

# **Geo-Statistical Analysis and Sub-Delineation of all Vegter Regions**

Report to the  
**Water Research Commission**

by

North West University

**WRC Report No. 2745/1/19**

**ISBN 978-0-6392-0098-9**

**January 2020**

**Obtainable from**  
Water Research Commission  
Private Bag X03  
GEZINA, 0031

[orders@wrc.org.za](mailto:orders@wrc.org.za) or download from [www.wrc.org.za](http://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 Vegter methodology is used to perform geostatistical analysis on the Vegter regions. This methodology has been well documented by Vegter himself. It is however a tedious process to obtain all the relevant data, process the data and calculate the required statistics. In the light of the aforementioned a software tool was developed to automatically perform the analysis.

Historically all the statistical analyses were done in the context of the simplified geology of the study area in question. However, the software tool also allows the Vegter analysis to be done in the context of quaternary catchment, groundwater occurrence, aquifer type, borehole depth, borehole elevation and aquifer vulnerability, all at the press of a button. This contextual analysis assists the user in understanding the study area in question by examining the different relationships.

Vegter goes further and states that sub-delineation of the Vegter regions are required, but does not describe this process. A delineation methodology was developed to address the shortcomings of previous methodologies by relating delineation classes to a physical parameter termed inferred transmissivity. The inferred transmissivity is a calculated parameter to estimate the actual transmissivity that would have been obtained by performing a pump test. The inferred transmissivity is obtained through a classification tree making use of borehole yield values. Four delineation classes for transmissivity (0-1, 1-5, 5-25, >25) are obtained from the classification tree.

The delineation is dependent on data density and interpolation methods used, but allows for continuous delineation across study area boundaries to neighboring study areas. The fact that a physical parameter is used in the delineation process ensures that delineation results are consistent across large areas as well as areas where high contrasting borehole parameters are present.

This page was intentionally left blank

## TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY</b> .....	<b>III</b>
<b>LIST OF FIGURES</b> .....	<b>VIII</b>
<b>LIST OF TABLES</b> .....	<b>XI</b>
<b>ACRONYMS</b> .....	<b>XII</b>
<b>1 INTRODUCTION</b> .....	<b>1</b>
1.1 PREAMBLE .....	1
1.2 VEGTER'S METHODOLOGY .....	1
1.3 VEGTER'S DELINEATION .....	2
1.3.1 Vegter's Depth to Water Level.....	5
1.3.2 Vegter's Borehole Probability Index .....	6
1.3.3 Vegter's Recharge Map.....	9
1.3.4 Vegter's Saturated Insterstices .....	9
1.3.5 Vegter's water quality maps .....	11
<b>2 DATA SOURCES</b> .....	<b>14</b>
2.1 BACKGROUND .....	14
2.2 BOREHOLE INFORMATION .....	16
2.3 SPATIAL INFORMATION.....	16
2.4 ELEVATION DATA .....	17
2.5 ONGOING DATA CAPTURE.....	19
<b>3 DATABASE DESIGN</b> .....	<b>20</b>
3.1 DATABASE SOFTWARE .....	20
3.2 RATIONAL DATABASE .....	20
3.2.1 Site Table.....	21
3.2.2 Static Level Table .....	21
3.2.3 Water Strike Table .....	22
3.3 DATABASE STATISTICS AND PREVIEWS .....	23

3.3.1	<i>Basic Statistics</i> .....	23
3.3.2	<i>Spatial Datasets</i> .....	32
<b>4</b>	<b>STATISTICAL ANALYSIS</b> .....	<b>39</b>
4.1	INTRODUCTION .....	39
4.2	SIMPLIFIED GEOLOