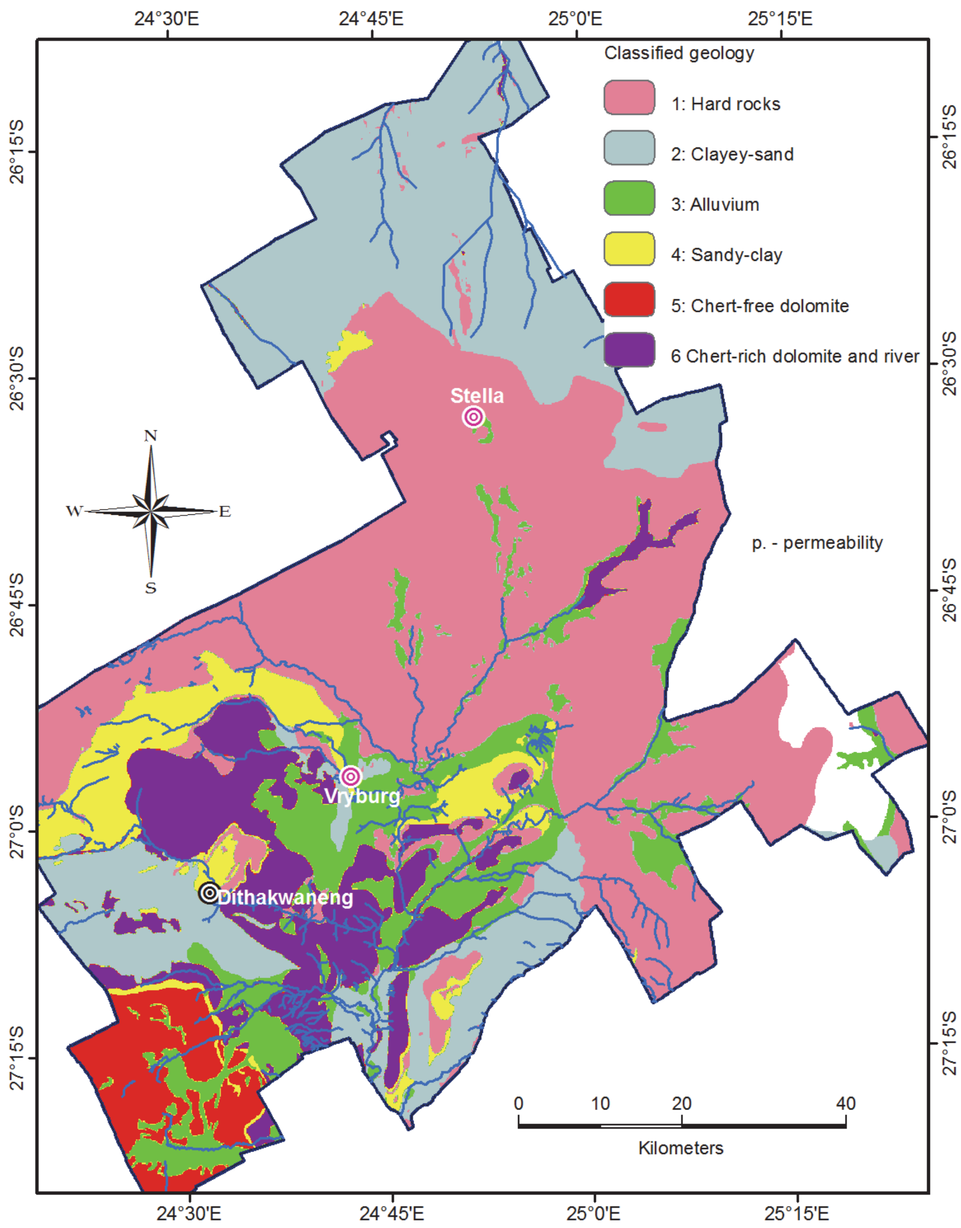


Figure 29: Reclassified geological map of the area.

7.2.2 Lineaments

Lineaments are surface manifestations of structurally controlled features, such as joints, straight course of streams and vegetation alignment. Lineaments are linear or curvilinear features and are identifiable on aeromagnetic map or Landsat imagery by their long, narrow and approximately straight alignments visible as tonal differences with respect to other terrain surfaces. Lineaments being weak zones, usually serve as conduits for movement or accumulation of groundwater in the subsurface; therefore, lineaments analysis of an area when extracted from the airborne magnetic and remotely sensed data give important information about the subsurface features that may control the movement and/or storage of groundwater.

Lineaments include tectonic structures and geomorphologic signatures such as topographic breaks. They have been used extensively for many applications: petroleum and mineral exploration, nuclear energy facility sitings, and water resource investigations. Previous studies have revealed a close relationship between lineaments and groundwater flow and yield (Lattman and Parizek, 1964; Mabee et al., 1994; Magowe and Carr, 1999; Fernandes and Rudolph, 2001). Generally lineaments are underlain by zones of localized weathering and increased permeability and porosity.

In this study, lineaments were extracted from a mosaic of Landsat 7 ETM satellite images using edge enhancement and filtering techniques. This was done manually using an interactive system based on pattern recognition. In addition, lineaments were derived from high resolution airborne magnetic data, SPOT5 imagery, geology map and digital elevation model.

Lineament length density and lineament intersection frequency were computed using Eq. 1 as illustrated in Figure 30. The lineament intersection frequency was computed using ArcGIS Spatial Analyst based on constructing grid cells over an area of interest. The cell size was chosen based on the average distance between adjacent lineaments. Based on this approach, a cell size of 1 km by 1 km was used to produce lineament length density and lineament intersection frequency map of the study area.

$$Linedensity = \sum_{i=1}^{N} L_i / \pi r^2$$

(1)

$$Linedensity = \sum_{i=1}^{N} L_i / \pi r^2$$

where L_i and Pi stand for line and intersection point, respectively.

The resulting computed value corresponds to the center of a grid cell, and this was the basis for generating discrete raster data for lineament intersection points (Figure 31). Similar approach was used for producing lineament length density. The length of lineaments can be an important parameter in controlling regional and local groundwater flow. The lineaments that extend over tens and hundreds of kilometers may act as groundwater divide similar to surface water divide which can be controlled by topographic relief or rock layering. On the other hand, lineaments that extend over short distance control local groundwater flow (Vries and Simmers, 2002).

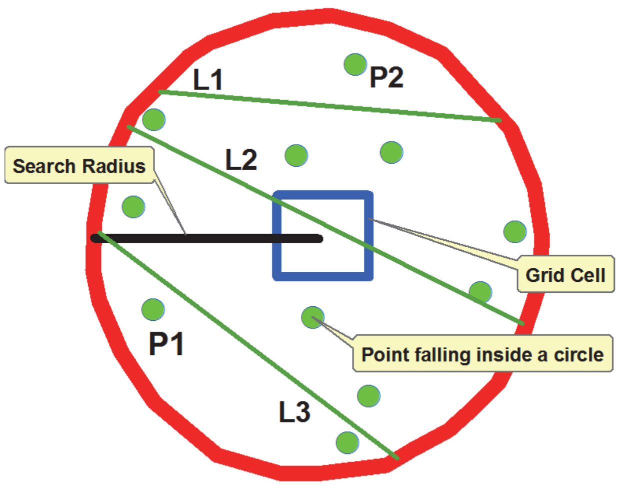
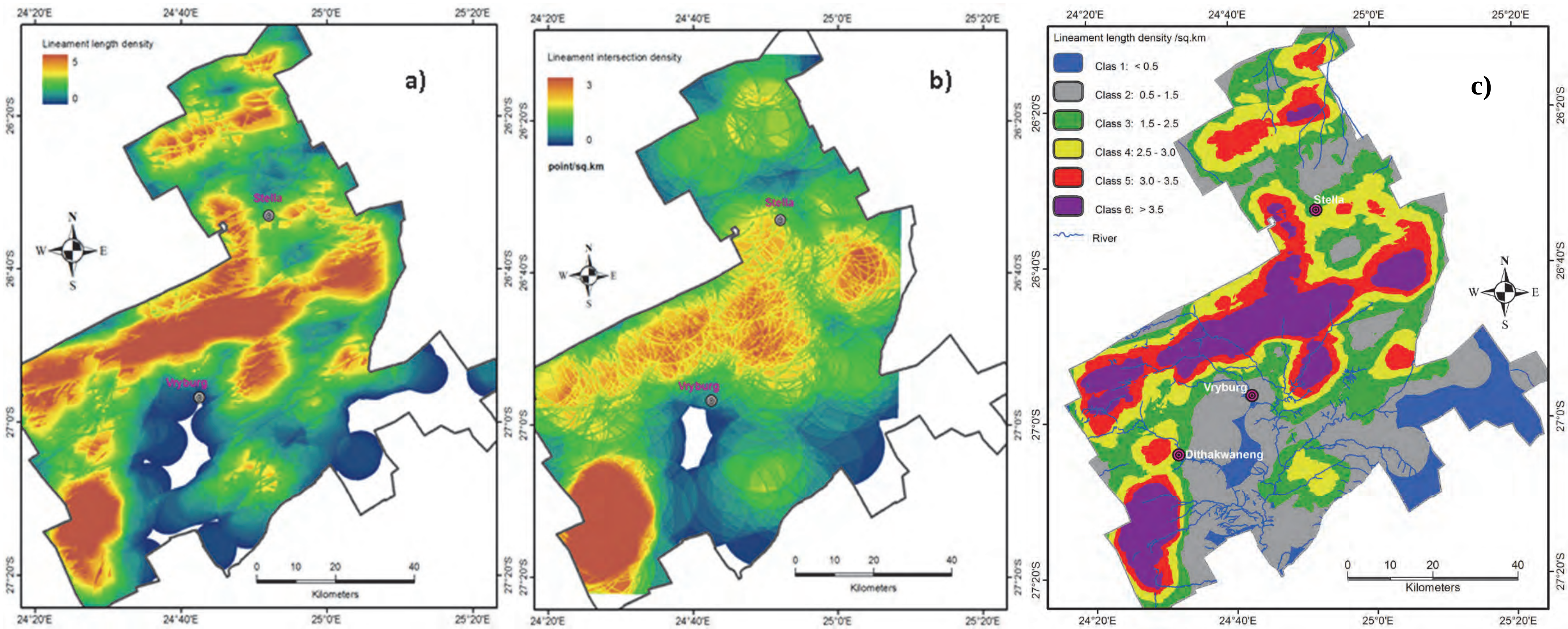


Figure 30: Simplified illustration of method of determination of lineament length density and lineament intersection frequency. L1, L2 and L3 stand for lineament length, whereas P1, P2 and P3 are lineament intersection points that fall inside a circle.



Figures 31: **a, b** and **c:** Lineament length density; lineament intersection density and classified lineament length density.

7.2.3 Land cover and land use

Land cover (LC) refers to the biophysical state of the earth's surface and upper subsurface in terms of broad categories such as cropland, natural or man-made forest, grassland, settlement, etc. (Schulze, 2003). However, LC can be altered to land use (LU) by human activities, primarily for purposes of agricultural production, settlement or other activities. According to Schulze (2003), both LU and LC can be described as a major controlling factors of watershed hydrology that partition rainwater into infiltrated water and runoff. For example, in an increasing order of runoff observed on the following LC/LU: forest, grassland, farmland, barren land and urban settlement (Anbazhagan et al., 2005).

Agricultural practices such as ploughing changes the soil's textural property and enhance infiltration. For example, Bromley et al. (2002) noted that infiltration increases beneath cultivated fields due to increase in soil permeability and moisture content. On the other hand, grazing if poorly managed (i.e. degradation through overstocking), results denudation and soil surface exposure as well as soil compaction which leads to overland flow (Schulze, 2003). In addition, a change in land use results in fluctuation of recharge thereby temporal and spatial variation of groundwater storage potential. The impact of LU/LC change on groundwater occurrence is most obvious in many hydrogeological environments and the impact also varies with the type of land cover and land use. Also, Leduc et al. (2001) noted that as a result of land use change in south Niger, an increase in recharge from the previous 5 mm/year to the present 20 mm/year has induced a long-term water table rise. This suggests that clearing vegetation changes the hydraulic properties of the topsoil allowing more infiltration and rise in groundwater level.

The land use data used in the present study was obtained from the Council for Geoscience. The data shows temporarily cultivated commercial and subsistence dry-land farms, human settlement, forest plantation, semi-grassland, bare soil, woodland, shrubs and degraded bush-land. Vegetation cover in the study area is dominated by small-leaved and finely branched shrubs and reeds. The density of vegetation varies greatly with rainfall, topography, lithology and soil type. The central part of the study area is dry-land with scattered trees and dry grassland.

Many details have to be considered when classifying LU and LC in terms of groundwater development. However, any attempt to understand all details of the impacts of each LU entity on groundwater development can only be realized through a simplified approach. In the present study, LU was classified based on borehole yield and soil type, considering aquifer performance is partly controlled by LU and LC. Based on this approach, the LC data was classified into six groups with respective class showing variability in groundwater potential (Figure 32). In this classification, the LU and LC located south and west of Vryburg are highly favorable for groundwater development, while the northern part of the area is low.

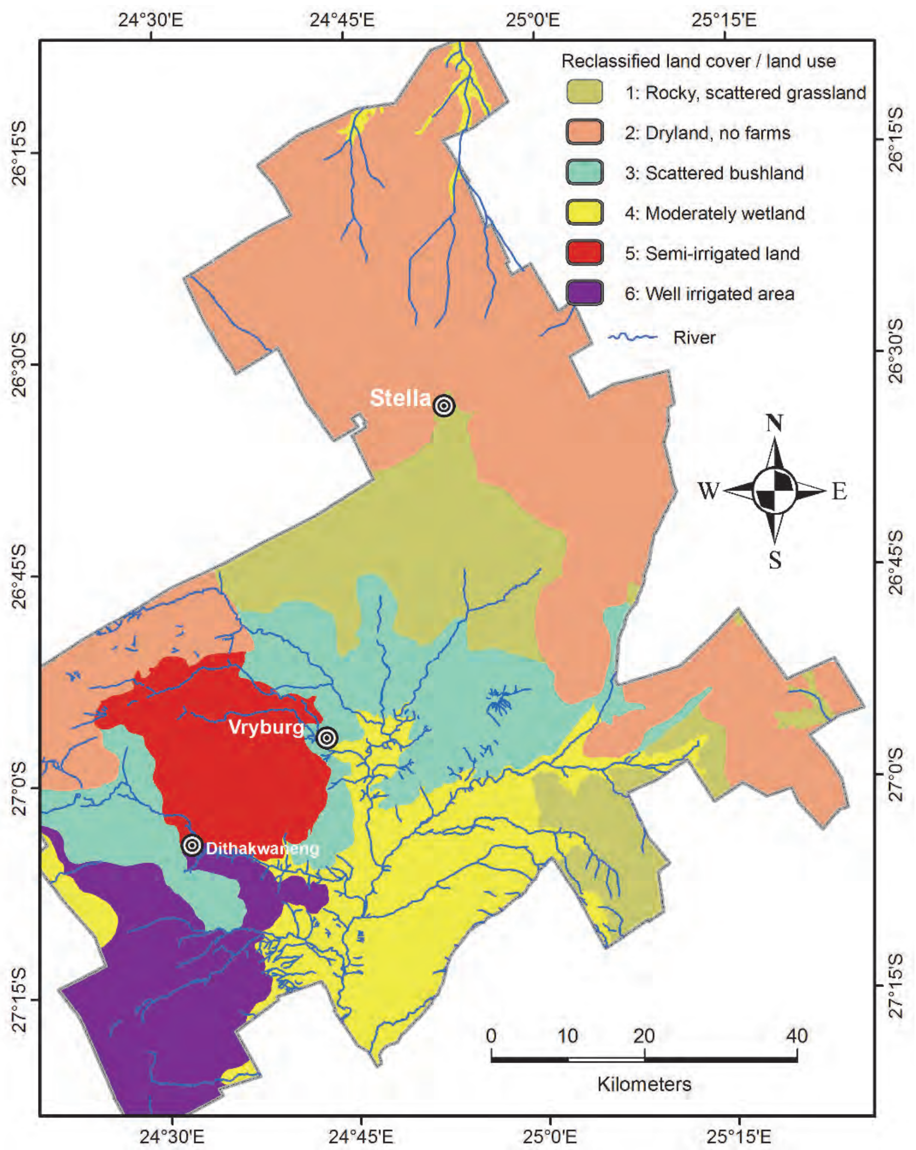


Figure 32: Reclassified land cover/land use map of the study area.

7.2.4 Drainage

Drainage pattern is one of the most important indicators of runoff and infiltration, because drainage texture, pattern and density are mainly controlled by the underlying lithology, structures and topography (Sener et al., 2005). According to Sener et al. (2005), drainage pattern is a reflection of the rate at which precipitation infiltrates compared to surface runoff. The infiltration-runoff relationship can be influenced by the permeability, which in turn a function of rock type and the availability of cross-cutting structures (Edet et al., 1998). Drainage density shows rock permeability and infiltration capacity, i.e. where rocks are highly permeable, infiltration increases thereby enhancing the development of groundwater potential. On the other hand, in areas where rocks are characterized by low permeability, insignificant infiltration takes place, and most of the precipitation generates surface runoff.

The drainage map of the study area was derived from ASTER global digital elevation model (DEM), orthorectified Landsat 7 ETM and SPOT5 images. Several digital filtering and colour contrasting techniques were applied to the satellite images to allow visual tracing of drainage lines. In the central and southern parts of the study area, drainage patterns are dendritic and sub-parallel to the regional strike direction of lineaments (*conf*. Figures 3d and 33) suggesting that the density and patterns of drainage lines appear to be controlled by the regional structural framework. The drainage map of the study area was converted into discrete data, i.e. raster format using similar approach discussed in Section 9.2. This was followed by identification of classification into six types based on drainage patterns and density. The resulting classified layer is shown on Figure 33.

7.2.5 Topography

Topography plays a major role in controlling local and regional groundwater flow-path as demonstrated by Pascal et al. (2013). To understand about groundwater occurrence, it is necessary to perceive the effect that Pascal (2013) called “the hydrogeologic environment” for groundwater occurrence, i.e. the effect of topography, geology and climate on groundwater flow and storage. In order to understand groundwater development, the three factors should be taken into consideration. Topography determines the amount of energy available to partition rainwater into infiltration and runoff (Pascal et al., 2013). According to this author, water runs off steeper slopes more quickly with less infiltration into the ground. Where topography decreases, the rainwater received per unit area will have sufficient residence time to infiltrate and recharge groundwater.

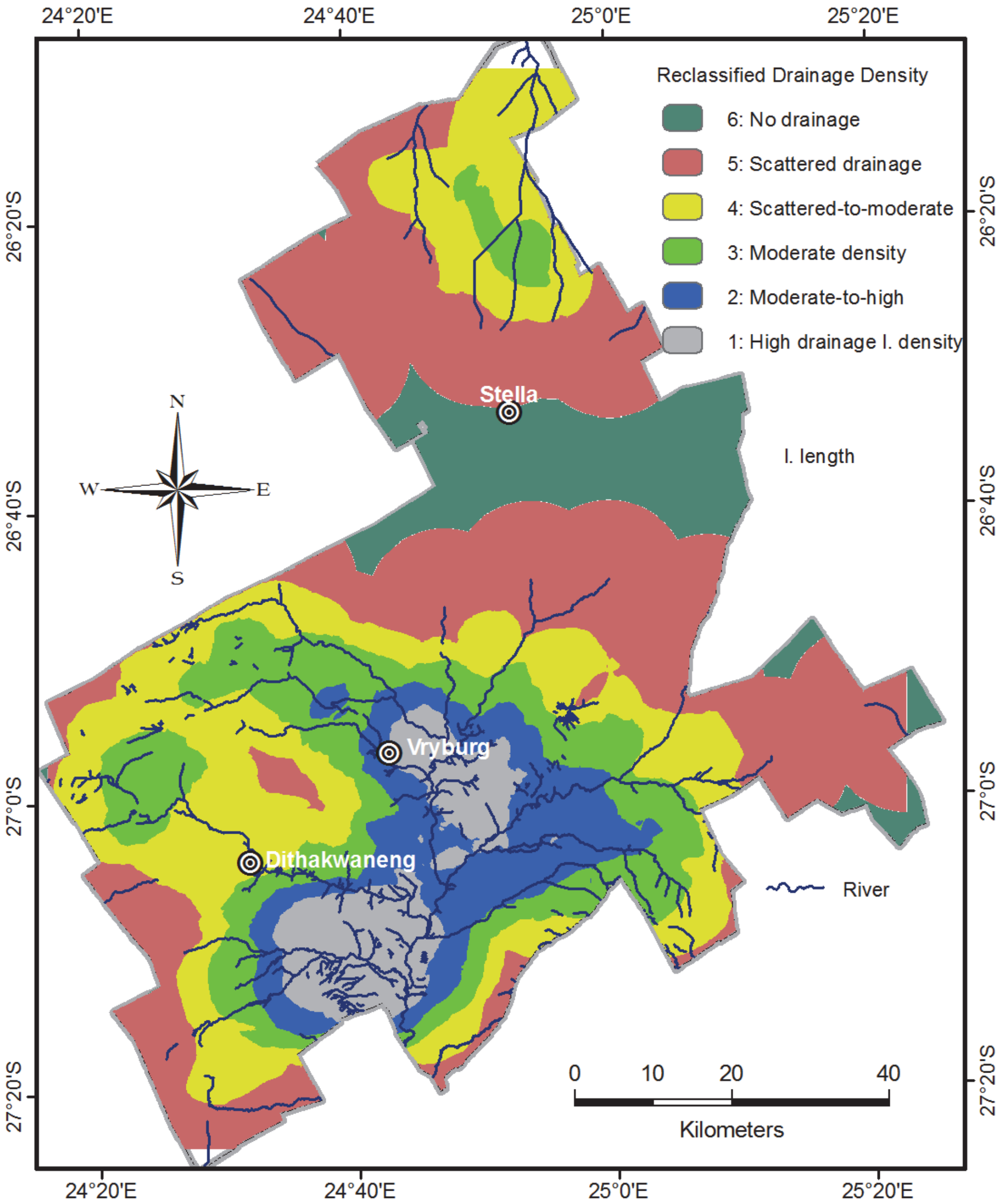


Figure 33: Rasterized and reclassified drainage density map of the study area.

In addition, flat topographic areas minimize denudation of weathering products (saprolite) where the rate of weathering process is higher than erosion which is an important condition for the development of groundwater in crystalline basement terrain such as the central part of the present study area. The occurrence of groundwater therefore depends on a number of interrelated factors such as topography, structure, rock type and soil permeability among others.

The topographic data used for the present study was retrieved from ASTER global DEM with a spatial resolution of 30 m by 30 m. The northeastern and southwestern parts of the area have relatively high topography varying in altitude from 1100 m to 1400 m a.m.s.l., while the central and southern parts are flat with an increased density of drainage lines. Physiographically, the area can be subdivided into number of sub-regions depending on topographic relief. Areas that are very close or within river channel can be described as good groundwater potential, whereas elevated regions with steep slope can be classified as poor groundwater potential. The gridded topographic data was classified into six categories based on variation of topographic relief, borehole yield and rock type. Each class was assigned a value that characterizes groundwater potential zones (Figure 34).

7.2.6 Soil

The soil properties are: texture, particle size, packing, cracking, sorting, crashing, clay and silt content, ratio of sand content and mineralogical composition. Soil composed of certain minerals (e.g. biotite), exhibits a swelling tendency, which results in local increase in volume that eventually favors cracking and fissuring thereby enhancing the soil permeability. In addition, the textural properties of soil invariably control the infiltration and recharge of groundwater. According to Schulze (2003), the soil surface roughness or coarse texture enhances infiltration capacity allowing groundwater recharge. The spatial variability of soil texture largely depends on the nature of the underlying rocks, physiography, climate, land use/land cover and weathering processes. In arid and semi-arid regions, similar to the present study area, soil formation and properties are the outcome of physical weathering and erosion that are acting on the underlying rocks. Soil also influences the type of land cover and land use which in turn determines groundwater occurrence. This suggests that soil, land use and land cover are interlinked in controlling the distribution of groundwater potential zones.

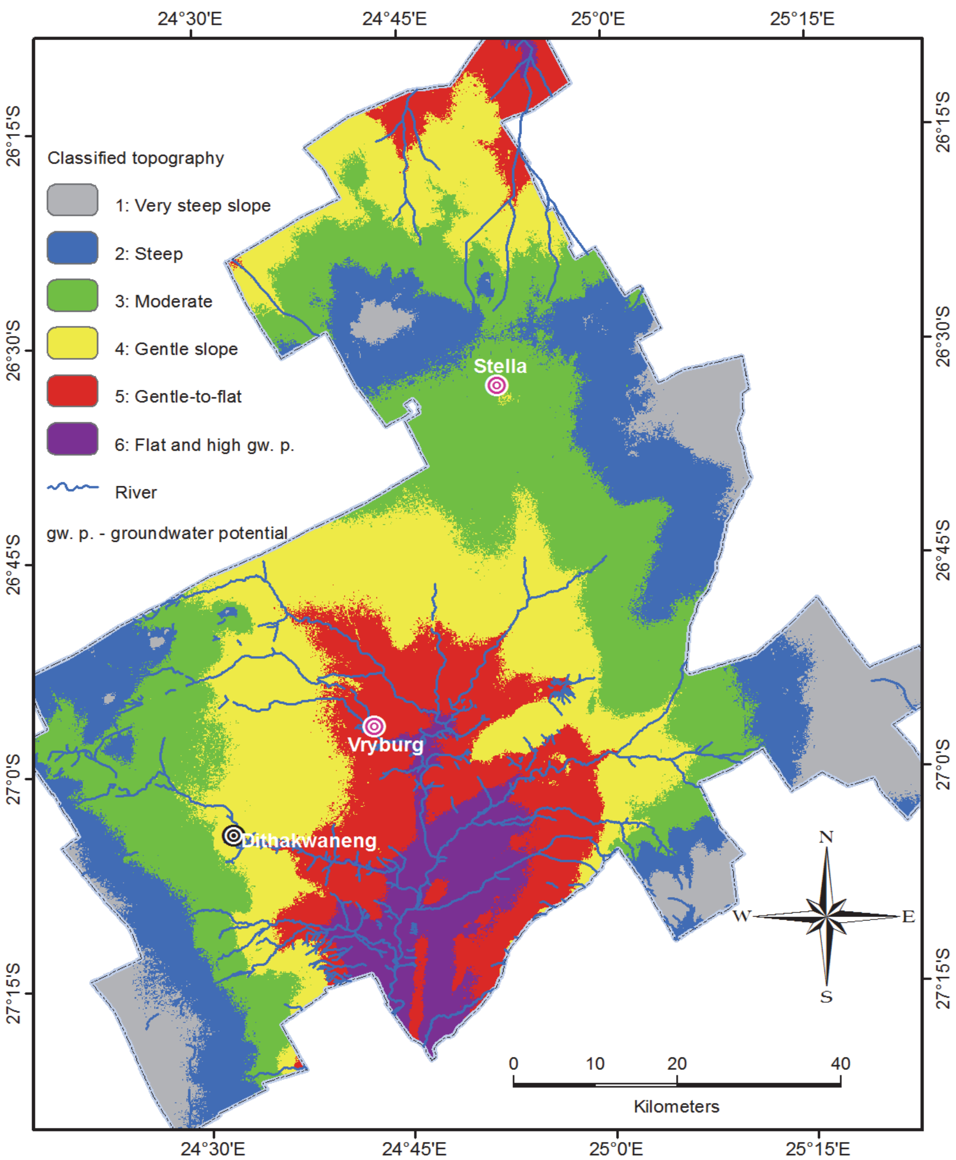


Figure 34: Reclassified topographic map of the study area.

Soil texture invariably controls the penetration of surface water into an aquifer. In the present study, soil data obtained from the Agricultural Research Council was combined with other data to characterize the groundwater potential of the area. The data was classified into six using the United States Department of Agriculture texture-based soil classification approach. The classified soil types are: clayey (~24%), silty-clay (~40%), silty-loam (~0.5%), sandy-loam (~0.1%), loam-sand (~26%) and sandy-soil (~10%) as shown in Figure 35. The northern and central parts of the study area consist of silty-loam to silty soil with increased clay content, whereas the southern part comprises of sandy-to sandy-loam soil. The soil type within the Kalahari Group in the northern part of the area is characterized by high infiltration capacity and high water retention properties due to the clay content, particularly during the summer wet season. These properties contributed to high evapotranspiration resulting in less groundwater recharge. In contrast, in the southern part of the area, soil type is strongly mixed, but largely sandy-soil that allows groundwater recharge, notably along the banks of the southward draining rivers (Figure 35).

In areas where the soil is friable and has increased sand content, the soil type was classified as permeable favoring infiltration and high groundwater recharge zone. Based on all soil attributes that characterize soil properties, a value ranging from one to six was assigned to each class to show variability in soil types in terms of texture and composition that defines the permeability (Figure 35). The classified thematic layer was used for further multivariate statistical analysis to map the groundwater potential of the area.

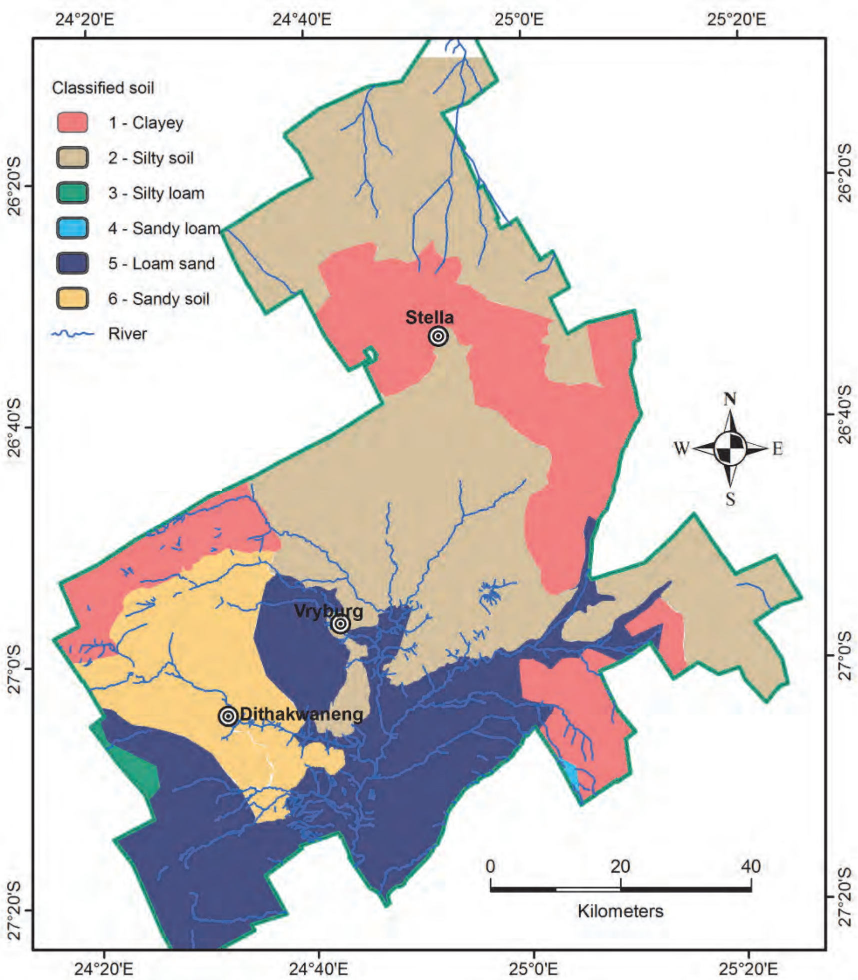


Figure 35: Reclassified soil map of the study area.

7.3 Integration of Thematic Layers

7.3.1 The use of PCA for Groundwater Potential Mapping

Six thematic layers, *viz.* geology, drainage, soil, lineament, topography and land use / land cover were used as input layers to multi-data integration algorithms. Prior to data integration, each layer was converted into raster format. This was followed by classification of each raster format thematic layers into 6 classes. Each class was scored depending on the relative importance in controlling groundwater distribution. Following this approach, the individual layers were assigned weighting coefficient which was calculated based on the following procedure.

Principal Component Analysis (PCA) is a covariance analysis between parameters. Covariance is a statistical approach that is used to determine the relationship between sets of data. So, when six layers are treated as having a certain degree of influence on groundwater occurrence, the covariance can be used to determine the relationship among the parameters. The PCA simplifies the complexity of computation of covariance for multi-datasets (e.g. the six layers that form 6 by 6 matrix). The PCA is a means of quantitatively analyzing the variation of data from the mean with respect to another dataset and allows reduction of the use of multi-dataset into reasonable number of layers without loss of information.

The PCA allows determination of eigenvectors and eigenvalues that are derived from the covariance matrix. Eigenvectors can be thought of as a measure of “preferential directions” of the dataset, or in other words, the main patterns (principal component) in the data. Whereas eigenvalues can be described as a quantitative assessment of how much a component represents the data. The higher the eigenvalues of a component, the more representative it is of the entire dataset, i.e. the first principal component contains the highest variance from all parameters used for calculation of PCA. As the order of the principal component (PC) increases, the information content decreases. The eigenvalues therefore measure the level of explained variance or indicate the percentage of the total variance. Eigenvectors on the other hand measure the weight or loading of a certain parameter in each component. An increase in an eigenvector in a certain component describes the degree of its influence, i.e. the parameter determines the trend of overall data distribution.

In the present study, PCA was applied to 6 input layers (Figure 36) that are hypothesized to be the main controlling factors of groundwater occurrence in the study area. Using PCA, eigenvectors for each principal component were computed (Table 7). The weighting coefficient used for integration of the 6 parameters was derived from the ratio of the sum of eigenvector in each component to the total sum of eigenvectors contributed from the 6 thematic layers. Table 7 presents the eigenvectors corresponding to each principal component that was used for calculation of the weighting coefficient. Table 8 shows the correlation matrix between pair of layers used for data integration.

Table 7: Eigenvectors for the 6 layers

Input layers	**Principal Components**					
	1	**2**	**3**	**4**	**5**	**6**
Slope	-0.00604	0.02983	0.52747	-0.07970	-0.84265	0.06655
Drainage	-0.01970	0.07195	0.27552	0.95213	0.09001	0.06222
Soil	-0.01152	0.04519	0.61027	-0.26192	0.45517	0.59128
Lithology	0.99970	0.00333	0.01903	0.01437	0.00375	0.00307
Lineament	0.00103	-0.99553	0.07816	0.05102	0.00969	0.01058
Land use/cover	-0.00670	0.02828	0.51668	-0.12523	0.27304	-0.80123

Table 8: Correlation matrix for the 6 layers

Input layers	**Correlation matrix**					
	Lithology 1	**Drainage 2**	**Soil 3**	**Topography 4**	**Lineament 5**	**Land use 6**
Lithology	1.00000	0.38160	0.61074	-0.20402	-0.01416	0.61744
Drainage	0.38160	1.00000	0.47143	-0.62883	-0.10366	0.43337
Soil	0.61074	0.47143	1.00000	-0.35350	-0.03465	0.77848
Topography	-0.20402	-0.62883	-0.35350	1.00000	0.01812	-0.24635
Lineament	-0.01416	-0.10366	-0.03465	0.01812	1.00000	-0.01284
Land use/cover	0.617744	0.43337	0.77848	-0.24635	-0.01284	1.00000

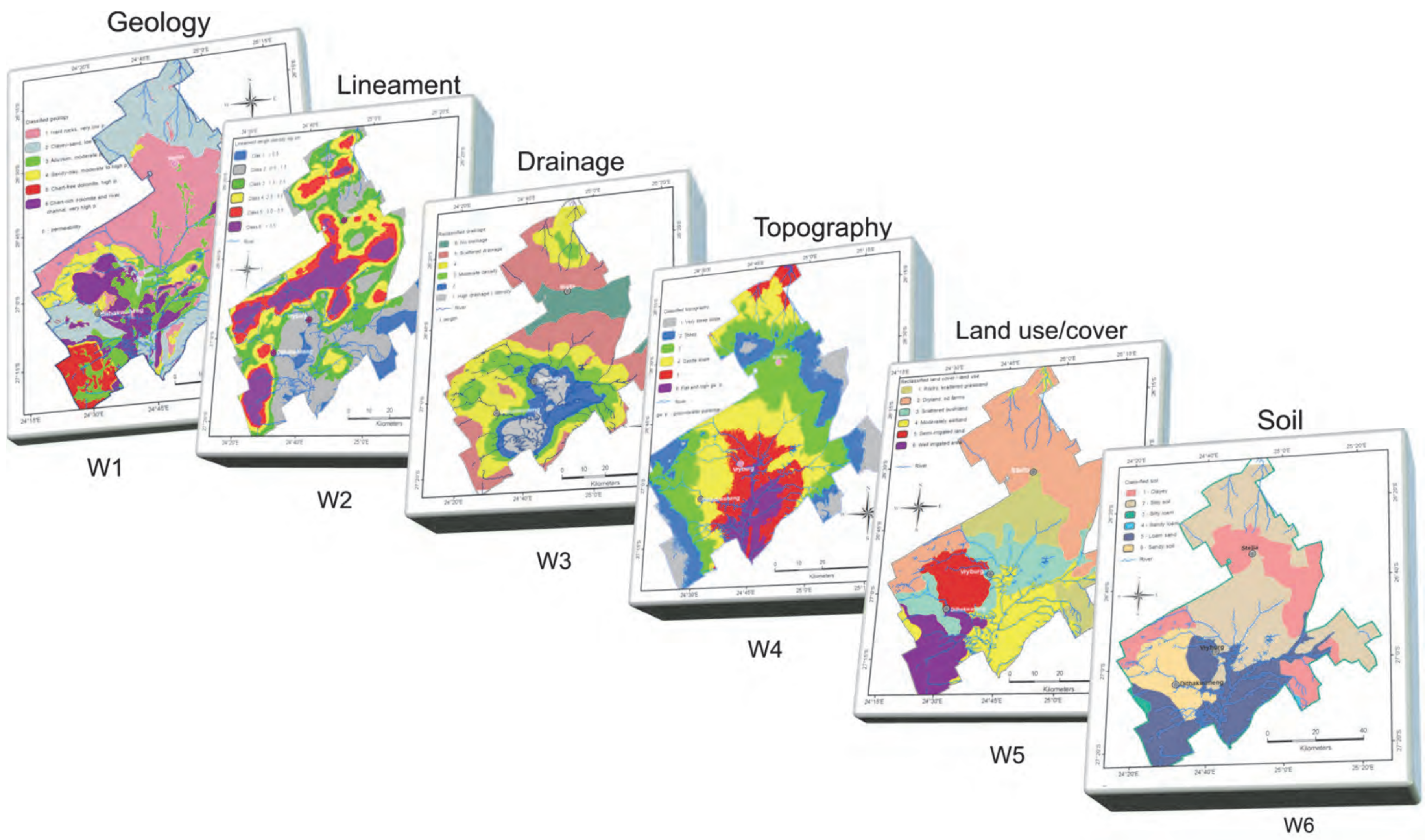


Figure 36: The 6 thematic layers and the weighting coefficients (w1, w2,…w6) used for modelling groundwater potential zones.

7.3.2 Groundwater Potential Mapping using PCA

Prior to integration of the 6 thematic layers, each class was scaled on an evaluation scale according to their importance to other classes in the layer (Table 7). The resulting raster obtained through data integration using PCA represents the model of groundwater potential of the area where high cell values correspond to potential target.

The higher the eigenvector in certain component, the more the parameter dominates the principal component. For example, in principal component 1 the maximum eigenvector is 0.99970 that corresponds to lithology (Table 7). This suggests that most of the data variance in principal component 1 is due to lithological loading. The maximum eigenvector corresponding to principal components 2 and 3 is -0.99553 and 0.61027, respectively which suggest that lineament length density and soil have maximum loading in principal components 2 and 3, respectively. The ratio of the sum of eigenvectors corresponding to each parameter to the grand total gives the weighting coefficient which describes the relative influence of each parameter on groundwater potential.

7.3.2.1 Determination of Weighting Coefficients

The weighting coefficient for each layer was calculated based on the following equation:

$$WC\ \% = \frac{\sum_{i}^{6} ev_i}{\sum_{j}^{6} \sum_{i}^{6} ev_i} * 100$$

where **WC%** is weighting coefficient, $\boldsymbol{ev_i}$ is eigenvector for a specific layer **i** that varies from 1 to 6.

Using the above equation, the weighting coefficient calculated for each layer is presented below:

Lithology: 26%

Drainage: 15%

Soil: 13%

Topography: 12%

Lineaments: 22%

Land use / land cover: 12%

The parameter with maximum weighting value and score suggests the most important factor in controlling groundwater potential in the area. Figure 37 shows the groundwater potential zones derived from the integration of the 6 layers. The zone indicated as '*very good'* and '*good'* potential cover 14% and 28%, respectively. They fall within the Malmani Subgroup that consists of dolomite and limestone. The proximity to river channels and low topographic slope partly contributed to high borehole yield within these zones, though the presence of dissolution channels significantly contributed to an increase in borehole productivity (> 10 l/s).

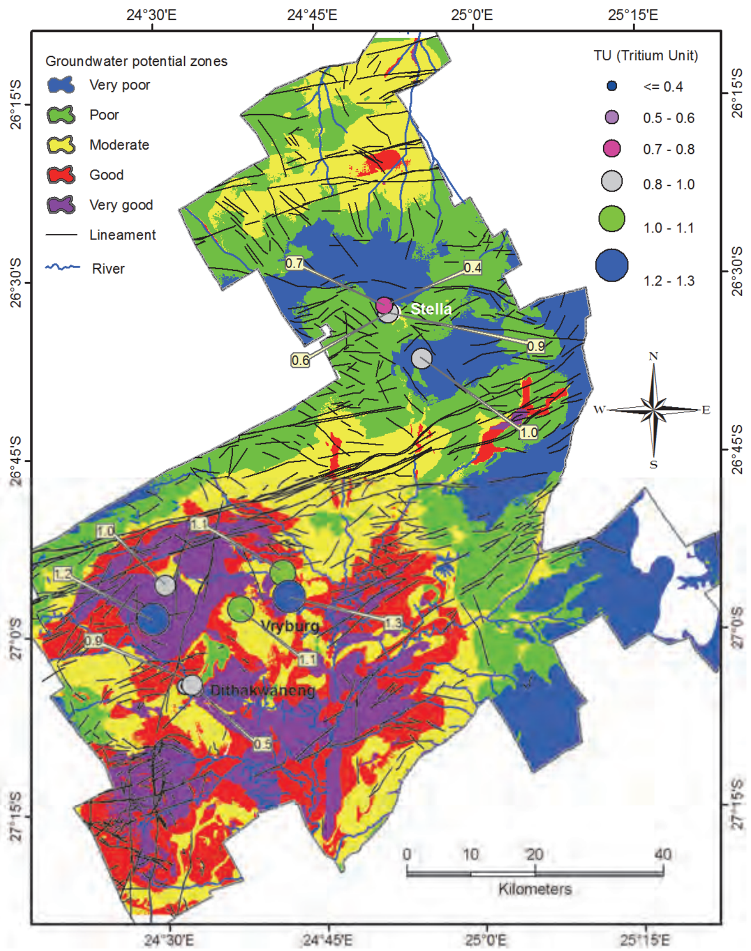


Figure 37: Groundwater potential zones superimposed on borehole yield. DD and LD stand for data gap in drainage density and lineaments density, respectively.

7.4 Weights of Evidence

7.4.1 Background and its Application

The Weights of Evidence (WofE) is a data-driven statistical modelling approach that uses known occurrences as a model training site to create the map of potential targets from weighted continuous or categorical thematic layers (Poli and Sterlacchini, 2007). A final response map can be derived by the combination of evidential themes in support of a hypothesis about the occurrence of a target (e.g. high yielding wellfield). According to Poli and Sterlacchini (2007), the WofE is based on prior and posteriori probability; the former is the probability that the area under investigation consists of high yielding boreholes prior to considering factors; the latter is an "adjustment" or response of prior probability, taking into consideration the evidence from one or more spatial patterns.

7.4.2 Formulation of WofE

Suppose an evidential theme geology "G" is used for characterization of groundwater potential. The WofE method requires training data to test the correlation between known occurrences of high borehole yield and geology. So, let us assume that we have training data "t", and it consists of the location of boreholes with yield > 10 l/s. Suppose that geology "G" is classified into 5 or N and each class can be represented as "Gi", where i varies from 1 to 5 or N depending on the variability of lithology in the area where we search for groundwater. We can then use Bayesian theorem to calculate the conditional probability to have training point "t" given the class "Gi". This conditional probability (CP) can be expressed using Bonham-Carter (1996) approach:

$$P(t|Gi) = \{P(Gi|t)*P(t)\}/P(Gi)$$

Where P(Bi|t) is the conditional probability to have "Gi" given t. P(t) is the prior probability to find "t" within the study area "AS", and P(Gi) is the prior probability to find the class "Gi" within the study area "AS".

According to Bonham-Carter (1996), the conditional probability to have “t” when the class “Bi” is not given can be expressed as:

$$P(t|Gi^0) = \{P(Gi^0|t)*P(t)\}/P(Gi^0)$$

Where $P(Gi^0|(t)$ is the conditional probability not to have the class “Gi” given “t”, P(t) is the prior probability to find “t” within the study area “AS” and $P(Gi^0)$ is the prior probability not to find the class “Gi” within the study area “AS”. In practice the above equations can be expressed in a more simplified form:

P(t) = area “t” / area “AS”
P(Gi) = area “Gi” / “AS”
$P(Gi^0)$ = area “Gi^0” / “AS”
P(t|Gi) = (area “t ∩ Gi” / area “Gi”) / P(Gi)
P(Gi|t) = (area “t ∩ Gi” / area “Gi”) / P(t)
$P(t|Gi^0)$ = (area “t ∩ Gi^0” / area “Gi^o” / area “Gi^0”) $P(Gi^0)$
$P(Bi^0|t)$ = (area “t ∩ Gi^0” / area “Gi^0”) / P(t)

According to Bonham-Carter (1996), the weight of each class of geology can be expressed in terms of weights $\mathbf{W^+}$, $\mathbf{W^-}$ and the difference between the two weights or contrast **C.** This can be expressed as:

$\mathbf{W^+} = \ln \{P(Gi\,|t) / P(Gi)\}$
$\mathbf{W^-} = \ln \{P(Gi^0\,|t) / P(Gi^0)\}$
$\mathbf{C} = \mathbf{W^+} - \mathbf{W^-}$

A value of contrast **C** equal to zero indicates that a specific lithology (class) of geology is not significant in controlling productive wellfield. A positive contrast indicates a positive spatial correlation, vice versa for a negative contrast. According to the WofE approach, the probability of finding productive wellfield depends on the linear combination of the weights of different lithologies. Then the value of posteriori probability “Pp(t) for unique condition unit (UCU) can be expressed as:

$$P_P = \sum_{i}^{N} W_i * \ln P_i(t)$$

Where W_i and P_i are weights and posteriori probability corresponding to a given evidential them. If N number of evidential themes are used, the posteriori probability will be the linear combination of all layers.

In the present study, the evidential themes used for modelling groundwater potential are: geology, lineaments, topography/slope, soil and land cover data. The WofE uses Bayesian statistics to calculate the strength of the spatial association between the training data (e.g. borehole with known yield) and the evidential themes (geology, lineament, DEM, etc.). Based on the spatial relationship, the weight corresponding to each class of the evidential theme was calculated. The algorithm calculates the weight ***W+***, i.e. the weight due to the presence of a high borehole yield, and ***W-***, the weight due to the absence of high borehole yield. The difference between **W+** and **W-** is known as the ***Contrast C***, which is a measure of the strength of the spatial relationship between the training points and each class of the thematic layer (geology, lineament, DEM, etc.). No relationship would result a 0 **C** value, i.e. the evidential theme does not indicate the presence or the absence of high yielding wellfield. The ratio of the weights contrast **C** to the standard deviation of the contrast provides the level of uncertainty in determining the weighting parameters. This approach therefore allows the user to adjust and select the most appropriate groupings or thresholds of the evidential themes that are used for groundwater potential mapping. The main aim is to maximize the contrast between **W+** and **W-**, i.e. **C** while ensuring the level of uncertainty is within an acceptable range.

The use of WofE for groundwater potential mapping is objective, i.e. it reduces subjective choice of weighting parameters. In addition, multiple layers can be combined, which optimizes the probability of identifying potential targets. Furthermore, it assumes conditional independence corresponding to each of the evidential themes used for calculation of posteriori probability. The WofE is one of robust spatial modelling approaches used for mapping groundwater potential.

In this study, 46 boreholes with yield > 10 l/s were used as training points to assess the groundwater potential of the area. The prior probability in finding groundwater is the ratio of the number of training points to the total area in sq. km. Where no other information available, this ratio can be interpreted as the prior probability of finding groundwater in the area. The initial probability can be successively modified or updated with the addition of new evidential theme used for calculation of posteriori probability. So, the posteriori probability from adding one piece of evidence can be treated as the prior for adding a new piece of evidence. Based on this approach, five evidential themes were combined together to give posteriori probability. Where the weights of evidential theme positive, but less than 1, the evidential theme results the posteriori probability to be less than prior probability.

The following are summary of the steps used for calculation of posteriori probability that shows groundwater potential map of the area:

Step 1: Construction of relevant database that are used for mapping groundwater potential.

Step 2: Generate raster data for each evidential theme.

Step 3: Classify each thematic layer based on hydrogeological attributes.

Step 4: Selection of suitable and coherent boreholes that can be used as a training site. Boreholes with yield greater than 10 l/s were extracted from the borehole database covering the study area. The extracted boreholes were superimposed on each classified evidential theme. The boreholes that fall close to each other within the same class corresponding to each evidential theme were excluded from the training points.

Step 5: Select suitable unit area for the training points. The unit area for the present study is chosen to be 0.25 km^2.

Step 6: Calculate weights and other statistics for each evidential theme.

Step 7: Use the Join Data command to append the newly created weights table to its evidential theme. Use the Value field in the evidential theme and Class field in weights to append or join.

Step 8: Display the layer with the Gen Class field. The Gen Class field shows which areas of the evidential theme shared greater association with the location of the training points.

Step 9: Check the *Contrast* which defines the limit of the association. Also check the *Studentized Contrast* if greater than 2 which indicates the confidence of the predicted results is greater than 98%. In addition, check Gen Class if it is 2 (favorable) and 1 (unfavorable).

Step 10: Calculate the response of the weighted evidential theme and generate posteriori probability showing groundwater potential of the area.

Step 11: Use conditional independence (CI test) to assess whether the WofE assumption has been violated.

Step 12: Calculate cumulative Area frequency and plot SRC curve. Use the SRC curve to symbolize posteriori probability.

Step 13: Generate groundwater potential map of the area using different symbols or colour.

7.4.3 Posteriori Probability Map

In this study, based on the above steps, groundwater potential zones were generated showing the posteriori probability that a unit area contains a training point. The posteriori probabilities resulting from the statistical analysis of the five evidential themes rage from 0.0000826 to 0.031827. The two methods used to demarcate thresholds for the posteriori probability classes are Cumulative Area-Posteriori Probability (CAPP) and Success Rate Curve (SRC), both of which can be derived from the Agterberg-Cheng-CI-Test (Agterberg and Cheng, 2002). The graph of percentage of cumulative area versus posteriori probability was plotted (SRC curve) to allow delineation of class breaks showing groundwater potential zones (Figures 38a and 38c). The class breaks were selected where a notable stepwise increase in posteriori probability relative to the cumulative study area (Figure 38c). Three class breaks were identified using this approach which allowed in creating four relative groundwater potential zones (Figure 38a). The class ranges from highly favorable to least favorable, where the most favorable class tends to be the greatest probability of groundwater occurrence. The high probable zones correspond to carbonate rocks, alluvial deposits and highly fractured hard rocks. Several high potential areas can also be observed along the banks of main river course. Conversely, areas shown by the least favorable are associated with the volcanic rocks of the Ventersdorp Supergroup and partly correspond to the Kalahari sand in the northern part of the area (Figure 38a).

7.4.4 Conditional Independence

An assumption is made when using WofE that there is conditional independence between the evidential themes used for predicting groundwater potential. Conditional independence is violated when the presence of one evidential theme influences the probability of another evidential theme. The validity of posteriori probability derived from the WofE depends on the degree of conditional independence. When the conditional independence is violated the model can over predict groundwater potential zones in some areas. According to Bonham-Carter (1994), if all the layers used for modelling groundwater potential are conditionally independent, the *calculated/expected* (*predicted*) number of sites will be equal to the number of boreholes used as training points. In addition, the number of training points divided by the number of predicted sites gives a diagnostic indication of conditional independence, i.e. CI ratio of < 0.85 suggests violation of the condition.

In the present study, the CI of the five evidential themes was assessed using Agterberg (1992) approach. The total number of observed training points is 46 and the predicted borehole sites are 33, that gives a CI ratio of 1.37. The result indicates the correlations among the evidential themes used for modelling is not significant. Although WofE assumes conditional independence in predicting potential borehole sites, in practice when working with geological data, conditional independence may be violated. However, the predictive power of the resulting posteriori probability can be tested by analyzing their success rate and prediction rate. This can be achieved by plotting success rate curve (SRC) and prediction-rate curve (PRC).

The SRC curve is a plot of the percentage of cumulative training sites versus cumulative training area, from high to low posteriori probability. It indicates how much percentage of all high borehole yields occupy the pixels of the raster map derived from the WofE (Fabbri and Chung, 2008). Furthermore, it shows how well the model and the evidential themes predict the groundwater potential of the area. In addition, the SRC can be used for classification of the posteriori probability map. The prediction rate curve is a graph of the cumulative testing sites versus cumulative area, from high to low posteriori probability. It is mainly used for model validation and quantifies how well prospectivity map delineates the

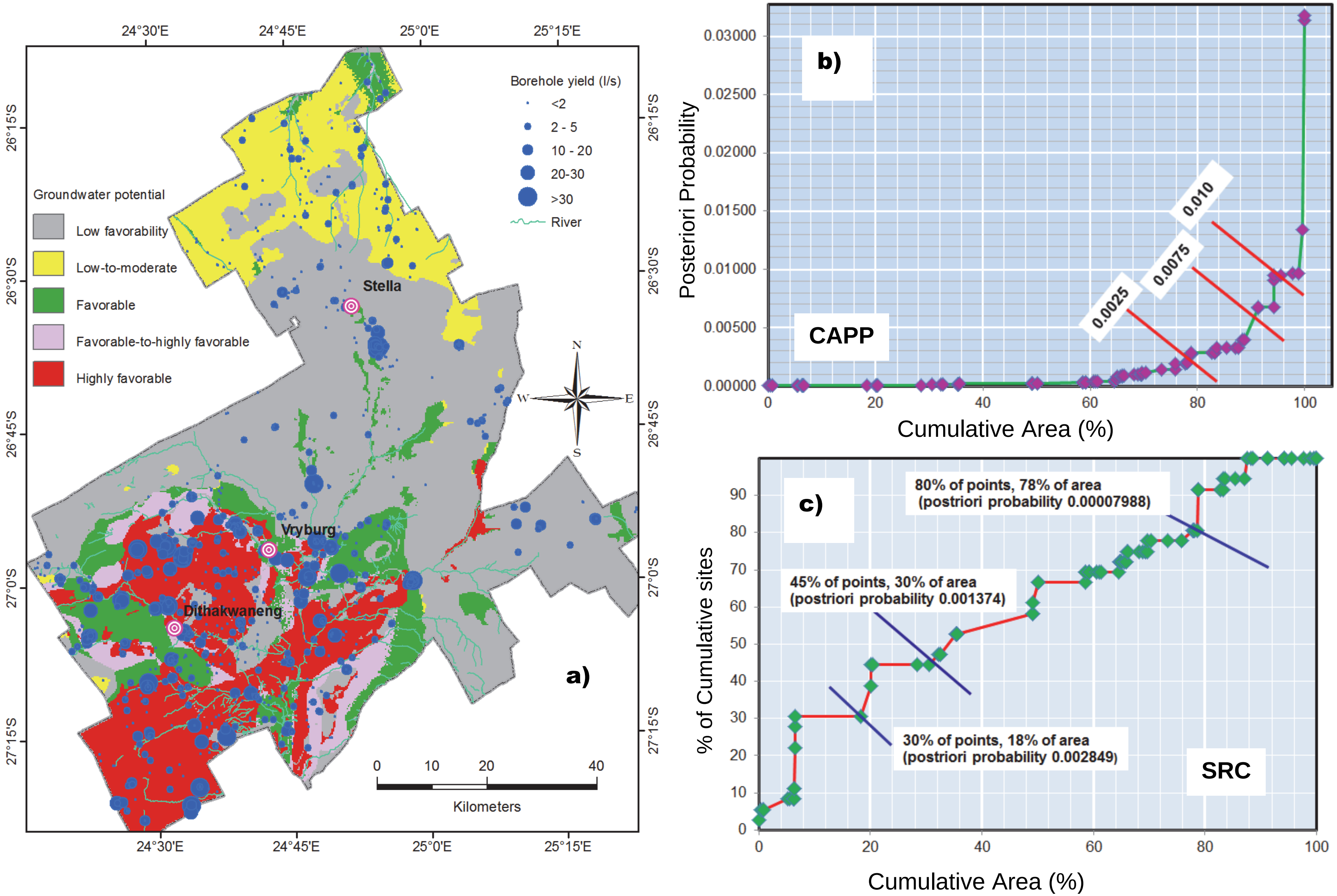


Figure 38: **a)** Posteriori probability map of groundwater potential derived from WofE. **b)** and **c)** Cumulative posteriori probability and number of predicted sited versus cumulative area, respectively.

training points. The PRC describes the ability of prospectivity map to predict features that are undiscovered or unknown. Figure 38c is an SRC curve plotted using the percentage of predicted potential sites versus cumulative area frequency. It shows the variation of cumulative predicted sites with cumulative area. There are three inflection points on this curve at which the slope gradually increases towards the end. Thus, four groundwater potential zones can be identified from the three inflection points. The four potential zones can be defined as *high potential area*, *moderate-to-high*, *moderate* and *low potential* zone, respectively.

The threshold probability corresponding to the upper and lower inflection points on Figure 38c are 0.002849 and 0.00010, respectively. The corresponding percentage of predicted number of potential sites can be 30% and 80%, respectively. This implies 30% of the total number of predicted sites exhibit high probability of groundwater occurrence and cover 18% of the total study area. On the other hand, 80% of the total number of predicted sites are characterized by poor probability, but cover 78% of the study area.

The area under the SRC curve can be used to provide an estimated *efficiency* score showing the measure of the prediction accuracy of the WofE (Swatzky et al., 2009b). A total area equal to 1 denotes perfect prediction accuracy of groundwater potential zones. Figure 38c shows ~88% efficiency of classification and the curve reflects how well the model classifies the distribution of groundwater potential zones.

7.4.5 Statistical Results

The calculated weights and the statistical results for lithology, lineaments, slope, land type and soil texture are shown on Table 9. The positive weight indicates areas where training points are likely to occur, conversely, a negative weight suggests areas where the training points are absent. The contrast column C defines the limit of the association between the groundwater potential sites and the evidential theme. It is an important measure of how well the overall evidential theme predicts the occurrence of groundwater. The higher the contrast, the greater significance in predicting the availability of groundwater. For example, on Table 9, the maximum contrast among lithology group corresponds to Class 3 (alluvial deposits), suggesting that the availability of coarse sedimentary materials associated with alluvial deposits is a good indication of groundwater occurrence.

The studentized contrast (confidence) is the ratio of the contrast and its standard deviation. The confidence value is greater than 2 corresponding to the first 5 classes in geology, suggesting ~99.5% level of significance in predicting groundwater occurrence. The studentized contrast indicates sandy-clay and alluvial deposits have a strong association with the high borehole yields used as a training points compared to other lithological units, indicating these rocks are the primary predictor of favorable sites. On the other hand, the volcanic rocks (Class 1) have negative contrast and low confidence level, suggesting the hard rocks are a strong predictor of areas where there is no training points (strongly negative weights contrast, Table 9). The weight contrast corresponding to the different classes of lineaments density is relatively low compared to lithologic units. The poorly interconnected dense lineaments partly contributed to the lack of weight contrast. However, Class 4 in the lineament group is the only exception where the studentized contrast suggests a strong association between lineament density and the training points (high borehole yield). This Class corresponds to the volcanic rocks of the Ventersdorp Supergroup, where interconnected fractures play a significant role in controlling borehole yield. As shown by the high weight contrast, morphologically flat areas strongly correlate with the training points. The variation in studentized contrast corresponding to the different categories of land types is consistent with the land use attributes, whereby irrigation return flows and wetlands (river banks) are indicative of the availability of the training points. As shown on Table 9, loam-sand and sandy-soil are good indicators of the availability of groundwater.

7.4.6 Confidence Map

There are two approaches that can be used to evaluate the uncertainty or the level of confidence of the posteriori probability map. The first method is the confidence corresponding to each of the classes of evidential theme as reported on Table 9 and discussed in the above section. The second type of confidence can be derived from the ratio of posteriori probability and its total uncertainty or standard deviation. Areas with a high posteriori probability tend to have higher confidence values and therefore have a higher level of certainty with respect to predicting groundwater potential zones. Figure 39 shows uncertainty in the posteriori probability map derived from the WofE model.

Table 9: Statistical results of WofE.

Geology														
Class	**Code**	**Area (sq.km)**	**Area Units**	**No Points**	**W+**	**σw+**	**W-**	**σW-**	**C**	**σC**	**Stud cont**	**Gen Class**	**Weight**	**σWeight**
6	6	762.31	3049.25	10	0.7301	0.3167	-0.1381	0.1691	0.8682	0.3591	2.4178	2	0.667300	0.189300
5	5	1053.44	4213.75	13	0.6688	0.2778	-0.1808	0.1769	0.8496	0.3293	2.5798	2	0.667300	0.189300
4	4	1498.94	5995.75	18	0.6415	0.2361	-0.2743	0.1926	0.9158	0.3046	3.0062	2	0.667300	0.189300
3	3	2272.31	9089.25	28	0.6673	0.1893	-0.5889	0.2426	1.2562	0.3077	4.0822	2	0.667300	0.189300
2	2	4113.38	16453.50	34	0.2670	0.1717	-0.5449	0.3016	0.8119	0.3471	2.3392	1	-0.588900	0.242600
1	1	7107.13	28428.50	45	-0.0002	0.1492	6.4469	14.1421	-6.4471	14.1429	-0.4559	1	-0.588900	0.242600
Lineament														
Class	**Code**	**Area (sq. km)**	**Area Units**	**No Points**	**W+**	**σw+**	**W-**	**σW-**	**C**	**σC**	**Stud cont**	**Gen Class**	**Weight**	**σWeight**
2	2	1428.81	5715.25	12	0.2254	0.2890	-0.0708	0.1742	0.2962	0.3374	0.8778	2	0.172100	0.156300
3	3	3427.25	13709.00	26	0.1235	0.1963	-0.1475	0.2296	0.2710	0.3021	0.8970	2	0.172100	0.156300
4	4	5148.63	20594.50	41	0.1721	0.1563	-0.9633	0.5002	1.1353	0.5240	2.1665	2	0.172100	0.156300
5	5	6237.38	24949.50	42	0.0040	0.1544	-0.0546	0.5778	0.0586	0.5981	0.0980	1	-0.963300	0.500200
6	6	6709.75	26839.00	45	-0.0002	0.1492	6.3893	14.1421	-6.3895	14.1429	-0.4518	1	-0.963300	0.500200
DEM														
Class	**Code**	**Area (sq. km)**	**Area Units**	**No Points**	**W+**	**σw+**	**W-**	**σW-**	**C**	**σC**	**Stud cont**	**Gen Class**	**Weight**	**σWeight**
1	1	602.81	2411.25	3	-0.2434	0.5777	0.0194	0.1526	-0.2628	0.5975	-0.4398	99	-0.117404	0.164515
2	2	1878.94	7515.75	11	-0.0807	0.3017	0.0268	0.1692	-0.1075	0.3459	-0.3108	99	-0.117404	0.164515
3	3	2414.19	9656.75	11	-0.3317	0.3017	0.1320	0.1692	-0.4637	0.3459	-1.3406	99	-0.117404	0.164515
4	4	1575.13	6300.50	10	0.0004	0.3165	-0.0001	0.1668	0.0006	0.3577	0.0016	99	-0.117404	0.164515
5	5	693.06	2772.25	9	0.7177	0.3339	-0.1174	0.1645	0.8351	0.3722	2.2437	5	0.717700	0.333900
6	6	84.69	338.75	2	1.3185	0.7092	-0.0328	0.1509	1.3512	0.7251	1.8636	99	-0.117404	0.164515
Land type														
Class	**Code**	**Area (sq. km)**	**Area Units**	**No Points**	**W+**	**σw+**	**W-**	**σW-**	**C**	**σC**	**Stud cont**	**Gen Class**	**Weight**	**σWeight**
1	1	1055.31	4221.25	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	99	-0.047793	0.229589
2	2	2896.69	11586.75	8	-0.8313	0.3537	0.3187	0.1624	-1.1500	0.3892	-2.9551	2	-0.831300	0.353700
3	3	1026.19	4104.75	8	0.2077	0.3539	-0.0387	0.1623	0.2464	0.3894	0.6328	99	-0.047793	0.229589
4	4	1063.63	4254.50	11	0.4909	0.3019	-0.1150	0.1692	0.6060	0.3461	1.7510	99	-0.047793	0.229589

5	5	523.63	2094.50	7	0.7483	0.3786	-0.0904	0.1602	0.8387	0.4111	2.0401	5	0.748300	0.378600
6	6	694.25	2777.00	12	1.0063	0.2893	-0.2021	0.1716	1.2083	0.3364	3.5922	6	1.006300	0.289300
Soil														
Class	**Code**	**Area (sq. km)**	**Area Units**	**No Points**	**W+**	**σw+**	**W-**	**σW-**	**C**	**σC**	**Stud cont**	**Gen Class**	**Weight**	**σWeight**
1	1	1600.88	6403.50	2	-1.6066	0.7072	0.2050	0.1526	-1.8117	0.7235	-2.5040	1	-1.606600	0.707200
2	2	3033.81	12135.25	9	-0.7414	0.3335	0.3212	0.1668	-1.0626	0.3729	-2.8497	2	-0.741400	0.333500
3	3	33.56	134.25	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	99	-3.277860	10.000294
4	4	9.00	36.00	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	99	-3.277860	10.000294
5	5	1770.13	7080.50	24	0.7809	0.2045	-0.4821	0.2183	1.2629	0.2991	4.2222	5	0.780900	0.204500
6	6	785.88	3143.50	10	0.7172	0.3167	-0.1365	0.1691	0.8537	0.3591	2.3775	6	0.717200	0.316700

The columns headings are described in acronyms, abbreviation and in the Glossary. The yellow shaded fields show *studentized contrast* ≥ 2 (98% *confidence as shown on Table 10*) and Gen Class ≥2 (favorable for groundwater occurrence). The blue shaded fields corresponding to the *studentized contrast* and *Gen Class* indicate poor confidence in prediction of groundwater occurrence and low favorability, respectively. Table 9 summarizes the range of *studentized contrast* and the level of confidence in predicting groundwater potential.

Below is description of Classes for each evidential theme:

Geology
1. Hard rocks, very low permeability
2. Clayey-sand, low permeability
3. Alluvium, moderate permeability
4. Sandy-clay, moderate-to-high permeability
5. Chert free dolomite
6. Chert-rich dolomite, high permeability

Lineament (density, sq. km)
1. < 0.5
2. 0.5-1.5
3. 1.5-2.5
4. 2.5-3.0
5. 3.0-3.5
6. >3.5

Soil texture
1. Clayey
2. Silty soil
3. Silty-loam
4. Sandy-loam
5. Loam-sand
6. Sandy-soil

Topography (slope)
1. Very steep slope
2. Steep
3. Moderate
4. Gentle slope
5. Gentle-to-flat
6. Flat

Land type
1. Rocky, scattered grassland
2. Dry-land, no farms
3. Scattered bush-land
4. Moderate wetland
5. Semi-irrigate land
6. Well irrigated land

Table 10: Studentized contrast and the level of confidence.

Confidence (%)	Studentized contrast
99.5	2.576
99	2.326
97.5	1.96
95	1.645
90	1.282
80	0.842
70	0.542
60	0.253

The uncertainty of the model is relatively large in the central part of the study area. A small population of training points along with missing data raises the total uncertainty of the response of the themes, which in turn lowers the confidence in the northern and central parts of the study area (Figure 39). In addition, the model shows moderate uncertainty corresponding to the Kalahari sand owing to lack of conditional independence of the evidential themes lithology and lineaments. In general, the confidence (posteriori probability) map of the response theme shows lower level of significance than those calculated for each evidential theme shown on Table 9.

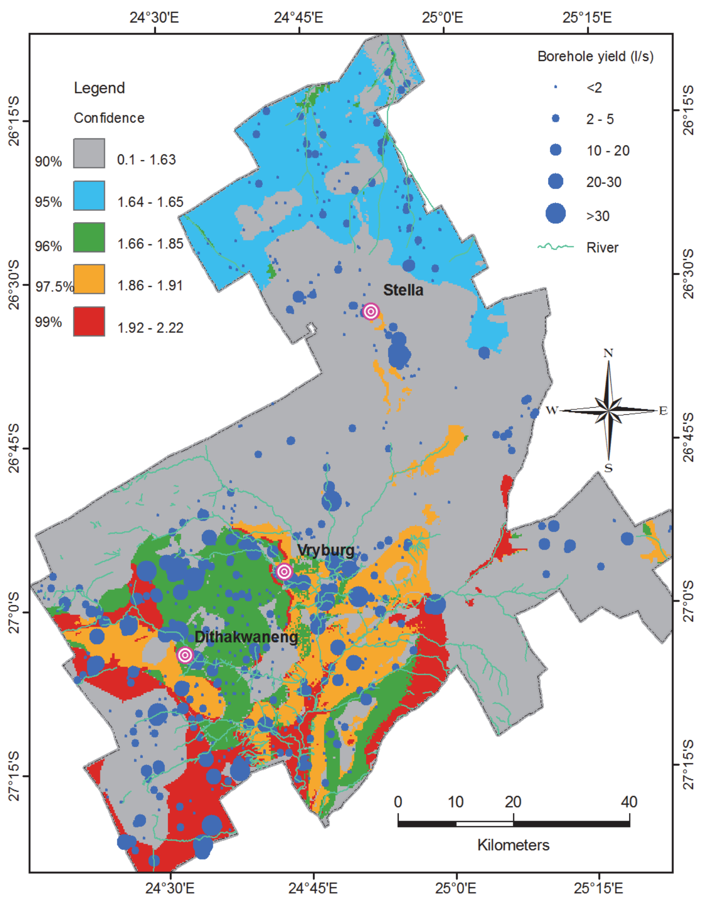


Figure 39: Confidence map derived from statistical analysis of the five themes.

7.5 Logistic Regression

7.5.1 Background

Logistic regression (LR) is a type of probabilistic statistical classification model used for predicting the outcome of a categorical dependent variable, i.e. a class label based on one or more predictor variables. The principle of LR is based on the analysis of a problem, in which a result measured with dichotomous variables such as 0 and 1 or *true* and *false*, is determined from one or more independent factors (Menrad, 1995). In LR analysis, the values of dependent variable are to be predicted or explained using the values of independent variables. The more independent variables are included, the more complete the model will be, but only when they play a major role in determining the dependent variable. The general consensus in the selection of parameters is that any independent variable can be *optional* (has a certain degree of affinity with the dependent variable), *complete* (is fairly represented all over the study area), *non-uniform* (varies spatially), *measureable* (can be expressed by any of the different types of measuring scales) and *non-redundant* (its effect should not account for double consequences in the final result).

The LR is a function that can be used to account for the inflated probabilities associated with conditional independence. LR is similar to linear regression, however, because the evidence is reduced to binary themes, the response variable can only be divided into two classes (i.e. presence or absence of training points), whereas linear regression can have continuous values ranging from 0 to 1. The results of LR do not differ greatly from the WofE. The main difference is the posteriori probabilities of the response theme with conditional independence problem improves when using LR. In general, the patterns of the response themes are nearly identical, but the use of LR provides better statistical results that explain the posteriori probability regardless of conditional independence issues.

In the case of groundwater potential mapping, the goal of LR would be to find the best fitting model that describes the relationship between the *presence* or *absence* of high yielding wellfields and a set of independent variables such as lithology, lineaments, geomorphology, land use and soil texture. LG generates the model statistics and coefficients that can be used to predict a *logit* transformation of the probability that the dependent variable is 1 (probability of occurrence of high yielding wellfield). The *logit* of the probability p between 0 and 1 can be expressed as:

$$Logit\ (p) = \log (p/1-p) = \log (p) - \log (1-p)$$

In the above equation, if p is the probability then $p/(1-p)$ is the corresponding *odds*, i.e. the *logit* of the probability is the logarithm of the odds. Generally, LG involves fitting the dependent variable using the equation of the form:

$$Y = Log_e \left[\frac{p}{1-p} \right] = Log\ it\ (p)$$

$$p = \frac{e^y}{1+e^y}$$

$$Y = C_o + C_1X_1 + C_2X_2 + C_3X_3 + \ldots\ldots + C_nX_n$$

Where p is the probability that the dependent variable (Y) is 1; $p/(1-p)$ is the so called *odds* or likelihood ratio, C_o is the intercept, and C_1, C_2, C_3………C_n are regression coefficients to be estimated, which measure the contribution of independent factors (X_1, X_2, X_3, …X_n). Y is the dependent variable expressed in terms of the linear combination function of the independent variables.

In order to appropriately interpret the meaning of eq. 1, one has to use the coefficients as a power to the natural log(e). The result represents the odds ratio or the probability that an event will occur divided by the probability that it fails to do so. If z is denoted as a binary response variable (0 or 1), value 1 (z=1) means the presence of high borehole yield and value o (z = 0) indicates the absence of a high borehole yield. In the above equation, P refers to the probability of occurrence of high yielding wellfield, i.e. z = 1. The function y can be expressed as logit (P), that is the log (to the base e) of the odds or likelihood ratio that the dependent variable z is equal to 1.

The probability P strictly increases when the value of Y increases. The regression coefficients C_1, C_2, C_3………C_n show the contribution of each independent variables on the probability value P. If the coefficient is positive, its transformed log value will be greater than 1, indicating that the event is more likely to occur. If a coefficient is negative, the latter will be less than one and the odds of the event occurring decreases. A coefficient 0 has a transformed log value of 1, and it does not change the odds one way or the other. For a positive coefficient, the probability plotted against the values of an independent variable follows an S-shaped curve. A mirror image will be obtained for a negative coefficient (Menrad, 1995).

The LG was applied to five independent evidential themes, *viz*. geology, lineament, topography, soil and land use. Each evidential theme was converted into raster using a grid cell size of 250 m by 250 m then classified into five depending on the values and the hydrological setting of the area. A total of 46 boreholes with yield > 10 l/s were used as a training site for calculation of the posteriori probability corresponding to each class for each evidential theme. The weight for each evidential theme was calculated and the result was used together with the training points as an input to logistic regression modelling algorithm. The result gives posteriori probability showing groundwater potential map of the study area as shown on Figure 40.

7.5.2 Statistical Results of LR

The relative importance of the evidential themes can be assessed using the LR coefficients c_1, c_2, c_3, …..c_n shown on Table 11. The coefficients corresponding to *lithology*, *lineaments*, *slope*, *land use* and *soil* are '*positive*', indicating that they are positively related to the probability of groundwater occurrence. The coefficients which strongly depart from 0 can be described as significant parameter in influencing the occurrence of wellfield with high yield. In this study, the coefficient corresponding to lineament is greater than the remaining evidential themes suggesting that the productivity of a wellfield is largely determined by cross-cutting structures. In addition, the high coefficient corresponding to lithology suggests its significance in controlling the distribution of high borehole yield. The coefficients corresponding to land use and soil are close to 0, indicating that the two evidential themes have little influence on groundwater occurrence.

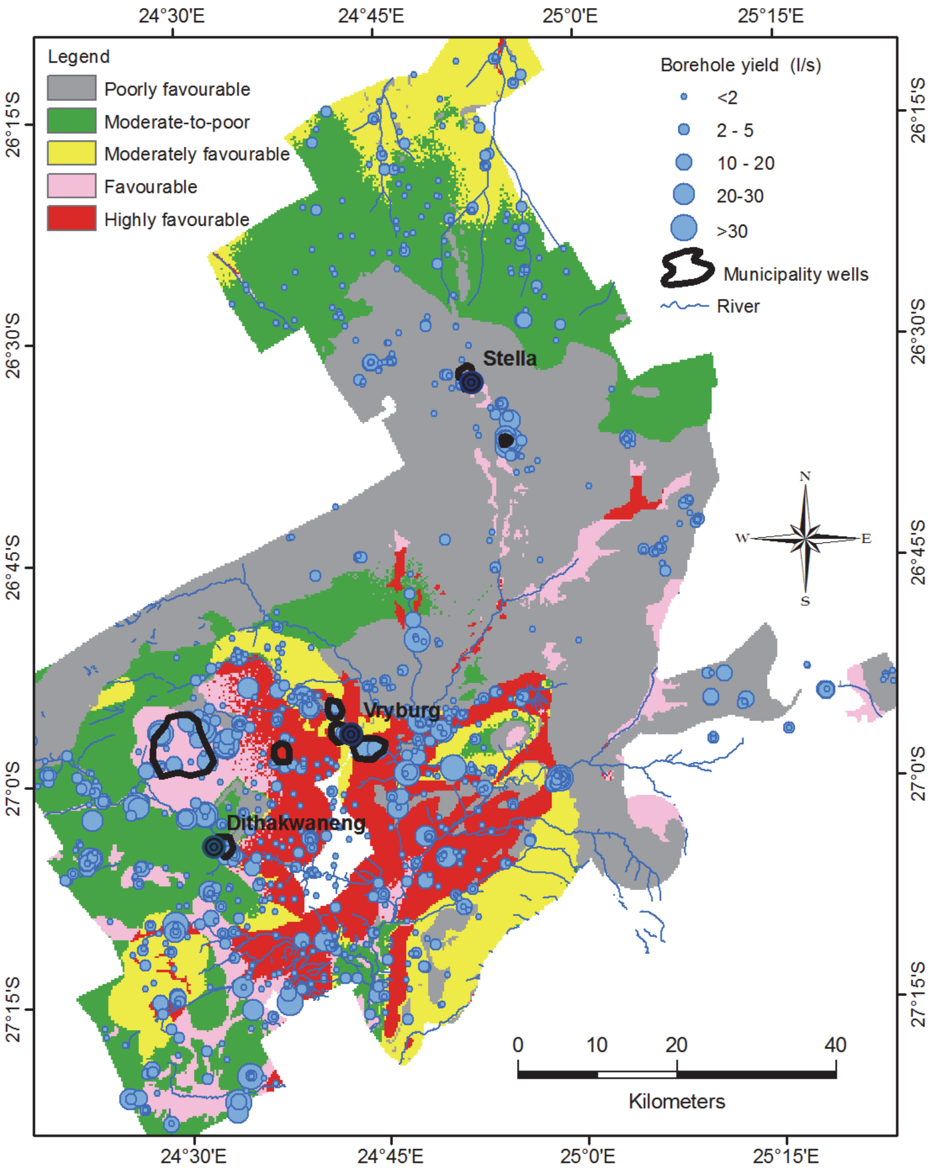


Figure 40: Posteriori probability map of groundwater potential derived from Logistic Regression

Table 11: Regression coefficients for the five evidential themes

Evidential theme	Coefficient
Lithology	0.899774
Lineament	0.905895
Slope/geomorphology	0.409097
Land use/land type	0.092465
Soil	0.00112

Table 12:Probability of groundwater potential zones and coverage

Probability range	**Description**	**Coverage (%)**
0.0-0.035	Poorly favorable	39
0.045-0.085	Moderate to-poor	26
0.155-0.210	Moderately favorable	15
0.25-0.457	Favorable	8
0.500-0.865	Highly favorable	12

7.5.3 Predicted Groundwater Potential

The final results of LR are numbers showing probability that varies from 0 to 1. The numbers close to 1indicate the likelihood of groundwater occurrence, whereas numbers that are close to 0 suggest the occurrence of groundwater is very unlikely. As sown on Table 11, the probability ranges from 0 to 0.865 and the high values are located in the southern part of the area and along the main river channels where gravel deposits act as favorable environment for groundwater occurrence. The area shown by low probability ranging from 0 to 0.035 covers 39% of the study area and largely coincides with the Ventersdorp volcanic rocks. This suggests that the availability of groundwater within the crystalline rocks entirely depend on fracture connectivity and density of cross-cutting structures. Very high probability ranging from 0.65 to 0.865 forms 12% of the study area, while moderate-to-high and high probability constitutes 23% of the study area. The region indicated as very high probability is owing to the presence of dissolution channels associated with dolomite and lime stone that are located in the southern part of the area.

In general, the LR method is very useful in predicting groundwater potential zones. As shown on Table 10, lithology, lineaments and topography are the best predictor variables for estimating the probability of groundwater occurrence. Lithology and lineaments have strong positive correlation, while soil and land use have little impact on groundwater occurrence.

7.6 Fuzzy Logic

7.6.1 Background and Approach

A fuzzy set theory was first formulated by Zadeh (1993). According to this author, a fuzzy set of A is a set of ordered pairs that can be expressed as:

$$A = \{[x, \mu_A(x)]x \in X \}$$

Where X is collection of objects, which can be lithology, lineament, soil type or slope, and μ_A (x) is the membership function or degree of compatibility of x in A. The membership function $\mu_A(x)$ maps x to the membership space. Every value of x is associated with the value of μ_A (x), and the ordered pairs [x, μ_A (x)] are known collectively as a fuzzy set (Zadeh, 1993). The shape of the fuzzy membership function need not be linear, it can take any analytical or arbitrary shape appropriate to the problem at hand. Additional insights into fuzzy logic can be found in Tsoukalas and Uhrig (1997), Burrogh and McDonnell (1998).

Fuzzy membership values must lie in the range of (0, 1), but there are no practical constraints on the choice of fuzzy membership values. Values are simply chosen to reflect the degree of membership of a set, based on subjective judgment. Values need not increase or decrease monotonically with class number.

The presence of various classes of a map can be expressed in terms of fuzzy membership of different sets, possibly storing them as several fields in the map attribute table. For example, the density of cross-cutting structures or the type of lithology on a map can be represented in terms of their fuzzy membership of a set *”favourable indicator for high borehole yield"* or “*not favourable*”.

Given two or more maps with fuzzy membership functions for the same set, a variety of operators can be employed to combine the membership values together. An et al. (1991) noted five operators that can be useful for combining thematic layers used for mapping

groundwater potential. These include: fuzzy AND, fuzzy OR, fuzzy Algebraic Product, fuzzy Algebraic Sum and fuzzy Gamma operator.
Fuzzy logic is conceptually easy to understand, and the mathematical concepts behind fuzzy reasoning are very simple. In addition, fuzzy logic is tolerant to imprecise data, i.e. everything is imprecise if you look closely enough, but more than that, most things are imprecise even in careful inspection. Fuzzy reasoning builds this understanding into the process rather than tacking it onto the end.

In the present study, fuzzy logic was applied to 5 thematic layers to produce the groundwater potential map of the study area. The evidential themes used for fuzzy logic analysis include: geology, lineaments, slope, soil and land use. Each evidential theme was expressed in terms of fuzzy set showing membership of "*favourability*" or "*no favourability*" for high borehole yield depending on the class attributes. If a certain class in an evidential theme is favourable for high borehole yield, it will be assigned a membership value of 1 or close to 1. In contrast, if a class of an evidential theme not favourable, it will be assigned a fuzzy membership value of 0 or close to 0. The membership value for each class in an evidential theme depends on the influence on borehole yield. For example, if the geological map of the area is classified on the basis of permeability/porosity, each class will be assigned a fuzzy membership value which varies from 0 to 1. Based on this approach, five evidential themes with each having discrete membership values were generated. The results show that the fuzzy membership values in each evidential theme reflect the relative importance in indicating high borehole yield.

There are several approaches that are used for assigning fuzzy membership values to a class in an evidential theme. Tsoukala and Uhrig (1997) used "Small", "Near" and "Large", while Masters (1993) used Gaussian approach for fuzzification. According to Tsoukala and Uhrig (1997), the fuzzification algorithm "Small" and "Large" are used to indicate small or large values of the classified data (crisp set) are members of the fuzzy set. The spread and mid parameters are subjectively defined to reflect the expert opinion.

In this study, following Tsoukala and Uhrigs' (1997) algorithm, "*fuzzy Large*" was used as a *fuzzy membership function*. The membership function was applied to the five evidential themes to create a membership values varying from 0 to 1. Following this step, the fuzzified thematic layers were combined using "*fuzzy AND*" operator (Tsoukala and Uhrig (1997). The resulting groundwater potential map is shown on Figure 41.

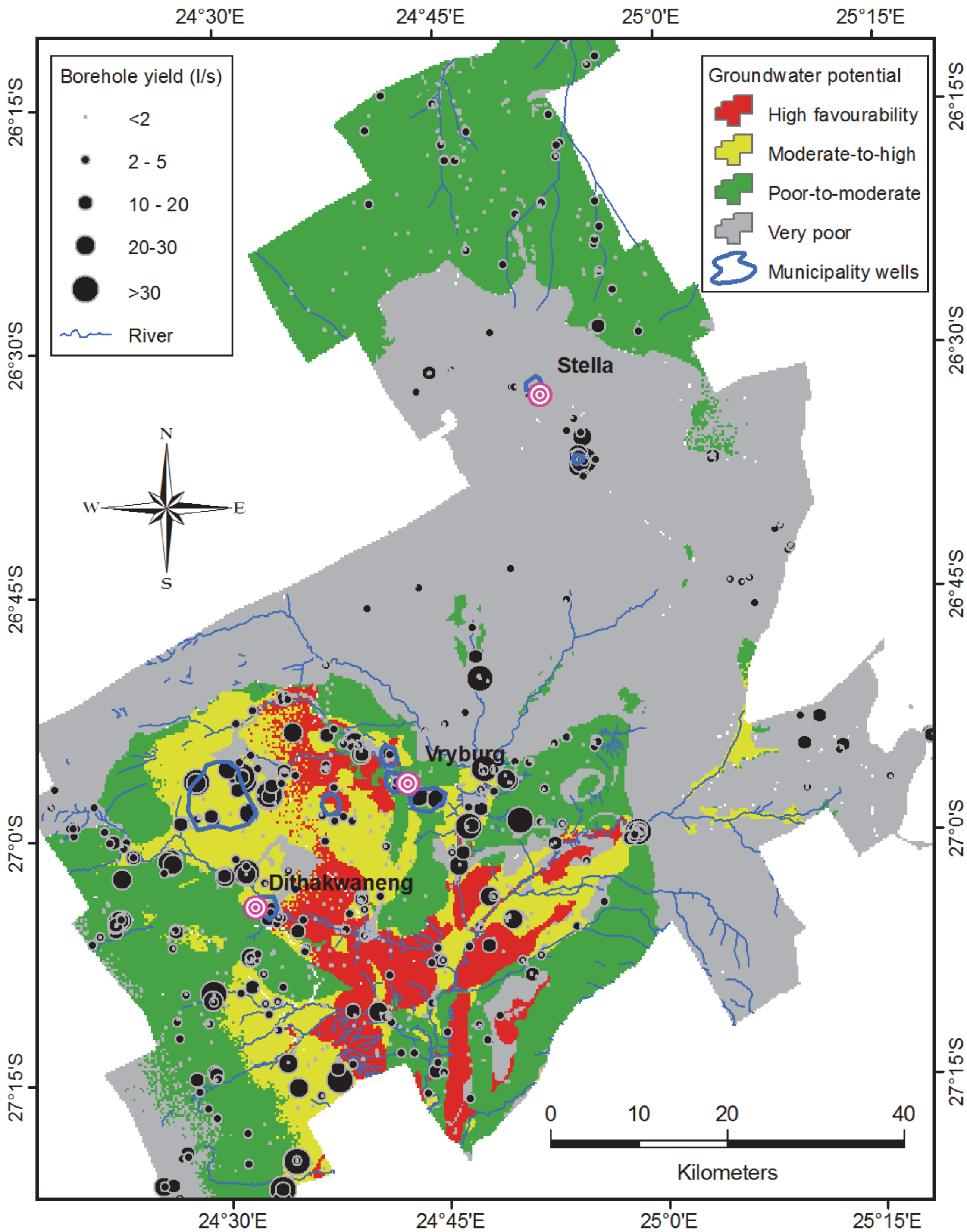


Figure 41: Groundwater potential map of the study area derived from fuzzy logic applied to five thematic layers.

7.7 Neural Networks

7.7.1 Theoretical Background

The brain consists of a large number of neurons, connected with each other by synapses. These networks of neurons are called neural networks, or natural neural networks. The human brain provides proof of the existence of massive neural networks that can succeed at those cognitive, perceptual, and control tasks in which humans are successful. The brain is capable of computationally demanding perceptual acts (e.g. recognition of faces, speech) and control activities (e.g. body movements and body functions). The capability of the brain is its effective use of massive parallelism, the highly parallel computing structure, and the ability to precisely process information. The human brain is a collection of more than 10 billion interconnected neurons. Each neuron is a cell that uses biochemical reactions to receive, process, and transmit information (Ajith, 2005). Figure 42 shows a simplified illustration of a biological neuron.

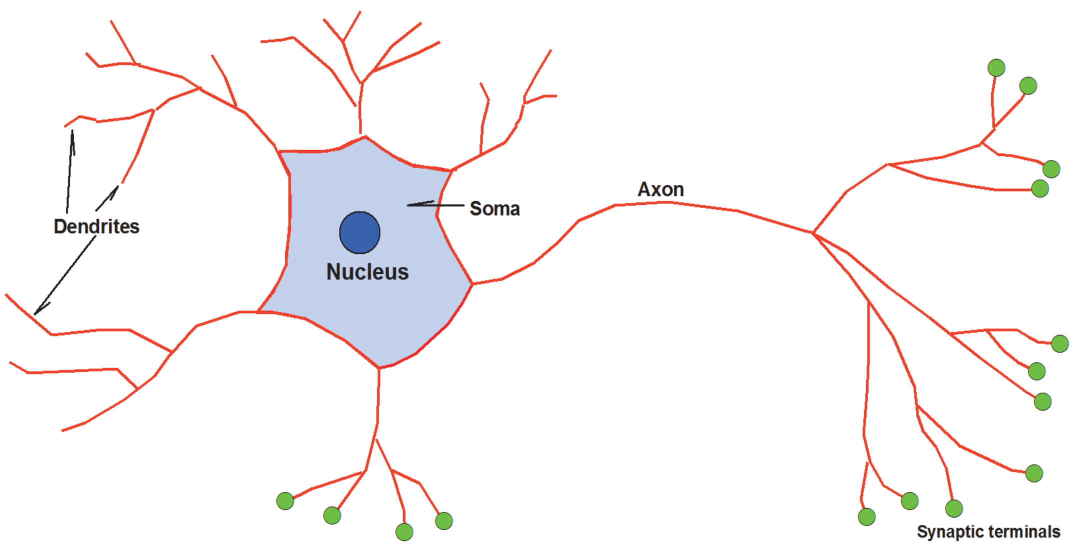


Figure 42: Simplified illustration of biological neuron (after Ajith, 2005).

Treelike networks of nerve fibbers called dendrites are connected to the cell body or soma, where the cell nucleus is located (Figure 42). Extending from the cell body is a single long fibre called the axon, which eventually branches into strands and sub strands which are connected to other neurons through synaptic terminals or synapses. The transmission of signals from one neuron to another at synapses is a complex chemical process in which specific transmitter substances are released from the sending end of the junction. The effect is to raise or lower the electrical potential inside the body of the receiving cell. If the potential reaches a threshold, a pulse is sent down the axon and the cell is "*fired*" (Ajith, 2005).

7.7.2 Artificial Neural Networks

7.7.2.1 Overview of Application

Artificial neural networks (ANNs) refer to computing systems whose central theme is borrowed from the analogy of biological neural networks. ANNs are new information-processing and computing techniques inspired by biological neuron processing (Lee et al., 1998). ANNs are used for applications where formal analysis is difficult or impossible, such as pattern recognition. They have the ability to learn and generalize from examples, and produce meaningful solutions to problems even when input data contain errors or are incomplete. ANNs adapt solutions over time to compensate for changing circumstances and to process information rapidly (Jain et al., 2004). In other words, they learn by adjusting the weights between the neurons in response to the errors between the actual output values and the target output values. At the end of the training phase, the neural networks provide a model that should be able to predict a target value from a given input value.

Three functional groups can be distinguished in the ANNs, i.e. the inputs receiving signals from the network's outside (e.g. yellow boxes in Figure 43) and introducing them into its inside, the neuron which process information (red box), and the neurons which generate the results (**F** in Figure 43). The model includes N inputs (e.g. X1, X2, X3, ….X_N), one output, a summation block and an activation block (Hola and Schabowicz, 2005).

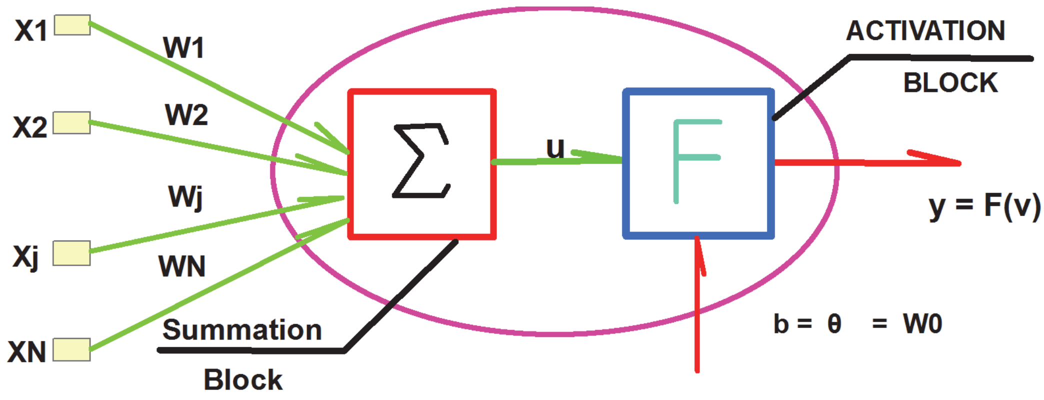


Figure 43: Simplified illustration of Artificial Neurons (after Hola and Schabowicz, 2005).

The network architecture consists of a set of nodes (neurons) connected by links and usually organized in a number of layers. The arrangement of the nodes is referred to as the network architecture. Each node in a layer receives and processes weighted input from a previous layer and transmits its output to nodes in the following layer through links. Each link is assigned a weight, which is a numerical estimate of the connection strength. The weighted summation of inputs to a node is converted to an output according to a transfer function (e.g. **F** and **fi** in Figures 43 and 44).

Most ANNs have three layers or more: an input layer, which is used to present data to the network; an output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors. Determination of appropriate network architecture is one of the most important, but also one of the most difficult tasks in the ANNs model-building process. Unless carefully designed an ANNs model can lead to over parameterization, resulting in an unnecessarily large network (Sudheer et al., 2002). Figure 43 shows schematic illustration of general ANNs model of three layers.

Capabilities of neural networks fall into three broad types of applications: (1) function approximation or regression analysis, (2) time series prediction and modelling, (3) classification, pattern and sequence recognition. Because of these capabilities and the powerful abilities of ANNs, there have been increasing applications of these methods in the earth sciences. During the last decade, various case studies show the wide application of the ANNs for mineral prospectivity mapping and groundwater exploration (e.g. Singer and Kouda, 1996; Harris and Pan, 1999 amongst others).

The basic principle of the application of ANNs for mapping potential target is, it consists of a set of connected nodes or units where each connection has a weight associated with it. ANNs "learn" from prior applications and during the learning phase, the network learns by adjusting the weights (Han and Kamber, 2001). If a poor solution to the problem is made, the network is modified to produce a better solution by changing its weights (Corseni et al., 2009).

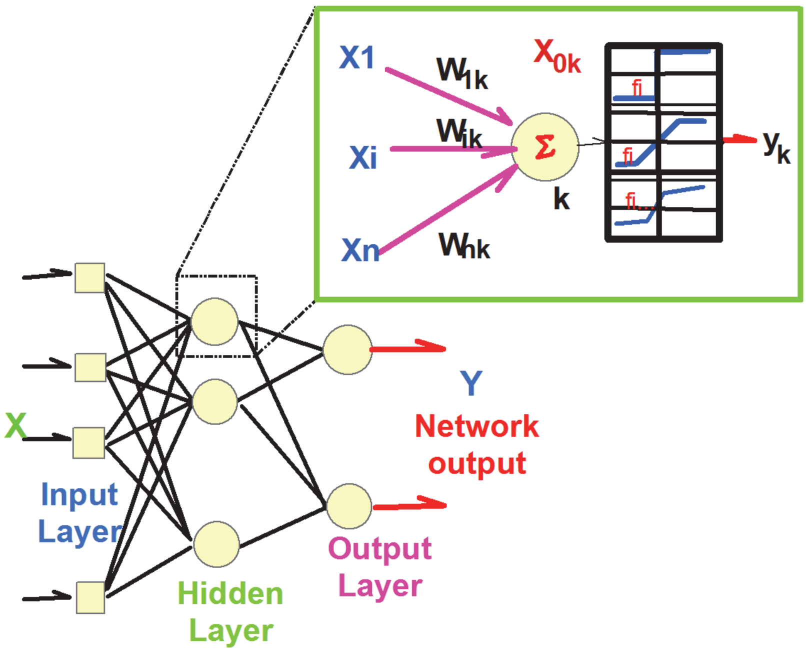


Figure 44: Schematic diagram showing a three layer ANNs and its neurons (after Claudius et al., 2005). Where X_1, X_2, X_n stand for input parameters to the Hidden Layer, W_{1k}, W_{ik} are input weights, X_{ok} is output from the Hidden Layer and an input to the Output Layer consisting of sigmoid or step function f_i.

Artificial neural networks have a capability of learning from data and mapping of data. These properties make neural networks a powerful tool for modelling nonlinear systems such as groundwater potential modelling. Neural networks can be trained to provide an alternative approach for problems that are difficult to solve.

One of the advantages of the application of ANNs is the tolerance to noisy data. On the other hand, one of the problems of ANNs is; like all statistical models, ANNs are subject to over-fitting when the neural network is trained to fit one set of data almost exactly (Russell and Norvig, 2003; Dunham, 2003). To minimize over fitting, the use of smaller neural networks can be recommended (Dunham, 2003).

Transfer Functions

A variety of transfer functions exist which can be used to construct an artificial neural network. In many software packages these transfer functions are already built-in and they are available to use. The most commonly used transfer functions include: Hard-Limit, Linear, Log-Sigmoid and Tan-Sigmoid transfer functions. The Hard-Limit transfer function limits the output of the neuron to either 0 if the net input argument is less than 0, or 1 if n is greater than or equal to 0. The Linear Transfer function calculates the neuron's output by simply returning the value passed to it. This neuron can be trained to learn an affine function of its inputs, or to find a linear approximation to a nonlinear function. The Sigmoid transfer function takes the input received from the neuron and reduces the value so that the output ranges from 0 to 1. The sigmoid transfer function takes the input and reduces the output into the range -1 to 1. The input value, similar to the logsig type values, can have any value between plus and minus infinity. This transfer function is commonly used in back-propagation networks.

7.7.2.2 Some Properties of ANNs

Capacity of ANNs

Artificial neural networks have a property called – "capacity", thus a capacity for neural network to model any given function can be quantified by this term. It is directly related to the amount of information that can be stored in the network (layers) and the complexity of internodes' connections.

Generalization and Over-fitting

During the development of artificial neural networks to create a system that can generalize, any possible type of example, the problem of overtraining has emerged. This problem arises due to the method in which the neural network is specified and constructed. The network as such has a significantly larger capacity than the parameters given to it, resulting in an over-specified system. In this instance the network requires more data than that which is available and must reiterate to determine its own internal variables such as the weights of the network and the bias of each node.

Over-fitting of networks may also results from "memorizing" instead of learning during the training process. In order to prevent the possibility of memorization to occur, calibration is utilized. Calibration is a parameter, which indicates that the network has trained enough thus stopping the iteration process. This can be achieved in two ways:

Calibration Based on Best Test Set

When training begins at the interval specified, the Network stops to read the test set and computes an average error for it. The error of the training set continues to decrease until it becomes flat whereas the test set error decreases to an optimal point after which it slowly increases. The Network could be saved at this optimal point based on the best test set.

Calibration Based on Minimum Error Events

Training was ordered to stop when the number of invents since minimum error for test set reaches a particular value. Calibration thus prevents over training of the network and thus reduces the training time.

Confidence Analysis of ANNs

Supervised neural networks that use a mean squared error (MSE) performance or cost function can use formal statistical methods to determine the confidence of the trained model. The MSE on a validation set can be used as an estimate of the variance. This value can then be used to calculate the confidence interval of the output of the network, if it is assumed that it follows a normal distribution. The confidence analysis of the network is statistically valid, as long as the output probability distribution stays the same and the network is not modified.

In unsupervised neural networks only inputs exist and no definition is given for the results. In these instances a comparison of the data after training is made to determine the stability of the network.

ANNs learning system

There are several types of ANNs learning system. These include supervised and unsupervised learning. In supervised learning, a known set of examples are given to the neural network and the aim is to find a function in the allowed class of functions that matches the examples as closely as possible. Thus, the model network must infer the mapping implied by the data; the cost function is used to relate the mismatch between mapping and the data. This implies that a prior knowledge of the domain area should be known. For classification and regression tasks, supervised learning is used, where the available data set consists of corresponding input and output values representing a characteristic pattern or underlying functional behaviour. This data set is the so-called training set. The adaptation of the weights is carried out by an optimization algorithm that tries to minimize the difference or error measured between the ANNs output based on the training set input values and their corresponding training set output value(s) (Govindaraju et al., 2000).

In unsupervised learning, the training of the network is entirely data-driven and no target results for the input data vectors are provided. ANNs of the unsupervised learning type, such as the self-organizing map, can be used for clustering the input data and find features inherent to the problem. Unsupervised learning allegedly involves no target values. In fact, for most varieties of unsupervised learning, the targets are the same as the inputs (Sarle, 1994). In other words, unsupervised learning usually performs the same task as an auto-associative network, compressing the information from the inputs (Deco and Obradovic, 1996).

7.7.2.3 Types of ANNs

When neural networks are trained, a particular input leads to a specific target output. Figure 45 shows a simplified illustration of back-propagation (BP) training algorithm used for training artificial neurons. In this illustration, the network is adjusted based on a comparison of the output and the expected target, until the network output matches the target. Typically many such input/target pairs are used, in the supervised learning, to train a network.

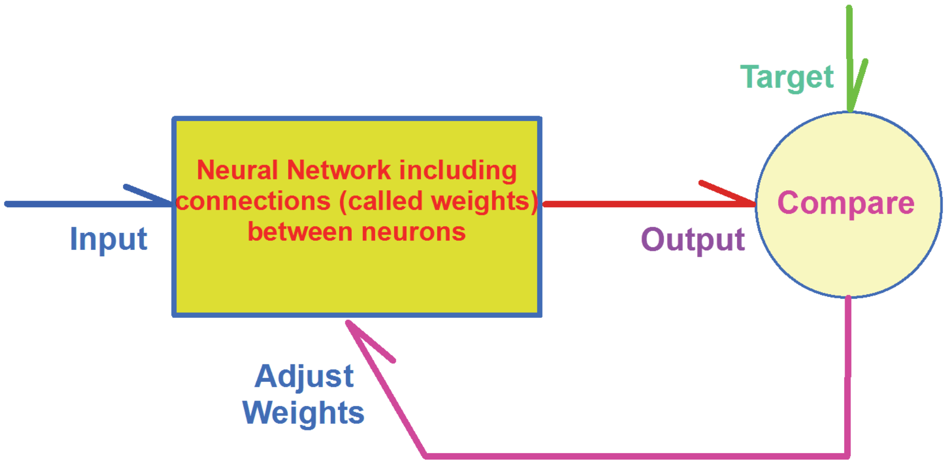


Figure 45: Simplified illustration of back-propagation training algorithm.

The network adopts as follows: change the weight by an amount proportional to the difference between the desired output and the actual output which can be expressed as:

$$\textbf{Delta Wi} = \textbf{L} * (\textbf{D}-\textbf{Y}) * \textbf{I}_i$$

Where L is the learning rate, D is the desired output, and Y is the actual output.

This relationship is known as ***Perceptron Learning Rule***. Back-Propagation (BP) which sometimes known as multi-layer perceptions (MLP) and Radial Basis Function Networks (RBFN) are both well-known developments of the Delta rule for single layer networks that is a development of the Perceptron Learning Rule. In summary, there are many types of ANNs and the actual structure of the network and the methods used to set the interconnection weights changes from one application to another. In this study RBFLN, probabilistic neural network (PNN) and fuzzy neural networks (FNN) are used for mapping groundwater potential zones.

7.7.2.4 Implementation of ANNs

It should be noted that although ANNs have great advantages, the application of neural networks are not so straightforward and a relatively good understanding of the underlying theory is required. The following three factors should be considered before using a neural network to investigate groundwater potential of an area using ANNs. An initial step is the choice of model, is there enough data available to model the application. Overly complex models tend to lead to problems with learning. Secondly the choice of learning algorithm for the neural network model plays a significant role in the success of the application. Finally, if the model, cost function and learning algorithm are selected appropriately, the resulting ANNs can be extremely robust. The method by which an artificial neural network learns or is trained can significantly affect the application and robustness of the network.

When using ANNs for mapping potential targets, the first step is the upload of training data or supporting evidence (SE). These can be the location of known high-yielding wellfields. The second is the production of thematic layers used for groundwater potential mapping, and these include geology, lineaments, slope, land use, soil amongst others. The third is the spatial crossing of casual factor maps or evidential layers in order to subdivide the study area into unique condition units (UCUs), which are “units with integer class values formed from the combination of two or more evidential maps, in which the class values represent uniquely occurring combinations of the classes of the input themes” (Kemp et al., 2001). Each UCU reflects a unique set of casual factors controlling the occurrence, or the absence of a given straining data.

The relationship between UCUs and training points or SE is established by means of self-organizing structure, in which a layer of neurons (points with which a mathematical function is associated) links feature vectors (that are strings of binary data obtained by the combination of SE dataset and UCUs), to an output map called “target vector” in which “probability” values are calculated non-linearly by progressive iterations between neurons. Iterations are controlled by interneuron connections, named synaptic weights (Haykin, 1994), which keep modifying themselves until each feature vector is mapped accurately to its target vector. For proper training of an ANNs, a non-supporting evidence data set, which represents the set of locations with the lowest probability of finding groundwater occurrence, is also required.

7.7.3 Application of ANNs to the present study

In this study, Probabilistic Neural Networks (PNN) and Radial Basis Function Link Neural Networks (RBFLN) were used to classify five evidential themes that were hypothesized to be the main factors controlling groundwater occurrence in the area. The training algorithm used is a multi-layer perception (MLP) consisting of an input layer, hidden layer and an output layer. The hidden and output layers process the inputs by multiplying each input by a corresponding weight, summing the product, and then processing the sum using a non-linear transfer function to produce a result. Each node in the hidden layer is interconnected to nodes in both the preceding and the following layers by weighted connections.

The 5 hydrogeological factors used as an input to the input vectors include: geology, lineament density, slope, land-use and soil texture (Figure 46). For spatial data modelling purposes, all the evidential themes were converted into raster with a grid cell size of 250 m by 250 m. Since there are many boreholes in the study area, the neural networks were trained using 46 boreholes with yield > 10 l/s. The training points are unevenly distributed throughout the study area. In addition, 20 barren sites with no or very poor borehole yield were randomly selected and used to train the neural networks as barren sites.

The determination of the number of nodes in the hidden layer is one of challenges in designing the architecture of artificial neural network used for predicting groundwater potential of the area. According to Shigidi et al. (2005), the number of hidden nodes depends on the number of variables, such as the size of the training dataset, the number of input and output variables, the complexity of the underlying function acting as a neuron and the activation function used. According to Shigidi et al. (2005), twice the number of input variables plus 1 is a sufficient number of hidden nodes that can predict an unknown target. In this study, the size of the training points is relatively low (46 points), and the number of input and output variables are 5 and 1, respectively. These show that simple-to-moderate architecture of ANNs training algorithm may provide optimal classification of the input layers that predict the groundwater potential of the area.

The neural networks used in this study is shown on Figure 46. It consists of input vectors L1, L2, L3, L4 and L5 with the corresponding output weights Wi1, Wi2, Wi3, …Win. Each vector consists of 20 nodes that are interconnected to the hidden layer consisting of 20 neurons or activation functions. The hidden layer comprised of Hi1, Hi2, Hi3, …Hi20 neurons with nonlinear radial basis function that are designed to step-up or step-down the weights feed into the hidden layer.

The hidden layer receives the output from the input vector and generates Weights Wo1, Wo2, Wo3,…Won, which are in-turn act as a feed-forward to the output layer. The output layer (OL) consists of a linear transfer function (a summation block) and generates the result **s**. The difference between **s** and the expected target **t** triggers the RBF to adjust the weights Wo1, Wo2, …Won, or terminate the training epoch. This process continues until the SSE and the number of iterations reaches the calibrated values. The algorithm trains the network until the minimal error between the desired and actual output values of the network achieved. Following the training, the network was used for classification of the five evidential themes showing groundwater potential zones.

The algorithm was run over number of cycles until the error stop improving to avoid the effect of over-fitting. The length of epoch required to terminate training the neurons is determined by the difference between the target **t** and the output **s.** The total sum of square of errors (SSE) was recorded repeatedly together with the training epoch. The graph of SSE versus number of iterations was plotted (Figure 47) and this was to determine the optimum number of iterations that minimize the possibility of over-fitting. Based on this approach, 20 neurons and 100 iterations were used to generate the result shown on Figure 48.

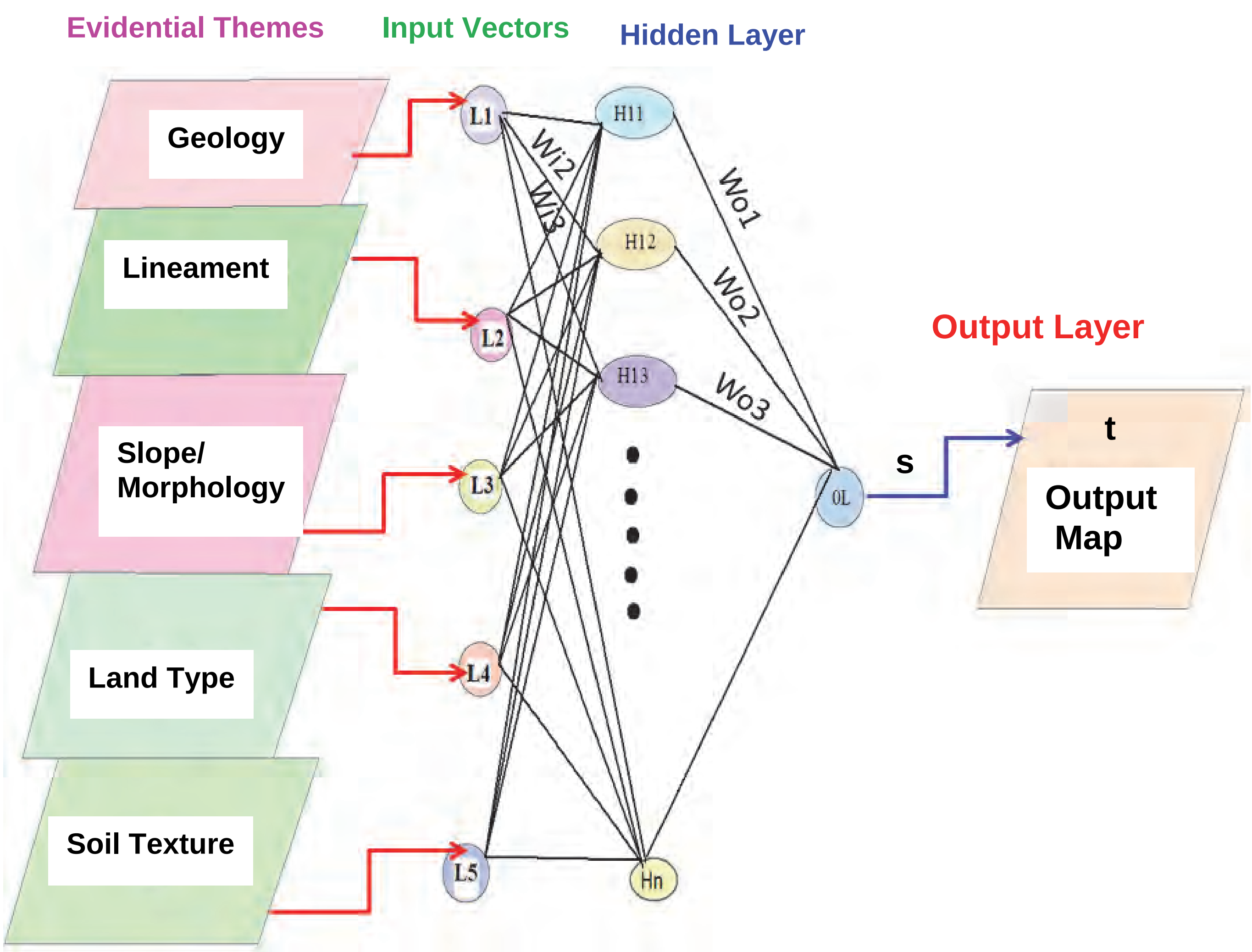


Figure 46: Simplified illustration of the architecture of ANNs training algorithm used in the present study.

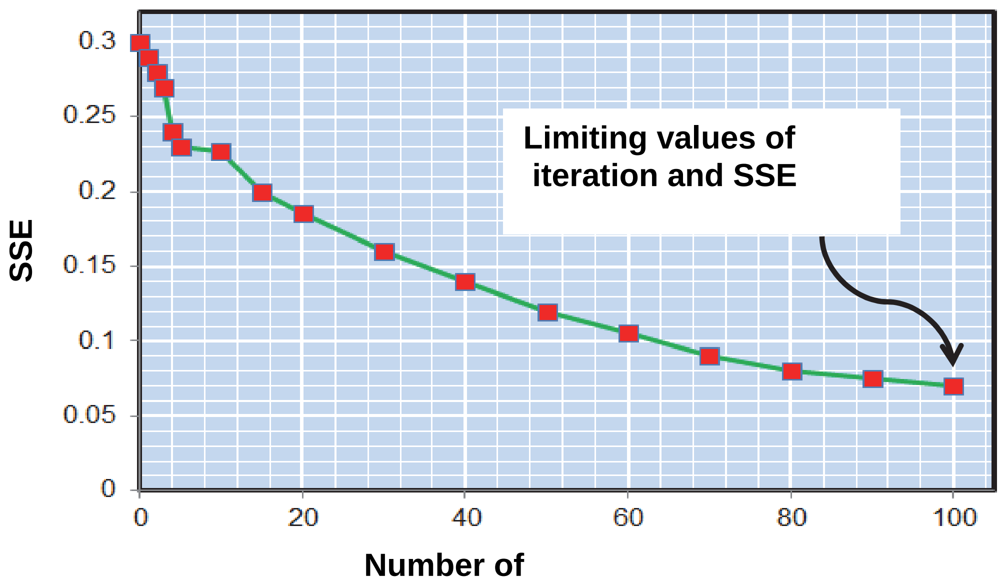
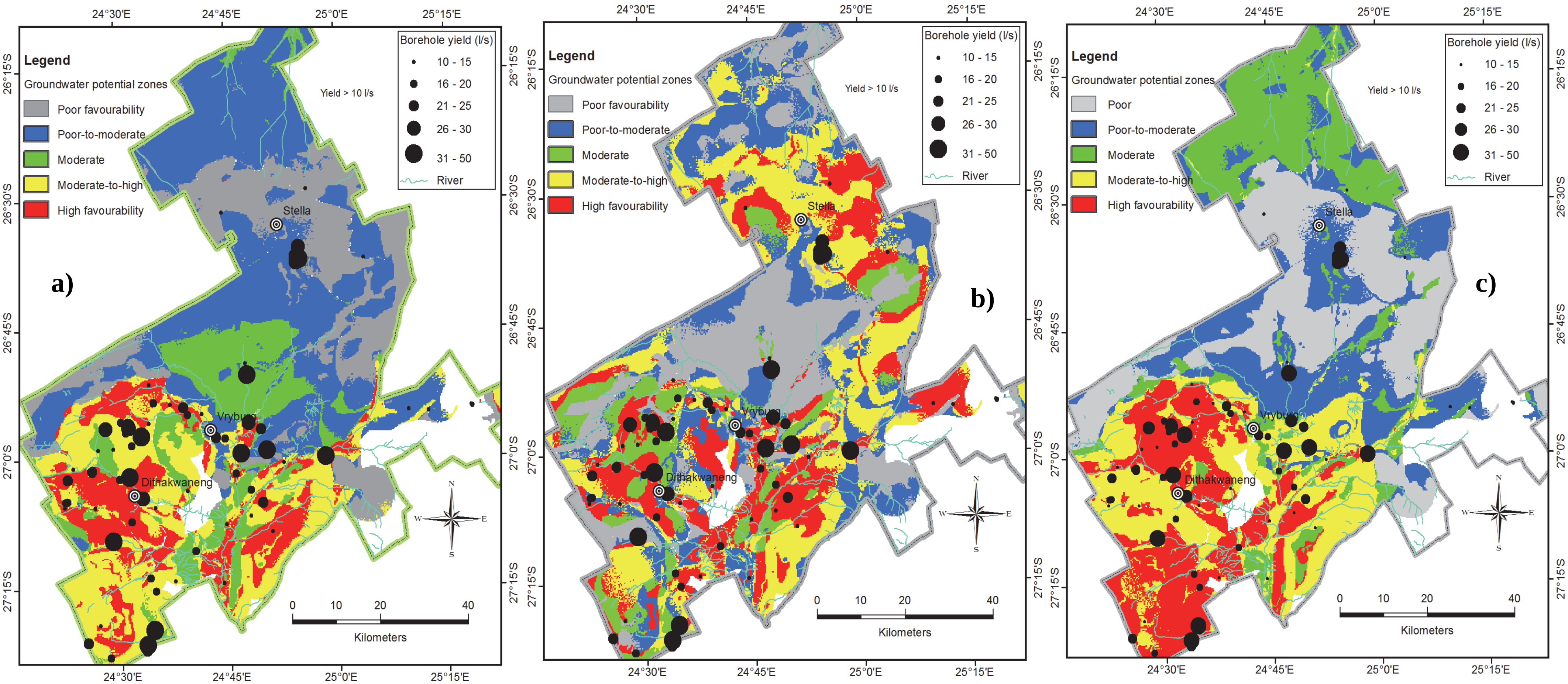


Figure 47: The performance of training session.



Figures 48: **a)**, **b)** and **c)** Groundwater potential zones derived from RBFLN, PNN and fuzzy neural network (FNN).

7.7.3.1 Sensitivity Analysis

Background

Sensitivity analysis indicates which input evidential themes are considered most important in RBFLN neural network. It can give important insights about the significance of the individual layer used for groundwater potential mapping. Sensitivity analysis often identifies evidential themes that can safely ignored in subsequent analysis, and key layers that must always be retained. However, it must be cared for reasons explained below.

Input layers are not, in general, independent – that is, there are interdependencies between evidential themes. Sensitivity analysis rates evidential themes according to the deterioration in modelling the performance that occurs if that layer is no longer available to the model. In so doing, it assigns a single rating value to each evidential theme. However, the interdependence between variables means that no scheme of single ratings per layer can ever reflect the subtlety of the true situation.

Consider, for example, the case where two input layers encode the same information (they might even be copies of the same variable). A particular model might depend wholly on one, wholly on the other, or on some arbitrary combination of them. Then sensitivity analysis produces an arbitrary relative sensitivity to them. Moreover, if either is eliminated the model may compensate adequately because the other still provides the key information. It may therefore rate the variables as of low sensitivity, even though they might encode key information. Similarly, an evidential theme that encodes relatively unimportant information, but is the only evidential theme to do so, may have higher sensitivity than any number of variables that mutually encode more important information.

Sensitivity of RBFLN used in this study

In this study, sensitivity analysis was done by alternatively excluding one evidential theme at a time from the training session and computing the sum of square of errors for the remaining 4 evidential themes. The resulting sum of square of error was compared with the SSE computed for the entire five layers used for groundwater potential mapping. This procedure was executed repeatedly for all the five evidential themes under the same initial parameters. Table 13 shows the result of sensitivity of the model in terms of the ratio of errors for each of the layer. The deterioration in the ratio of error indirectly measures the sensitivity of the RBFLN in predicting groundwater probability using the 5 evidential themes.

Table 13: Error ratio and rank for the five each of the five evidential theme.

	Geology	Lineament	Slope/Morphology	Land use	Soil
Error Ratio	2.58	1.29	1.16	0.91	1.27
Rank	1	2	4	5	3

The high error ratio corresponding to geology suggests the deterioration in error due to the lack of geological data. This shows that lithological units are significant parameters in determining the result of RBFLN used for groundwater potential mapping, which is consistent with the results of WofE and LR discussed in Section 10. Lineament and soil are also relatively sensitive parameters for characterization of groundwater potential of the area. Conversely, land use is less significant compared to the other parameters for groundwater potential mapping.

Results

Using the RBFLN and PNN, the relative importance and weight of 5 layers were calculated. The lithological units showed the highest significant factor followed by lineament density. In contrast, the land use indicates the lowest rank in controlling groundwater distribution. This result shows that lithological units are the main factors in determining the distribution of groundwater in the area, which is consistent with the distribution of high borehole yield that are mainly confined to carbonate rocks and coarse sedimentary units located in the southern part of the area. In addition, the significance of lithology is also supported by the studentized contrast derived from the WofE and regression coefficient obtained using the LR. The 3 modelling algorithms point to the presence of dissolution channels and coarse sedimentary rocks that largely control the distribution of groundwater in the area. The overall groundwater potential derived from the 3 approaches are coherent and nearly similar patterns of groundwater distribution.

8. Statistical Correlation

Analysis of the relationship between lineament-intersection frequency and borehole yield is very important when assessing the influence of secondary porosity on movement and storage of groundwater. In order to quantify the relationship between the two parameters, borehole yields were superimposed on vectorized and classified raster lineament intersection frequency, i.e. two independent layers of the same feature with a different spatial data model. The classified raster data model of the lineament-intersection frequency was used as a base map to quantify the relationship between the two parameters in each class. The cases in which the values of borehole yield and lineament-intersection frequency fall close to each other were manually sampled onto a table corresponding to each of the six classes. How well does the lineament-intersection frequency within the area of study determine borehole yield? In order to answer this question, correlation coefficients for each of the six classes were calculated using the following equation:

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(Y_i - M_Y\right)\left(LC_i - M_{LC}\right)}{\sigma_Y \sigma_{LC}}$$

Where R is correlation coefficient, N is total number of data, Y_i and LCi are borehole yield and lineament-intersection frequency, respectively, M_Y and M_{LC} are the statistical means of the two parameters, and σ_Y and σ_{LC} are the standard deviation for borehole yield and lineament-intersection frequency, respectively.

Prior to calculation of statistical correlation between lineaments and borehole yield, the geology of the study area was subdivided into 4 sub-groups based on broad hydrogeological properties. The basement rocks consisting of Gold Ridge Formation, Archaean granite intrusives and Ventersdorp mafic lavas were treated as one category while the Transvaal carbonates (limestone and dolomite) were considered separately from the rest of other rocks. The Karoo glacial diamictite were grouped as separate from Kalahari windblown sand.

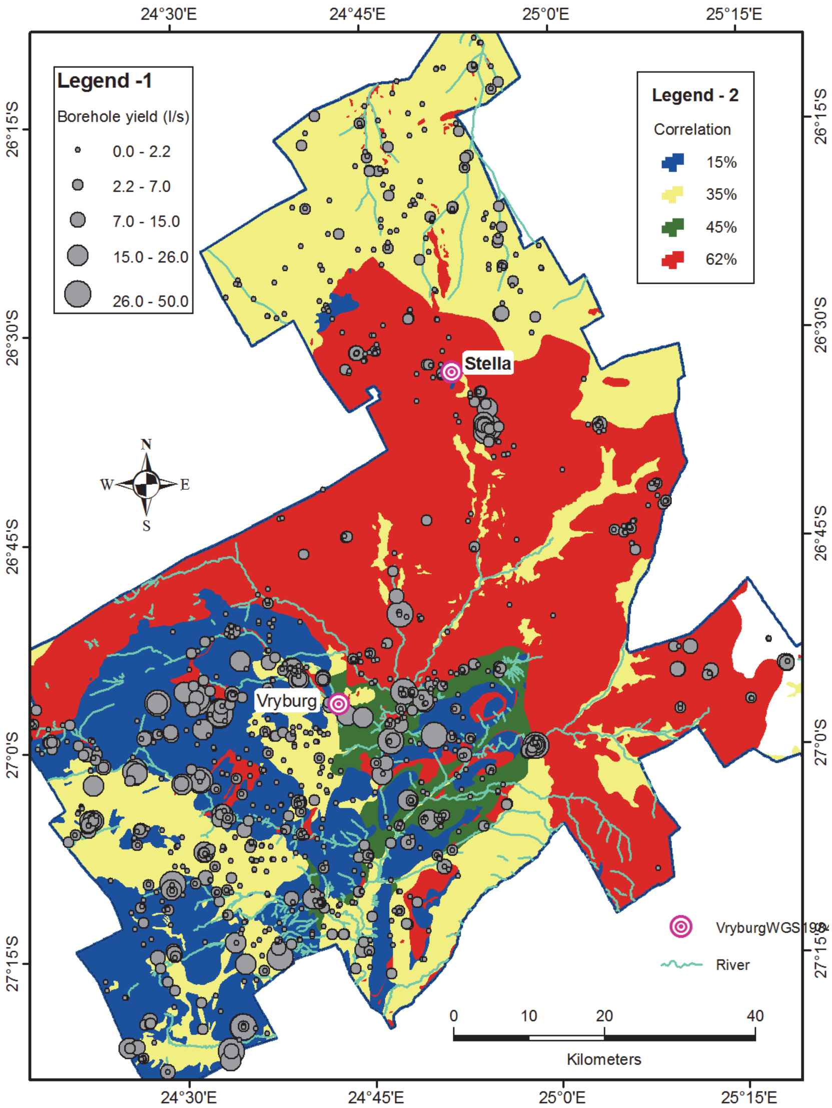


Figure 49: Result of statistical correlation between borehole yield and lineament-intersection frequency superimposed on geological map of the area.

Table 14: Summary of results of statistical correlation for each of 4 sub-region.

Sub-region	Total number of borehole with yield > 3 l/s	Correlation
Quaternary Kalahari sand	50	35%
Karoo glacial diamictite	44	45%
Transvaal Carbonates (limestone and dolomite)	138	15%
Ventersdorp mafic lavas	58	62%

Statistical correlation between lineament intersection-frequency and borehole yield was computed for four sub-regions (Figure 49 and Table 14). A total of 1988 boreholes were used out of which 1215 have recorded measured yield. The total number of boreholes throughout the study area with yield >3.0 l/s is 300 and the number of these boreholes with yield > 3.0 l/s falling in each sub-region were counted and used as a parameter for comparison with the results of statistical correlation (Table 14).

9. Discussion of Results

The quality of groundwater in Naledi Local Municipality has been assessed based on analysis of hydrochemical and environmental isotopes data. In addition, multivariate statistical approaches were used to identify the most prospective sites for groundwater occurrence. Furthermore, statistical methods were applied to 5 evidential themes and the results were used to investigate the control of groundwater distribution in the area. Below is the summary of discussion of the most important findings.

9.1 Groundwater Quality

The quality of groundwater in the study area was assessed using major cations and anions. In addition, selected dissolved materials and physical parameters that include; fluoride, nitrate, total alkalinity (total hardness), total dissolved solids, EC, pH and sodium adsorption ratio (SAR) were used to investigate the quality of groundwater in the area.

Borehole inventory involving analysis of the concentration of chloride, fluoride, pH, EC, alkalinity and nitrate concentrations were carried out. The results show the water quality throughout the study area is fresh, with the exception of some boreholes located around Stella. In this area, number of boreholes show elevated concentration of chloride, fluoride, nitrate, TDS and EC relative to the standard values prescribed by SANS241 (SANS, 2011). In addition, several boreholes located at Armoedsvlakte, Swartfontein and Biesjiesvlakte wellfields show elevated chloride, TDS, and fluoride content. These boreholes can be recommended to be abandoned from domestic water use until the concentration of the dissolved materials is verified based on standard laboratory analysis. The boreholes located around Dithakwaneng and Huhudi are characterized by elevated alkalinity (very high total hardness), though this may not be a threat to human health, but can affect the taste of drinking water quality and reduce the effectiveness of detergents. The boreholes that are located at Stella wellfield and north of Vryburg may not be recommended for irrigation owing to elevated concentration of SAR. In addition, sodium concentration also increasingly high within the same wellfield which may pose severe health effects to infants.

The analysis of major cations and anions shows the water type in the study area is generally dominated by Ca-Mg-HCO_3 type, which can be interpreted as fresh, relatively shallow and recently recharge groundwater with the exception of some boreholes. The exceptions are the boreholes that exhibit enrichments in Mg and Cl resulting from water-rock interactions and rapid rate of evaporation. Also, the aridity of the climate partly contributed to an elevated chloride content. These boreholes are located around Stella and north of Vryburg. In general, the analysis of the hydrochemical data shows the presence of fresh water in many parts of the study area.

The stable isotopic composition of groundwater (σ^2H and $\delta^{18}O$) in the study area suggests the precipitation that recharged groundwater have extensively undergone evaporation, which is consistent with elevated concentration of chloride content in number of boreholes that are located around Stella. However, the limited environmental isotopes data make difficult to understand the interaction between groundwater and surface water.

9.2 Groundwater Potential

Multivariate statistical approaches involving WofE, ANNs, PCA, LG and fuzzy logic were applied to 5 evidential themes. Several statistical results were generated using each of the five modelling algorithm. The results allowed in identifying the factors that control the distribution of groundwater in the area. The WofE, PCA, ANNs and LG provided nearly similar results showing lithologic units and fracture concentration are the main controlling factors. The influence of the lithology is mainly dominant in the southern part of the area, while the significance of fracture concentration is mainly shown within the hard rocks terrane consisting of the Ventersdorp volcanic rocks, Archaean granite gneisses and the greenstone belt. The wellfields which fall under this category include Stella, Middlekop, Spitzkop and the central part of the area. In addition, the WofE, PCA and LR show the significance of gravel deposits south of Stella and Vryburg.

The success rate curve was plotted and used to classify, the posteriori probability map derived using the WofE, LR, RBFLN and PNN. Based on this approach, 3 class-breaks were identified which allowed classification of the results. In addition, borehole yields were classified into five and superimposed on the classified posteriori probability image derived from WofE, LR, RBFLN, PNN and fuzzy logic. The resulting map shows a number of groundwater potential zones varying from “*very good*” to “*very poor*” (Figures 37, 38a, 40, 41, and 48). The maps show that the southern part of the study area is characterized by good groundwater potential, though some isolated patches of “*very good*” coincide with major river channels located north and northeast of Vryburg. The zone shown as ‘*very good’* and ‘*good’* groundwater potential covers ~17% and ~22% of the study area, respectively. The superimposed borehole yield also confirms the results derived from multivariate statistical modelling approaches, whereby high borehole yields (> 15 l/s) fall within carbonate rocks consisting of dolomite and limestone. In addition, follow-up geophysical surveys carried out at selected sites confirm the presence of conductive layers varying in depth from 20 m to 35 m (Appendix A). The high conductivity possibly indicates the presence of water-bearing formation, in particular dolomite north of Vryburg and highly fractured granite just south of Stella. The high yielding wellfields can be attributed to dissolution of carbonate rocks by water that percolates through pre-existing fractures leading to enlarged fracture apertures, and consequently resulting in the development of large cavities that can store and supply significant amount of water (e.g. Armoedsvlakte and Swartfontein wellfields).

The development of pre-existing fracture zones within Transvaal carbonate rocks and Ventersdorp lava has been addressed by several authors (e.g. Anahaeuser and Walraven, 1999; Good and de Wit, 1997). According to Anahaeuser and Walraven (1999), during the Archaean age the Kraaipan greenstone belt was subjected to lateral compressional tectonic forces, which was accompanied by the development of several faults and fracture zones (e.g. TML fault belt) that extend over tens and hundreds of kilometers. Dietvorst (1988) also noted that the rocks of the Ventersdorp and Transvaal Supergroups in the western part of the Kaapvaal craton were block faulted during ‘*cratonic updoming*’ at ~2.1 billion years ago. As a result of these processes most of the rocks were deformed resulting from extensional and compressional forces. This observation is consistent with the results of Bird et al. (2006) which show the distribution and patterns of stress and strain rate in the Northern Province including Southern Africa. According to these authors, the area has been subjected to NW-SE directed compressive horizontal principal stress which generated series of fault systems ranging from strike-slip to normal faults. These structures may have important implications in terms of controlling groundwater recharge and discharge zones, particularly within the hard rocks terrane of the present study area.

Surface water or the mildly acidic rain water (relatively low pH) percolates along the structural features (i.e. fractures and faults) and dissolves the underlying carbonate rocks if any. This suggests that pre-existing structures within carbonate rocks as well as the hard rocks played a significant role in the development of high yielding wellfields in the southern and central part of the area as shown from the results of multivariate statistical analysis. In addition, younger and coarse sedimentary rocks of the Karoo Supergroup and Kalahari sand were deposited atop the Transvaal carbonate rocks which enhanced the seepage of rainwater. Areas that are shown as '*very good*' groundwater potential therefore suggest the presence of deeply penetrating fault systems (e.g. TML) or the presence of coarse sedimentary units overlying the carbonate rocks that act as pathways for rainwater to replenish the underlying karstic aquifers.

About 50% of the total number of boreholes with yield >3.0 l/s fall within the southern part of the study area where correlation between borehole yield and lineament intersection-frequency ranges from 10% to 20%. The influence of lithology (dissolution cavities in particular) can also be explained in terms of the results of principal component analysis that shows maximum data variance in principal component 1 corresponds to lithology compared to other thematic layers. In addition, the WofE and LG modelling indicate maximum weights contrast and high regression coefficient corresponding to lithology (the carbonate rocks in particular). Furthermore, the results of RBFLN, PNN and fuzzy logic show high favorability for groundwater occurrence within carbonate rocks. Also, the results model sensitivity analysis of RBFLN show the significance of rock types and lineaments in controlling localization of groundwater occurrence. These suggest that in the southern part of the study area groundwater occurs in areas where the carbonate rocks were initially subjected to lateral compressional forces prior to the development of dissolution channels. This argument is consistent with results from similar groundwater investigation in southeastern Botswana that shows the significance of rock types and pre-existing structural fabric in controlling high well yield (Dietvorst et al., 1991). According to these authors, productive wellfields are situated in areas where geological structures transect the carbonate rocks, thus the zone of structural intersections within carbonate rocks represent one of Botswana's largest groundwater reservoirs with borehole yields up to 25 l/s.

The Ventersdorp volcanic rocks, Kraaipan greenstone and the surrounding granite gneisses largely cover the central part of the study area, and 20% of the total number of boreholes with yield >3.0 l/s fall in this region (Figure 49). The borehole yield throughout these rocks spatially varies over a wide range which can be attributed to many factors that include fracture density, fracture aperture, fracture size, fracture connectivity, mineralogy of the rocks, degree and depth of weathering, hydraulic properties of the overlying rocks/soil and proximity to river channels. For example, the boreholes located just south of Stella have a moderately high yield due to the intersection of the NE-SW and NW-SE striking lineaments (Figures 3d, 37 and 49). The influence of the density of lineament can also be explained in terms of the results of statistical correlation which shows that 62% of high borehole yield within the crystalline basement rocks are attributed to the density of fracture network (Figure 49). This is consistent with the previous studies elsewhere north of the present study area where cross-cutting structures significantly contributed to the productivity of boreholes that were drilled within the Ventersdorp lava (Tessema et al., 2012). There are several NE-SW and NW-SE striking lineaments that transect the Ventersdorp lava which contributed to the productivity of wellfields. These structures are the result of multiple episodes of deformation of Archaean basement rocks which resulted internal loss of cohesion.

Previous borehole loggings conducted within the Ventersdorp volcanic rocks revealed the presence of thin (0.2-4 m) sandy-regolith and saprolite layers overlying the fractured bedrock (Haddon, 2005). The thinning of the weathered horizon is possibly attributed to the arid-to-semiarid nature of the climate whereby the rate of weathering process is slower than denudation. In addition, several lines of evidence suggest that subsidence of the Kalahari Basin was closely associated with uplift at ~126 Ma, followed by major erosional events during early Cretaceous (e.g. Moore et al., 2009; Moore, 1999). Also, Patridge and Maud (1987) and Patridge (1998) postulated minor uplift along the Transvaal-Griqualand Axis (TGA) which occurred during early Miocene, followed by extensive erosion. This suggests that the presence of thin weathered layer within crystalline basement rocks of the study area attributed to extensive denudation and little preservation of the overlying unconsolidated weathering products. Location of high yielding wellfields within the hard rocks of the study area therefore depends on the availability of dense and interconnected cross-cutting structures.

The ability of fractures to act as conduits for groundwater flow can be determined by the degree to which the fractures are interconnected, thus fracture connectivity increases with increasing fracture length and fracture density, hence fracture intersection increases (Cook, 2003). According to this author, the hydraulic conductivity of a fractured crystalline rock is a function of fracture density whereby hydraulic conductivity increases proportional to an increase in the cubic power of fracture aperture. Figure 30c shows that the maximum concentration of fracture density occurs within the Ventersdorp rocks compared to elsewhere in the study area. This is consistent with the results of statistical correlation that shows high borehole yields coinciding with areas that exhibit high concentration of fissures, fractures and joints (Figure 49). The intersection of the NE-SW and NW-SE striking lineaments located south of Stella town is a typical example of where high borehole yield presumably enhanced by fracture connectivity. However, the high borehole yield around Stella town (Middlekop, Spitzkop and Stella wellfields) is not shown-up on the multivariate statistical modelling results of PCA, fuzzy logic, FNN and RBFLN (Figures 37, 41, 48a and 48b). This is probably due to the gridding parameters used for converting the five layers from vector to raster format, i.e. grid cell size of 250 m by 250 m is possibly greater than the spatial extent of the north-south striking fracture-controlled wellfields located just south of Stella. The modelling algorithms used for mapping groundwater potential, thus cannot uniquely resolve the presence of features with width less than 250 m.

Fifteen percent of the total number of boreholes with yield >3.0 l/s fall within the Kalahari Group. However, statistical correlation between borehole yield and lineament density in the Kalahari Group does not show the influence of lineaments on borehole productivity. The sites with high borehole yield (>3.0 l/s) are located close to flood plains or river channels (Figures 38 and 40), suggesting that lineament density and sedimentological properties (grain size, roundness of grains, grain sorting, grain shape and grain packing) are less important in controlling groundwater development compared to other factors. Previous borehole logging (Haddon, 2005) shows that the overlying Kalahari Group consists of clay-dominated sand, which absorbs large volume of rainwater, but poorly transmits through the unsaturated zone leading to loss of water by evaporation before it percolates and recharges the underlying permeable units. This is consistent with the sparse distribution of drainage lines within the Kalahari Group, indicating the influence of the underlying rocks in partitioning rainwater into surface runoff and infiltration. As suggested by Salman (1983), the stream pattern is a reflection of the rate at which precipitation infiltrates compared to runoff. The sparse drainage lines within the Kalahari Group therefore suggest the influence of clay-dominated overlying sedimentary unit.

Although high groundwater potential zones are primarily associated with carbonate rocks, the results derived from calculation of statistical correlation between lineaments density and borehole yield suggests that high yielding wellfields are largely controlled by fracture concentration within the hard rocks. The significant contrast in groundwater potential between the southern and northern parts of the area can be attributed to the presence of two end-members controlling the development of groundwater, i.e. dissolution channels in the south and fracture connectivity and density in the central and northern parts of the area.

9.3 Conclusion and Recommendation

The assessment of groundwater quality and quantity is a major step towards ensuring the sustainable use and management of one of the most basic needs of human being. Understanding the amount of groundwater resource and its quality assists in creating awareness amongst decision makers to use, manage and to protect groundwater without adversely affecting its future demand. In this study various datasets were used to assess the groundwater potential of the area. Hydrochemical and environmental isotopes data were used to evaluate the quality of groundwater and surface water. In addition, multivariate statistical techniques were applied to 5 evidential themes that were hypothesized to be the factors that control groundwater occurrence in the area. Based on integration of the results of analysis of hydrochemical, isotopes and multivariate statistics, the following conclusions are drawn:

The quality of groundwater in the study area is ideal for domestic and agricultural use with the exception of boreholes located around Stella and north of Vryburg. The boreholes identified as exceptions can be recommended to be abandoned until verified for the concentrations of dissolved materials.

The multivariate statistical approaches used in this study provided valuable information about groundwater potential zones. The results show differences in fracture set populations, coupled with contrast in primary and secondary permeability of various rock types significantly contributed to variability in groundwater potential throughout the study area. The most probable groundwater potential zones cover ~17% of the entire study area and mainly associated with carbonate rocks located in the southern part of the area.

The presence of pre-existing structures together with younger and coarse sedimentary rocks deposited atop the Transvaal carbonate rocks played a significant role in the development of karstic aquifers. This suggests that mild acidic rainwater percolates through the unsaturated zone and dissolve the carbonate rocks along pre-existing fractures leading to the development of large dissolution cavities that supply significant volume of water. In addition, the localization of high groundwater potential within the carbonate rocks is consistent with the results of principal component analysis that shows the dominance of lithology in data variance corresponding to principal component 1. Furthermore, the results of WofE and LG are in agreement with the PCA showing rock types and crosscutting structures played significant role in controlling groundwater distribution in the area.

The zone of high groundwater potential within the hard rock terrane coincides with areas where there are maximum concentration and connectivity of fractures, joints and fissures. In addition, the results of statistical correlation provide supporting evidence about the significance of crosscutting structures in controlling groundwater distribution.

Lithology and structures are less important compared to other factors in controlling groundwater development within the Kalahari Group. This is due to the presence of unconsolidated overlying Kalahari Group consisting of clay-dominated sand, which absorbs large volume of rainwater, but poorly transmits through the unsaturated zone leading to loss of water through evaporation before rainwater recharges the underlying permeable units

The multivariate groundwater potential modelling and statistical correlation used in this study provided valuable information about hydrogeological parameters that control the development of groundwater in the area. These approaches are very effective and can therefore be used as a sound scientific basis for understanding the control of groundwater occurrence in similar hydrogeological settings.

Routine monitoring of water level and abstraction volumes on a monthly basis can be recommended for managing the sustainable use and groundwater protection. In addition, monitoring of rainfall on a daily basis at known meteorological stations and analysis of water quality bi-annually strongly recommended to monitor temporal variations of water quality. In addition, detailed environmental isotopes study is very important and strongly recommended in order to understand the interaction between groundwater and surface water. This ensures planning for the sustainable use and protection of groundwater resource in the area based on the inference drawn from environmental isotope study.

REFERENCES

Adams, S., Titus R., Pietersen, K., Tredoux, G., Harris, C., 2001. Hydrochemical characteristics of aquifers near Sutherland in the Western Karoo, South Africa. Journal of hydrology, 241, 91-103.

Agterberg, F.P., 1992. Combining indicator patterns in weights of evidence modelling for resource evaluation. Nonrenewable Resources, 1, 39-50.

Agterberg, F.P. and Cheng, Q., 2002. Conditional independence test for weights of evidence model. Natural Resource Research, 11, 249-255.

Ajith A., 2005. "Artificial Neural Networks Oklahoma State" University, Stillwater, OK, USA. Handbook of Measuring System Design, John Wiley & Sons, Ltd.

Anbazhagan, S., Ramasamy, S. M. and Gupta. D.S., 2005. Remote sensing and GIS for artificial recharge study, runoff and planning in Ayyar basin, Tamil Nadu, India. Environmental Geology 48, 158-170.

Anhaeuser, C.R. and Walraven, F., 1999. Episodic granitoid emplacement in the western Kaapvaal Craton: evidence from the Archaean Kraaipan granite-greenstone terrane, South Africa. J. Afr. Earth Sci., 28, 289-309.

Banks, D. and Robinson, N., 2002. An introduction to groundwater in crystalline bedrock. Geological Survey of Norway, Trondheim, Norway.

Bartolino, J.R. and Cole, J.C., 2003. Groundwater resources of the middle Rio Grande Basin. US Geological Survey Circular, 1222.

Bird, P., Ben-Avraham, Z., Schubert, G., Andreoli, M., and Viola, G., 2006. Patterns of stress and strain rate in southern Africa. J. Geophysical Res., 111, 1-14.

Bonham-Carter, G.F, 1996. Geographical Information Systems for Geoscientists, Modelling with GIS. Elsevier Science Inc., TARRYTOWN, NY.

Brandl, G., Cloete, M. and Anhaeuser, C.R., 2006. Archaean greenstone belts. In: Jonson, M.R., Anhaeuser, C.R. and Thomas, R.J. (Eds.). The Geology of South Africa. Geological Society of South Africa, Johannesburg/Council for Geoscience, Pretoria, 9-56.

Burke, K., Kidd, W.S.F., Kusky, T.M., 1985. Is the Ventersdorp System of Southern Africa related to continental collision between the Kaapvaal and Zimbabwe cratons at 2.64 Ga ago? Tectonophysics, 115, 1-24.

Burrough, P.A., and McDonnell, R.A., 1998. Principles of geographical information systems, New York, Oxford University Press, 333p.

Chilton, P.J., Foster, S.S.D., 1995. Hydrological characterization and water supply potential of crystalline aquifers in tropical Africa. Hydrogeology Journal 3, 36-49.

Clark, I.D., Fritz, P., 1997. Environmental isotope in hydrology. Lewis, New York, 328.

Cook, P.G., 2003. A guide to regional groundwater flow in fractured rock aquifer. Printed by CM Digital, 234 Currie Street, Adelaide, South Australia, pp 20-35.

Crsini, A., Cervi, F., and Ronchetti, F., 2009. Weights of Evidence and artificial neural networks for potential groundwater spring mapping: an application to the Mt. Modino area, Northern Apennines, Italy. Geomorphology, 111, 79-87.

Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus, 16:436-468.

Deco, G. and Obradovic, D., 1996. An Information-Theoretic Approach to Neural Computing. NY: Springer-Verlag.

Department of Water Affairs and Forestry, 1996. South African Water Quality Guidelines (2nd Edition) Volume 1: Domestic Use.

De Vries, J.J., Simmers, I., 2002. Groundwater recharge: an overview of processes and challenges. Hydrogeology Journal 10, 5-17.

Dietvorst, E.J.L., De Vries, J.J., Gieske, A.G., 1991. Coincidence of well fields and tectonic basins in the Precambrian shield area of southeast Botswana. Ground Water, 29, 869-877.

Dietvorst, E.J.L., 1988. Early Precambrian basement uplift and block faulting along the western margin of the Bushveld Complex, south eastern Botswana. Journal of African Earth Science 7,641-651.

Drury, S.A., 1987. Image Interpretation in Geology. (London: Allen & Unwin), 243pp.

Van Dyk T., 1993. Geohydrological investigation of SADF, Vryburg. Unpublished DWAF Report, 53P, GH3813.

Dunham, M. H., 2003. Data Mining Introductory and Advanced Topics. Prentice-Hall.

Edet A.E., Okereke C.S. Teme, S.C., Esu E.O., 1998. Application of remote sensing data to groundwater exploration: a case study of the Cross River State, South eastern Nigeria. Hydrogeology J. 6: 394-404.

Edmunds, W., Darling, D., Kinniburgh, K.S., Mahgoub, S., 1992. Source of recharge at Abu Deliag, Sudan. J. Hydrology, 13, 1-24.

Fabbri,A.G., Chung, C.J., 2008. On blind tests and spatial prediction model. Natural Resource Research, 17 107-118.

Farquharson, F.A. and Bullock, A., 1992. The hydrogeology of basement complex regions of Africa with particular reference to southern Africa. . *In*: Wright, E.P. and Burgess W.G. (Ed) The hydrogeology of crystalline basement aquifers in Africa. The Geological Society Special Publication, 59-78.

Fitterman, D. V., and Stewart, M. T., 1986. Transient electromagnetic sounding for groundwater, Geophysics, 51(4), 995-1005.

Freeze, R.A. and Cherry, J.A., 1979. Groundwater. Prentice Hall Englewood Cliffs, NJ.

Gat, J.R. (1996). Oxygen and hydrogen isotopes in the hydrologic cycle. Annu. .Rev. EarthPlanet. Science, 24, 225-262.

Gat, J.R. and Galai, A., 1982. Groundwater of the Arava valley: an isotopic study of their origin and interrelationship. Israel J. Earth Sci., 31, 25-38.

Gong, H., Pan, Y. and Xu, Y., 2012. Spatio-temporal variation of groundwater recharge in response to variability in precipitation, land use and soil in Yanqing Basin, Beijing, China. Hydrogeology J., 20: 131-1340.

Good N. and De Wit M.J., 1997. The Thabazimbi-Murchison Lineament of the Kaapvaal Craton, South Africa: 2700 Ma of episodic deformation. J. Geological Society London, 154: 93-97.

Govindaraju R. S., 2000. Artificial neural network in hydrology. Journal of Hydrologic Engineering, 5(2), 124-137.

Greenbaum, D. 1987. Lineament Studies in Masvingo Province, Zimbabwe, British Geological Survey Report WC/87/7.

Greenbaum, D. 1992. Structural influences on the occurrence of groundwater in SE Zimbabwe. In: Wright, E. P. & Burgess, W. G. (eds) 1992, Hydrogeology of Crystalline Basement Aquifers in Africa, Geological Society Special Publication No 66, pp. 77-85.

Gupta, S.K. and Deshpande, 2005. Groundwater isotopic investigations in India: What has been learned? Current Science, 89, 825-835.

Haddon, I.G., 2005. The sub-Kalahari geology and tectonic evolution of the Kalahari Basin in southern Africa. Unpublished PhD thesis, University of the Witwatersrand, South Africa, 343p.

Han, J., Kamber, M., 2001. Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers.

Hem, J.D. (1989). Study and interpretation of the chemical characteristics of natural water. US Geological Survey Water Supply Paper, 2254, 263

Holland, M. and Withűser, K. T. 2011. Evaluation of geologic and geomorphologic influences of borehole productivity in crystalline bedrock aquifers of Limpopo Province, South Africa. Hydrogeology J. 19:1065-1083.

Hola J. and Schabowicz K., 2005. Application of Artificial Neural Networks to determine concrete compressive strength based on non-destructive test. Journal of civil engineering and management, 11 (1), 23-32.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, 2nd Edition. John Wiley and Sons Inc., 375, New York, 375p.

Haykin, S., 1994. Neural networks: A comprehensive foundation, 1st ed. Prentice Hall, New Jersey, Pearson Education.

Jain, S. K., Singh, V. P., and Van Genuchten, M.T. 2004. "Analysis of soil water retention data using artificial neural networks". Journal Hydrologic Engineering. 9(5), 415-420.

Jeng D. and Cha H. 2003. Application of Neural Network in Civil Engineering, school of Engineering, Griffith University Gold Coast Campus. Australia.

Kendall, C., McDonnel, J., 1998. Isotope tracers in catchment hydrology. Elsevier Science, Amsterdam, pp 51-86.

Kemp, L.D., Bonham-Carter, G.F., Raines, G.L., Looney, C.G., 2001. ArcSDM: ArcView Extension for Spatial Data Modelling using Weighs of Evidence, Logistic Regression, Fuzzy Logic and Neural Network Analysis. http://ntsserv.gis.brcan.gc.ca/sdm/

Kleinbaum, D.G. and Klein, M., 2002. Logistic Regression a self-learning Text. 2nd Edition. Springer-Verlag, New York, 513p.

Koch, M., and Mather, P.M., 1997. Lineament mapping for groundwater Resource assessment: a comparison of digital Synthetic Aperture (SAR) imagery and stereoscopic Large Format Camera (LFC) photographs in the Red Sea Hills, Sudan. Int. J. Remote Sensing, 18, pp. 1465-1482.

Lattman, L.H. 1958. Technique of mapping Geologic Fracture Traces and Lineaments on Aerial Photographs, Photogrammetric Engineering, 19, 4, pp.568-576.

Lattman, L.H. and Parizek, R.R. 1964. Relationship between fracture traces and the occurrence of groundwater in carbonate rocks. Jour. Of Hydrology, 2, pp. 73-91.

Leduc C., Favreau G., Schroeter P., 2001 Long-term rise in a Sahelian water table: the continental terminal in southwest Niger. J. Hydrol 243: 43-54.

Lee S., Cho S. and Wong M., 1998. "Rainfall Prediction Using Artificial Neural Networks". Journal of Geographic Information and Decision Analysis 2 (2), 233-242.

Mabee, S.B., Hardcastle, K.C., and Wise, D.W., 1994. A Method of Collecting and Analyzing Lineaments for Regional-Scale Fractured Bedrock Aquifer Studies, Ground Water, 32, 6, pp.884-894.

Mabee, S.B., 1999. Factors influencing well productivities in glacial metamorphic rocks. Groundwater 37:88-97.

Mabee, SB, Curry, P.J., Hardcastle, K.C., 2002. Correlation of lineaments to groundwater inflows in a bedrock aquifer tunnel. Groundwater, 32 (6): 884-894.

Masters, T., 1993. Neural network recipes in C^{++}: New York, Academic Press, 285p.

Masters, T., 1995. Advanced algorithms for neural networks: a C++ sourcebook: John Wiley and Sons, P431.

McCulloch, W.S., and Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity: Bulletin of Mathematical Biophysics, 5, 115-133.

Menrad, S., 1995. Applied logistic regression analysis. Sage University Paper Series on Quantitative Applications in Social Sciences, 106, 98p.

Moore, A.E., Blenkinsop, T., Cotterill, F., 2009. Southern African topography and erosion history: plumes or plate tectonics? Terra Nova 21, 310-315.

Moore, A.E., 1999. A reappraisal of epirogenic flexure axes in southern Africa. South African Journal Geology 102(4), 363-376.

Motebe, N.S., 2003. Hydrogeological chapter of the 1:250,000 Vryburg (2624) Geological Map. Unpublished CGS Technical report, 33p.

Muqtda, A., Rahid, U., Azizu, B., Habibah, L and Anton, A. K. 2012. Evaluation of groundwater quality using regression model. J. Applied Science Research, 8(1), 251-260.

Nel, J.M. and Wyk, E. van, 1999. Preliminary assessment of the groundwater potential of the Middlekop/Appleby Aquifer, Stella District, North Western Region. Unpublished DWA technical report. P43.

Nel, J.M., 2001. Assessment of the groundwater potential of the Middle Kop/Appleby aquifer, Stella District, North West Province. Unpublished M.Sc thesis, University of Free State, 78p.

Nelles, O., 2000. Nonlinear System Identification. Springer-Verlag.

Pascal, G., Philip, D., Etienne, B., Jean-Raynald, D. and Tanguy, B., 2013. Partitioning a regional groundwater flow system into shallow and deep regional flow compartments. Water Resource Research, 49, 1-13.

Patridge, T.C., Maud, R.R., 1987. Geomorphic evolution of southern Africa since the Mesozoic. South African Journal of Geology, 90, 179-208.

Patridge, T.C., 1998. Of diamonds, dinosaurs and diastrophism. 150 million years of landscape evolution in southern Africa. 25th Alex Du Toit Memorial Lecture. South African Journal of Geology 101, 167-184.

Poli, S. and Sterlacchini, S., 2007. Landslide Representation Strategies in susceptibility studies using Weights of Evidence modelling Technique. Natural Resource Research, 16, 121-134.

Roering, C., Barton, J.M. and Winter, H. de la R., 1990. The Vredefort structure: perspective with regard to new tectonic data from adjoining terranes. Tectonophysics, 171, 7-22.

Russell, S., Norvig, P., 2003. Artificial Intelligence A Modern Approach. Prentice-Hall.

Salman, A.A.B., 1983. Using Landsat imagery interpretation for underground water prospection around Qena Province, Egypt. Int. J. Remote sensing 4:179-189.

Sander, P., Chesley, M.M., Minor, T.B., 1996. Groundwater assessment using remote sensing and GIS In a rural groundwater project in Ghana: Lessons learned. Hydrogeology Journal 4(3), 40-49.

Schulze, R.E., 2003. On land Cover, land use and their impacts on hydrological responses. In: Schulze, R.E. (Ed) Modelling as a Tool in Integrated Water Resources Management: Conceptual Issues and Case Study Applications. Water Research Commission, Pretoria, RSA, WRC Report 749/1/02, Chapter 4, 84-97.

Sener, E., Davraz, A., Ozcelik, M., 2005. An integration of GIS and remote sensing in groundwater investigations: A case study in Burdur Turkey. Hydrogeology Journal 13,826-834.

Shigidi, A. and Garcia, L.A., 2005. Parameter Estimation in Groundwater Hydrology Using Artificial Neural Networks. Journal of Computing in Civil Engineering, 31, 281-289.

Sklash, M., 1990. Environmental isotope studies of storm and snowmelt runoff generation. In: Anderson, M.G., Burt TP. (eds.) Process studies in hill-slope hydrology. Wiley, Chichester, pp. 401-435.

Solomon, S. and Quiel, F., 2006. Groundwater study using remote sensing and geographic information systems (GIS) in the central highlands of Eritrea. Hydrogeology J., 14:729-741.

Stettler E.H., De Beer J.H. and Blom M.P., 1989. Crustal domains in the northern Kaapvaal as defined by magnetic lineaments. Precambrian Res. 45:263-276.

Sudheer K. P., Gosain A. K., Mohana R. D. and Saheb S. M. (2002) "Modelling evaporation using an artificial neural network Algorithm" Hydrol. Process. 16, 3189-3202.

Sawatzky, D. L., Raines, G. L., Bonham-Carter, G. F., 2009b. Spatial Data Modeller (SDM) for ArcMAP 9.3: Help Files.

Teeuw, R., 1995. Groundwater exploration using remote sensing and a low-cost geographic information system. Hydrogeology Journal 3, 21-30.

Tessema, A., Mengistu H., Chirenje, E., Abiye, T., Demlie, M., 2012. The relationship between lineaments and borehole yield in the North West Province, South Africa: results from geophysical studies, Hydrogeology Journal 20, 351-368.

Titus, R., Friese, A. and Adams, S., 2009. A tectonic and geomorphic framework for the development of basement aquifers in Namaqualand – a review. In: Titus, R., Beekman, H. and Adams, S. (Ed.). The basement aquifers of Southern Africa, Water Research Commission Report, TT428/09, 5-18.

Tsoukalas, I.H., and Uhrig, R.E., 1997. Fuzzy and neural approaches in engineering: New York, John Wiley and Sons, Inc, 587p.

van der Westhuizen, W. A., de Bruiyn H. and Meintjies P.G., 1991. The Ventersdorp Supergroup: an overview. J. African Earth Sciences, 33, 83-105

van der Westhuizen, W. A., de Bruiyn H., 2000. High temperature ash flow – wet sediment interaction in the Makwassie Formation, Ventersdorp Supergroup, South Africa, Precambrian Research, 101, 341-351.

Van der Westhuizen, W., A., de Bruiyn, H., Meintjies, P.G., 1991. The Ventersdorp Supergroup: an overview. Journal of African Earth Sciences 33, 83-105.

Vegter, J.R., 1995: An explanation of a set of national groundwater maps; Report TT 74/95, Water Research Commission Report, Pretoria.

Vengosh, A., Heining, S., Ganor, J., Mayer, B., Weyhenmeyer, C.E., Bullen, T.D. and Paytan, A., 2007. New isotope evidence for the origin of groundwater from the Nubian Sandstone Aquifer in the Negev, Israel. Applied Geochemistry, 22, 1052-1073.

WHO (World Health Organization). 2006. Guidelines for Drinking-water Quality. Third Edition. Volume 1, Recommendations. Geneva, ISBN 92 4 154696 4.

Williams, A.E. (1997). Stable isotope tracers: natural and anthropogenic recharge, Orange County, California. Journal of Hydrology, Science Direct,201, issues 1-4, 230-248

Winter, H, de la R., 1976. A lithostratigraphic classification of the Ventersdorp succession. Trans Geol. Soc. S. Afr., 79, 31-48.

Wright, E.P., 1992. The hydrogeology of crystalline basement aquifers in Africa. In: Wright, E.P. and Burgess W.G. (Ed) The hydrogeology of crystalline basement aquifers in Africa. The Geological Society Special Publication, 1-27.

Yen, J., Langari, R., 1999. Fuzzy Logic: Intelligence, Control, and Information. Prentice-Hall.

Zadeh, L.A., 1993. A computational approach to fuzzy quantifiers in natural language: Computers and Mathematics, 9, 149-184.

Appendix – A

Results of Follow-up Geophysical Surveys

Time domain electromagnetic (TDEM) and vertical electrical soundings were carried out at selected sites in Stella and Vryburg as well as Dithakwaneng. One-dimensional modelling of the data constrained by borehole log provided valuable information about the depth and thickness of conductive layers that are interpreted in terms of possible water-bearing formation.

The results show conductive layers are located at a depth ranging from 20 m to 50 m with the exception of time domain EM soundings carried out at Vryburg RDP houses where the depth of conductive layers varies from 15 m to 45 m. Below, the results of follow-up geophysical surveys carried out at the three sites is presented.

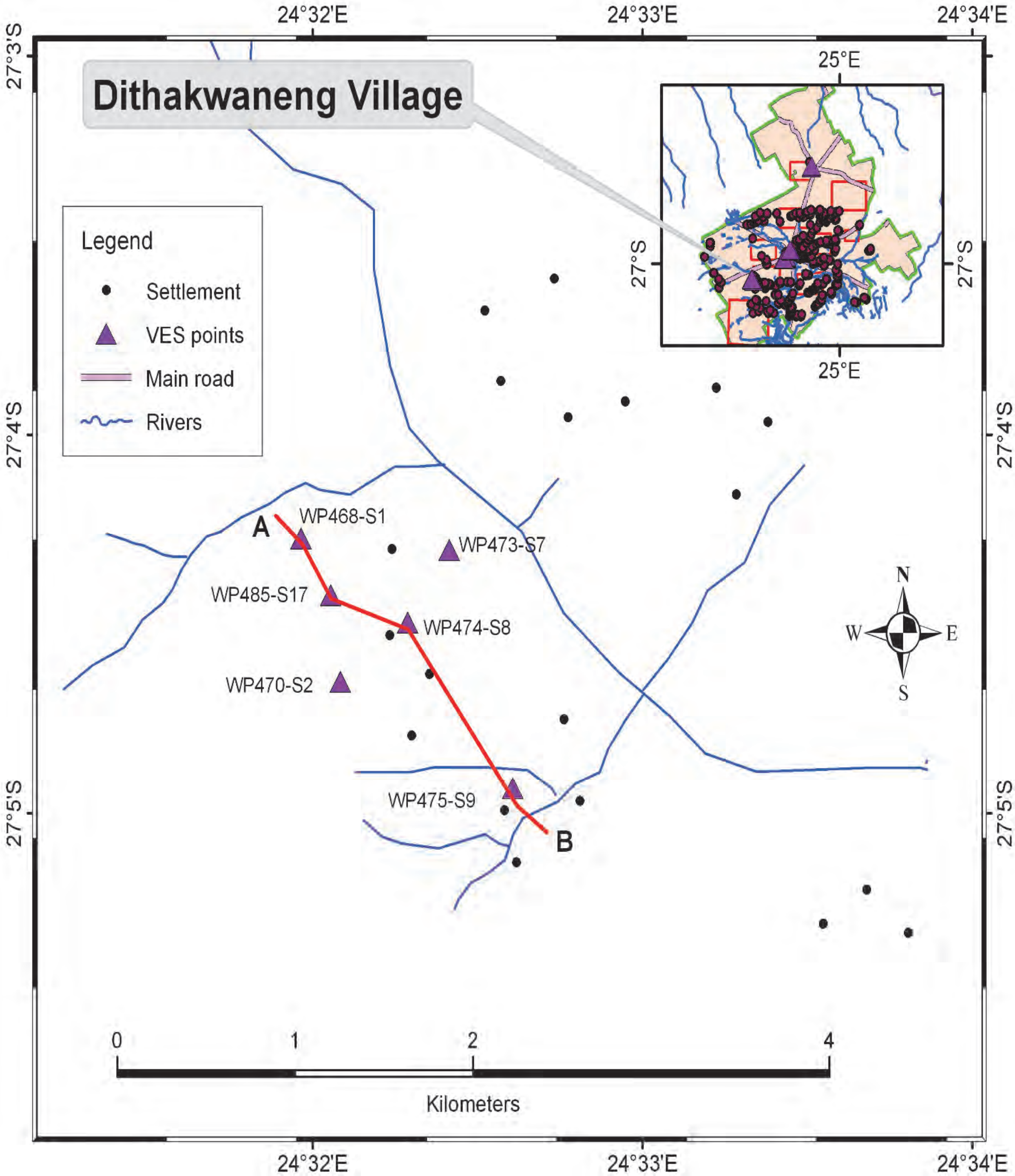
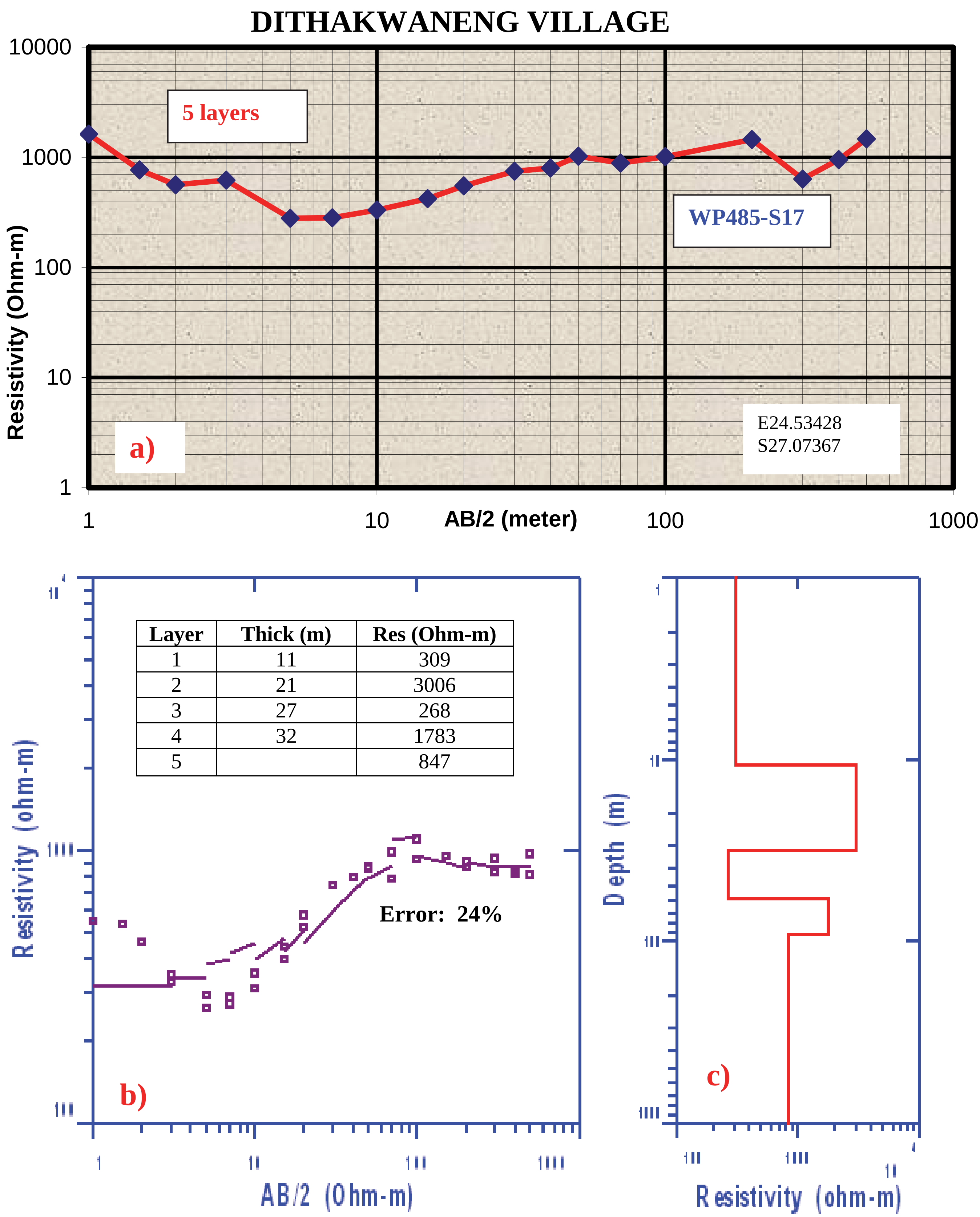


Figure A1: Location map of follow-up geophysical surveys



Layer	Thick (m)	Res (Ohm-m)
1	11	309
2	21	3006
3	27	268
4	32	1783
5		847

Figure A2 Resistivity – depth profile at sounding S17 (Dithakwaneng Village)

STELLA TOWN

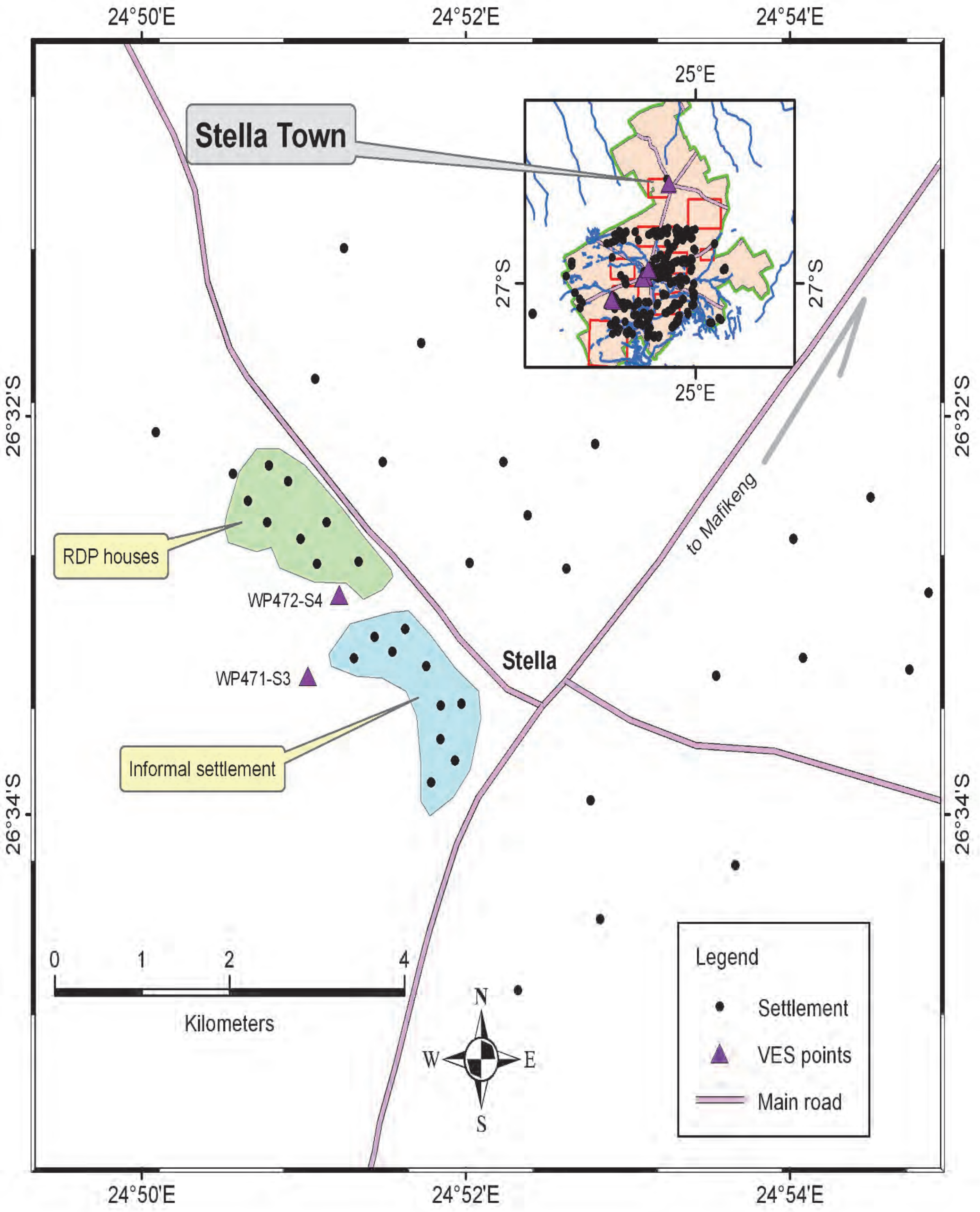
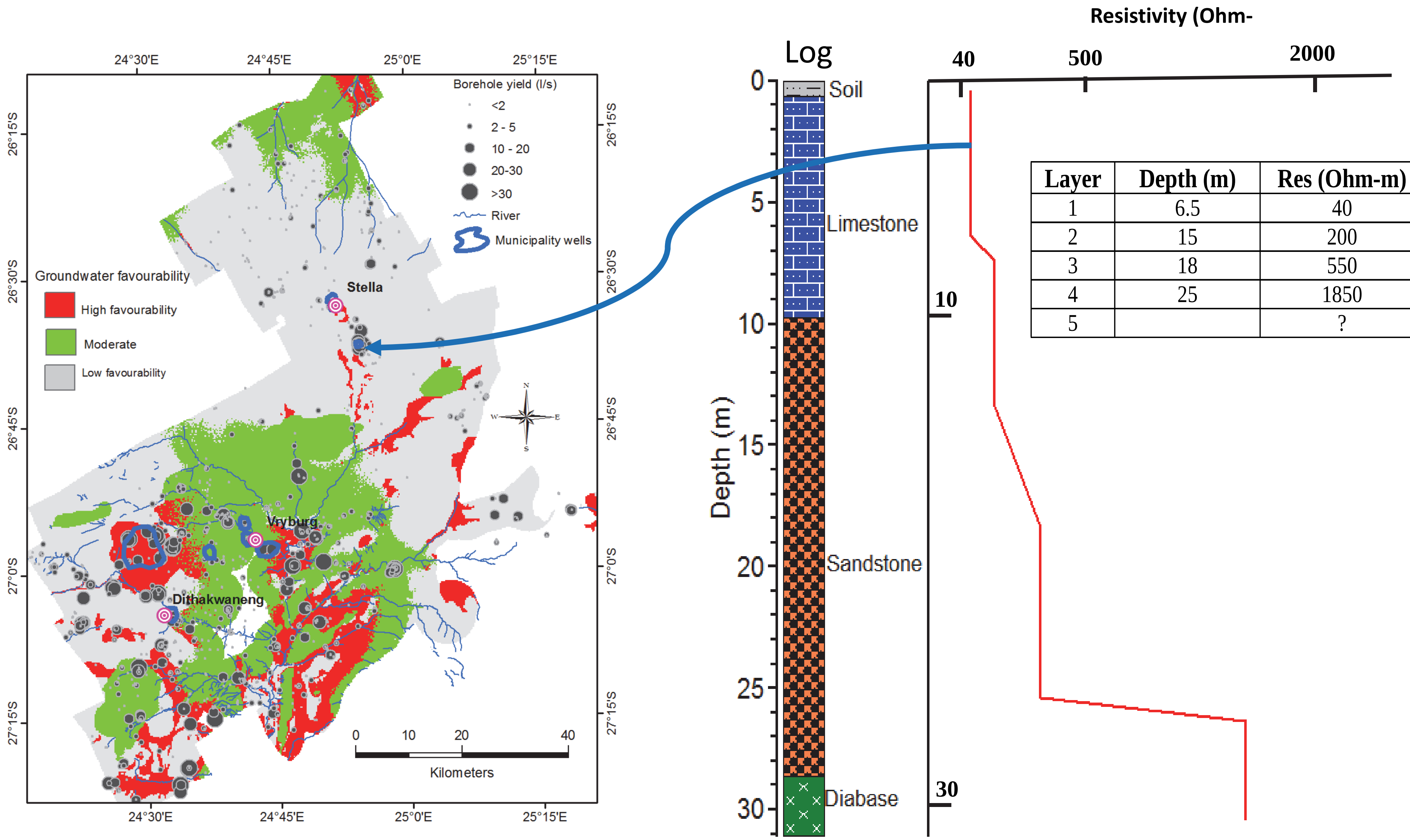


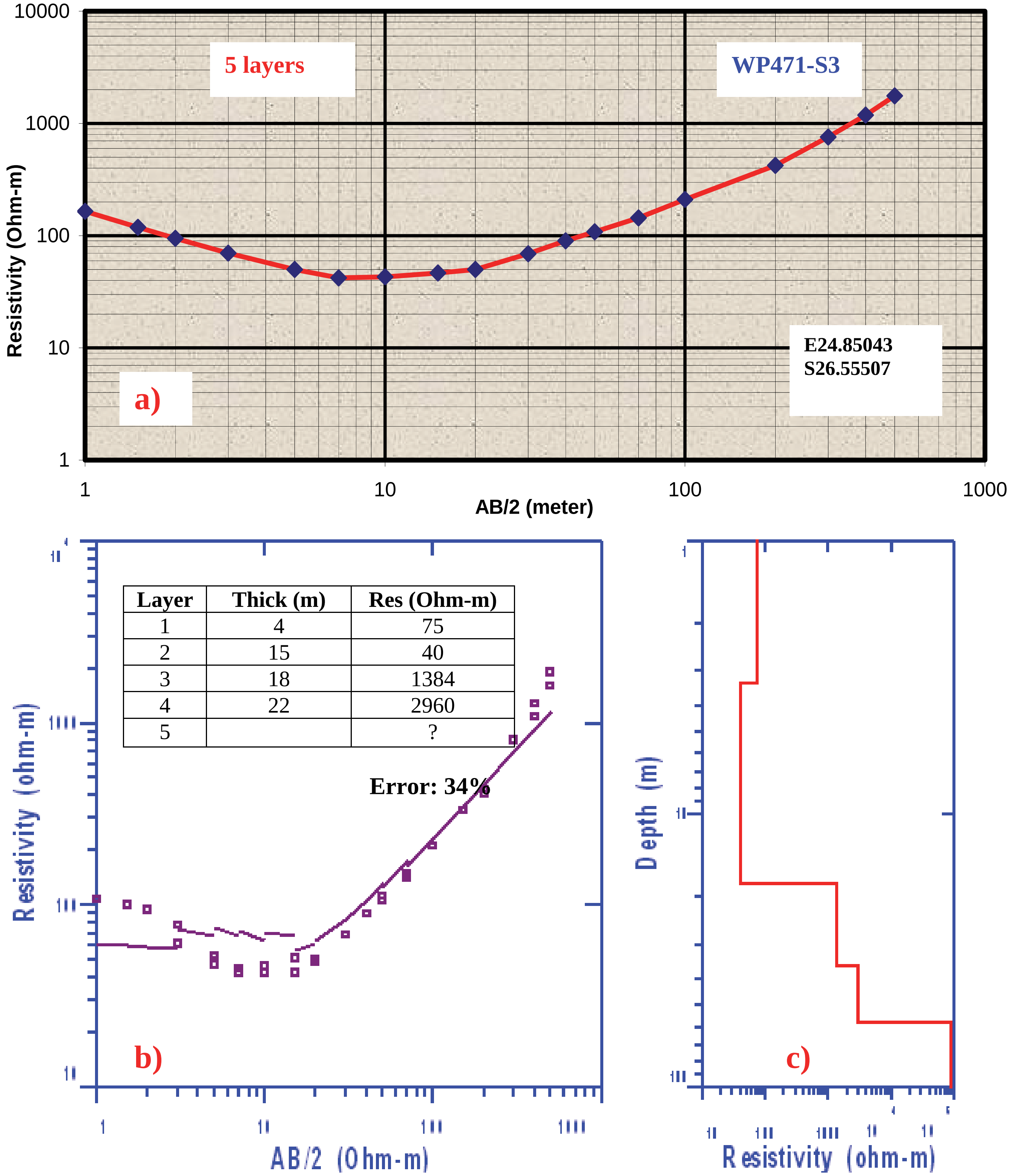
Figure A3: Location map of geophysical surveys at Stella town.



Layer	Depth (m)	Res (Ohm-m)
1	6.5	40
2	15	200
3	18	550
4	25	1850
5		?

Figure A4: Comparison of 1D modelling of Time domain EM with borehole log measured at Stella town.

STELLA TOWN



Layer	Thick (m)	Res (Ohm-m)
1	4	75
2	15	40
3	18	1384
4	22	2960
5		?

Figure A5: 1D resistivity – depth profile derived from inversion of Schlumberge vertical electrical sounding carried out at Stella town (Informal settlement).

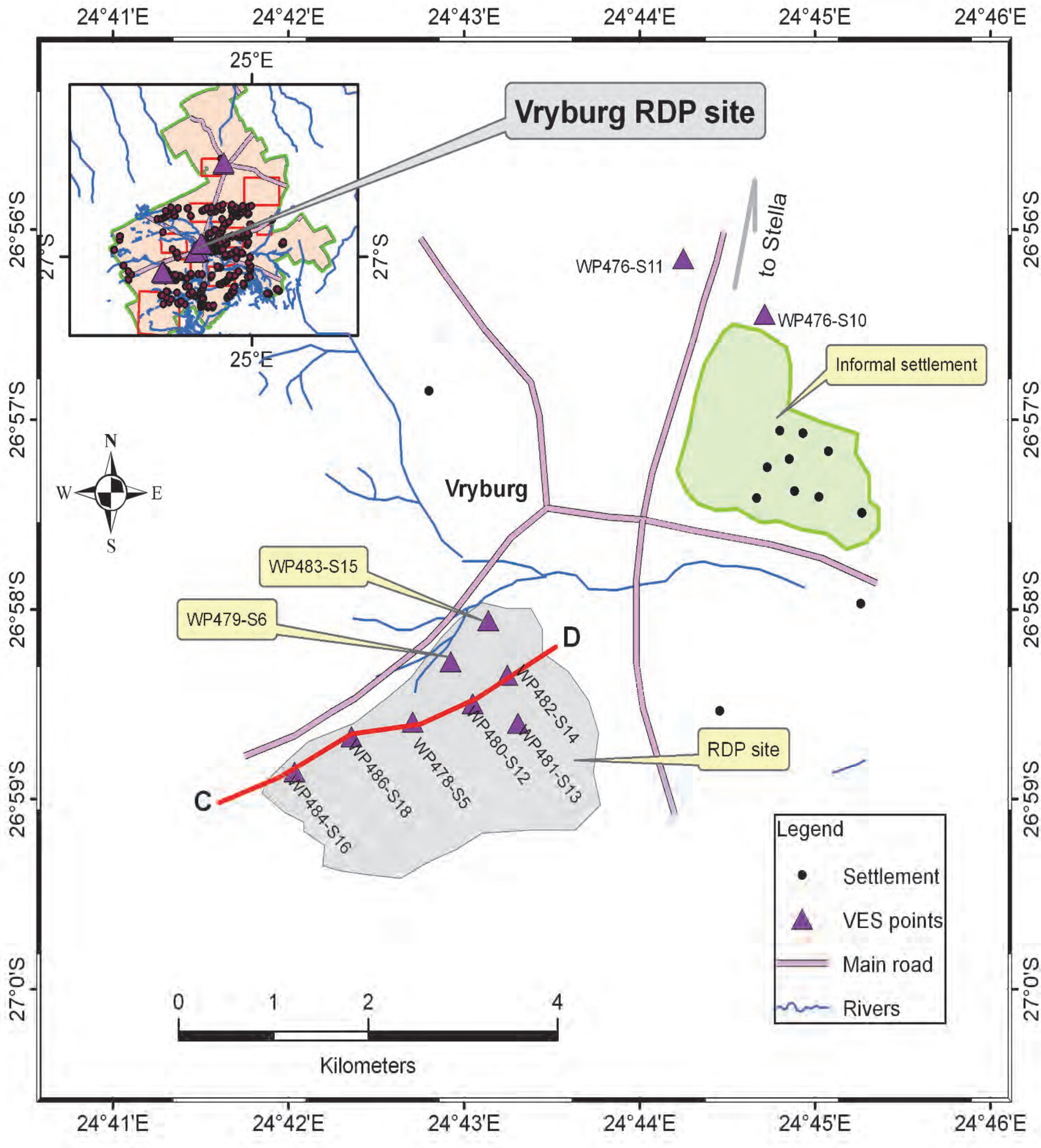
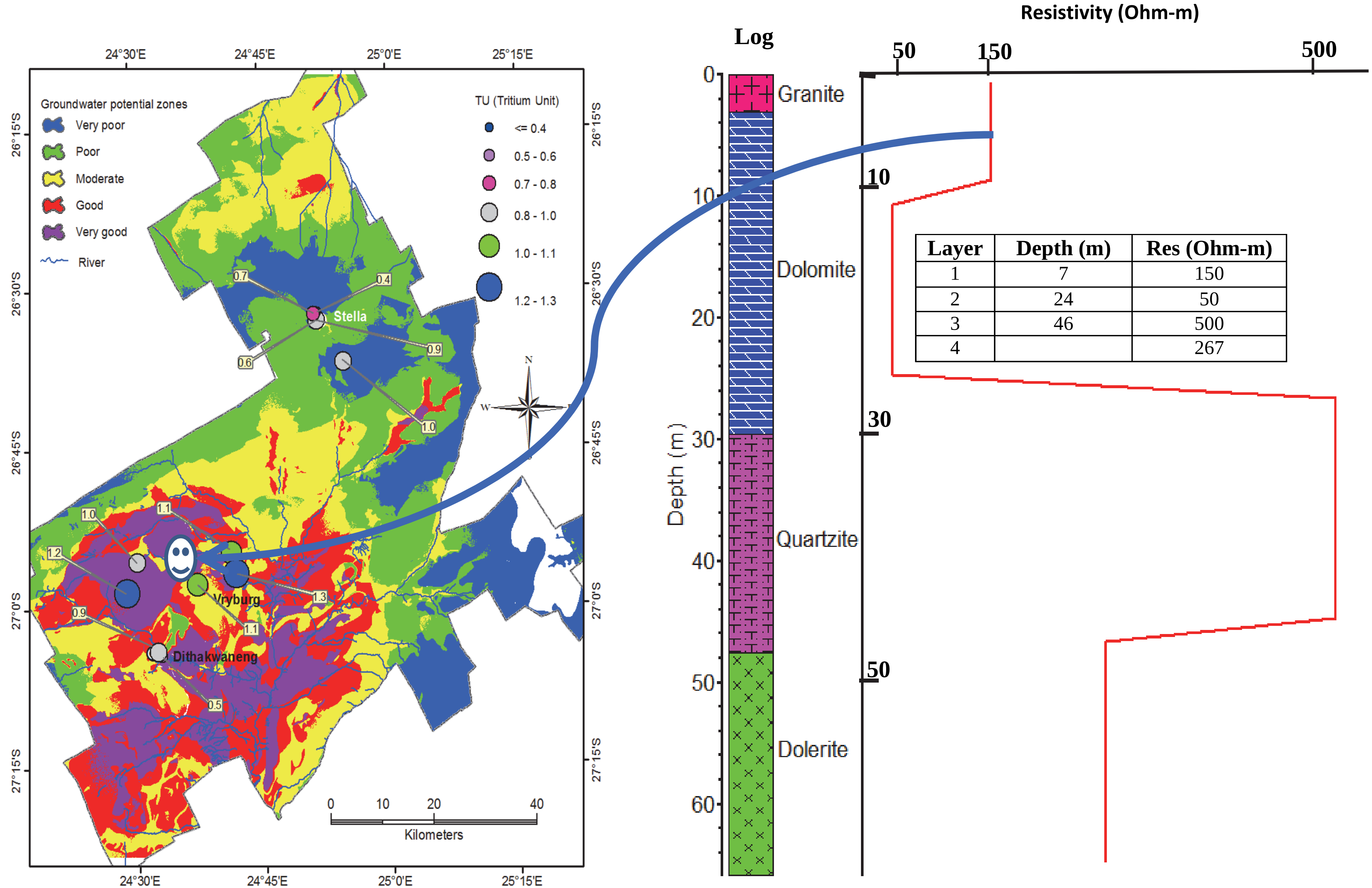


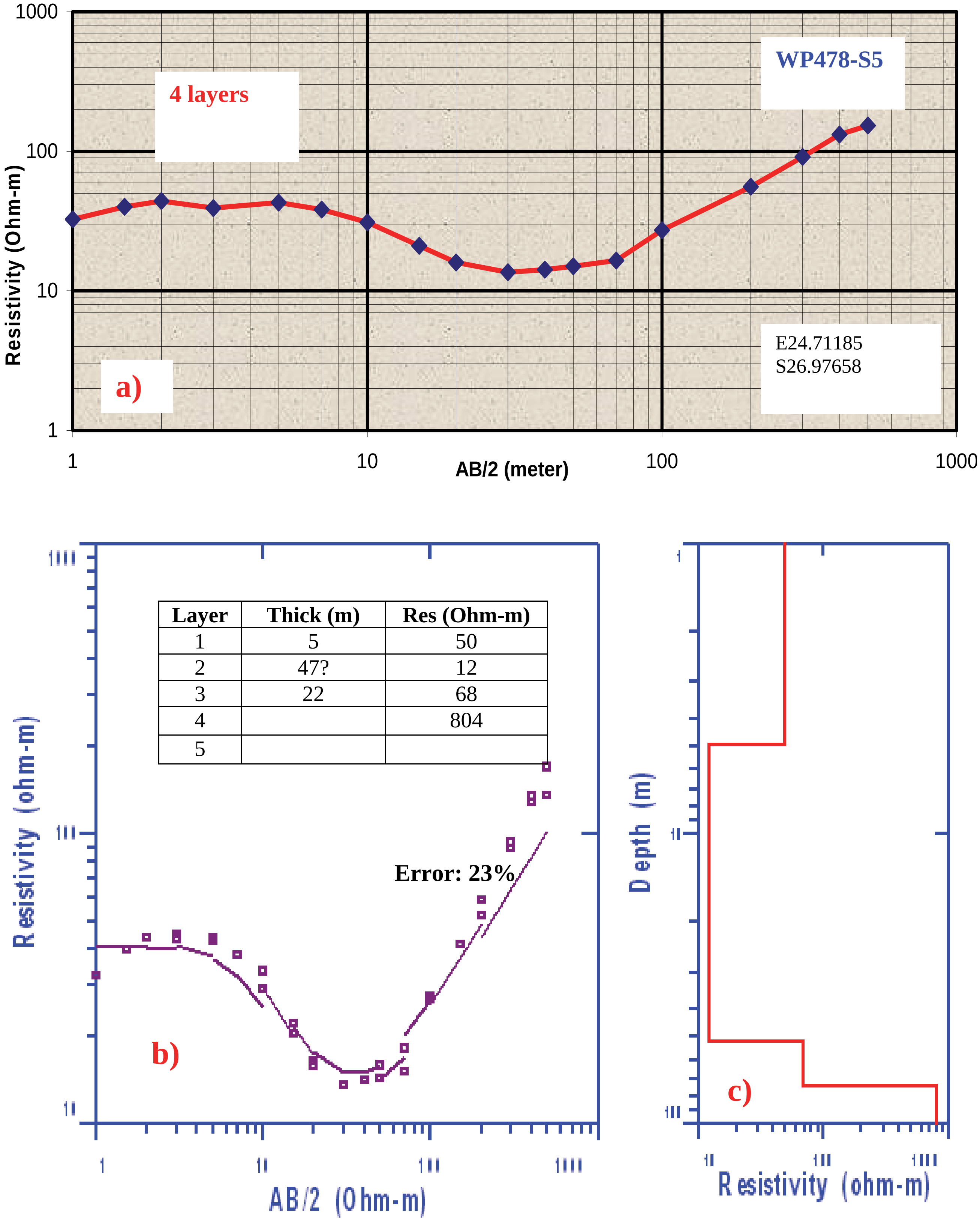
Figure A6: Location map of geophysical surveys at Vryburg RDP houses.



Layer	Depth (m)	Res (Ohm-m)
1	7	150
2	24	50
3	46	500
4		267

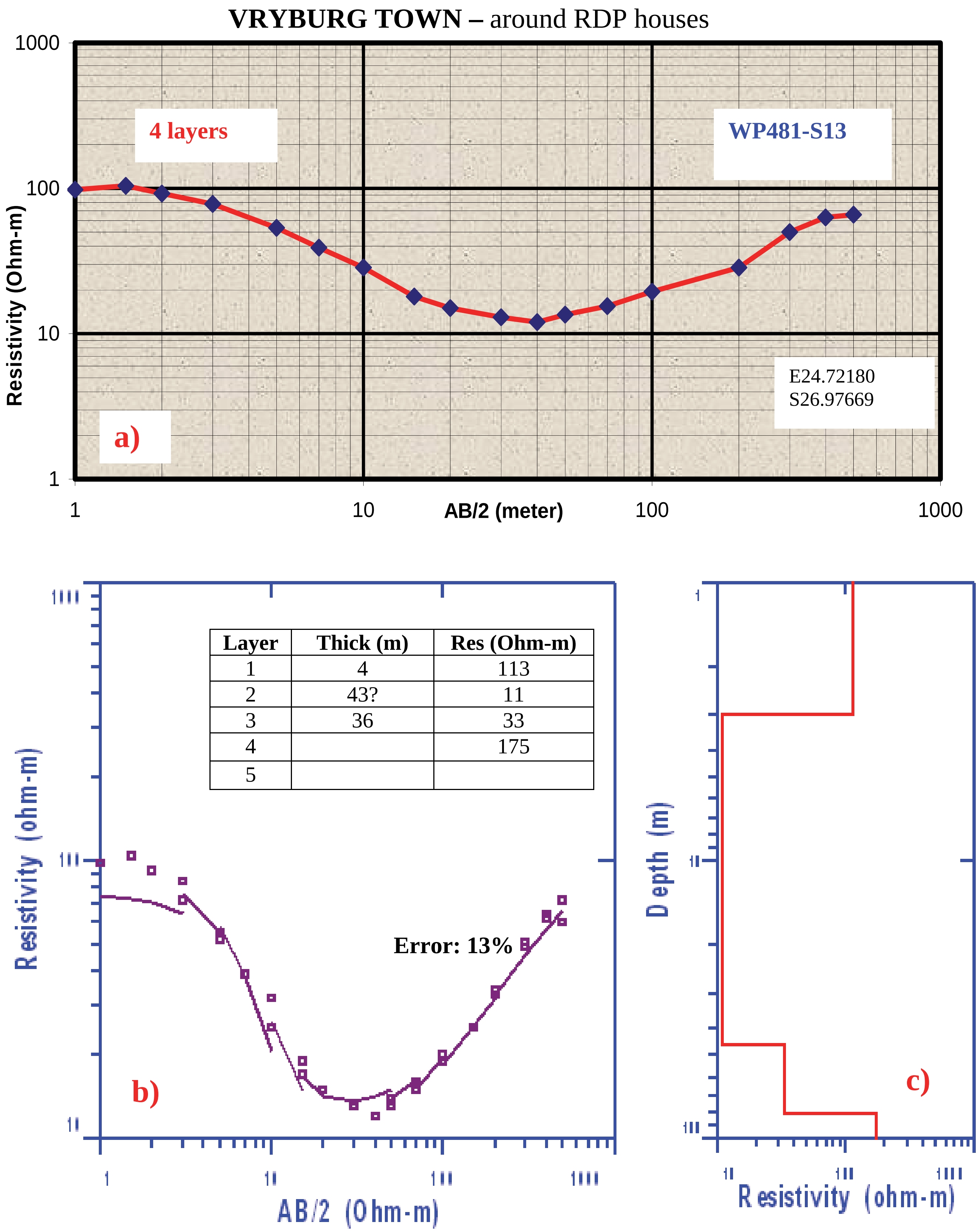
Figure A7: Comparison of 1D model of Time Domain EM with borehole log measured at Vryburg RDP houses.

VRYBURG TOWN – southern part of the town close to RDP houses



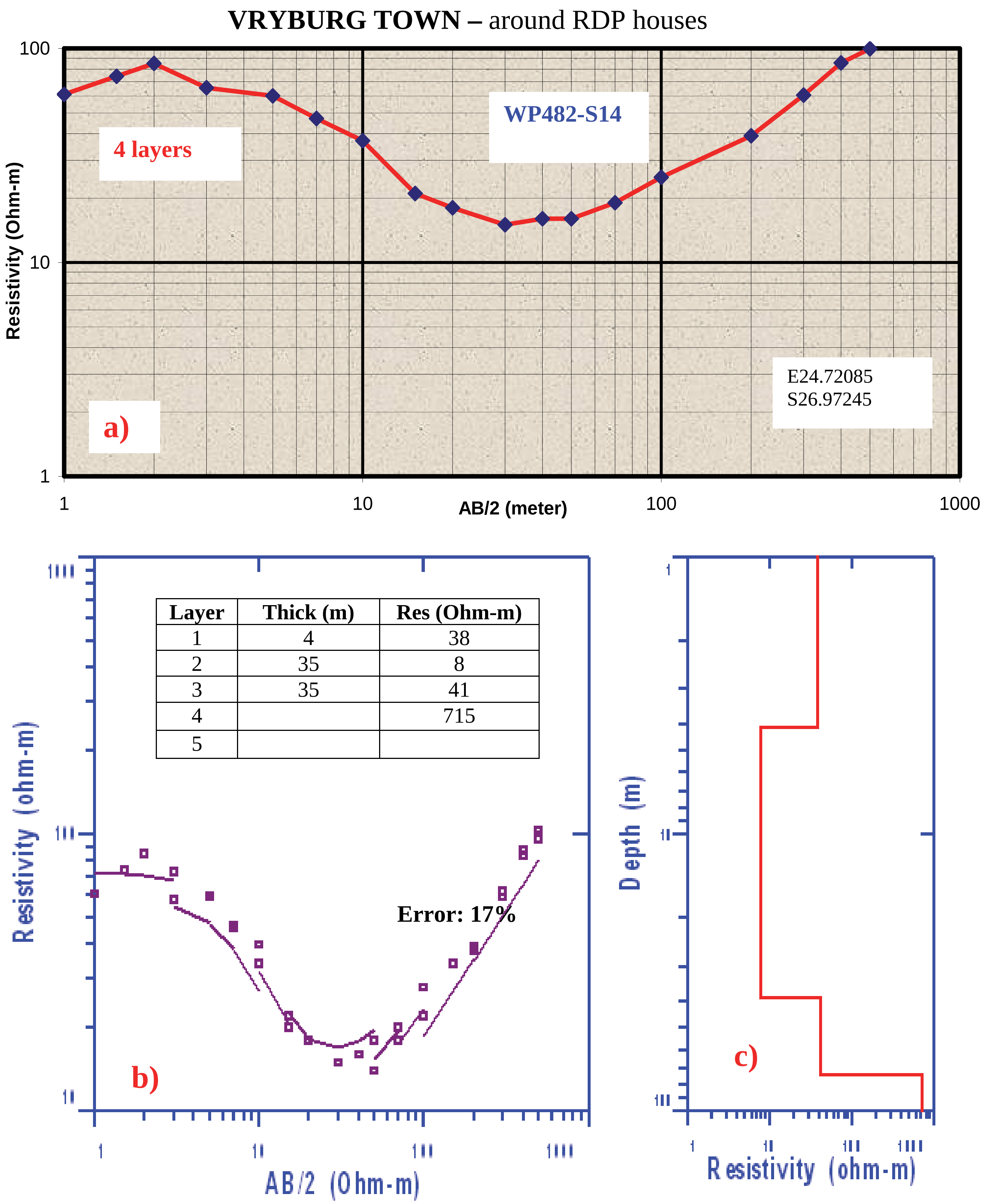
Layer	Thick (m)	Res (Ohm-m)
1	5	50
2	47?	12
3	22	68
4		804
5		

Figure A8: 1D resistivity – depth profile derived from inversion of Schlumberge vertical electrical sounding carried out at Vryburg town (near RDP houses).



Layer	Thick (m)	Res (Ohm-m)
1	4	113
2	43?	11
3	36	33
4		175
5		

Figure A9: 1D resistivity – depth profile derived from inversion of Schlumberge vertical electrical sounding carried out at Vryburg town (around RDP houses).



Layer	Thick (m)	Res (Ohm-m)
1	4	38
2	35	8
3	35	41
4		715
5		

Figure A10: 1D resistivity – depth profile derived from inversion of Schlumberge vertical electrical sounding carried out at Vryburg town (around RDP houses).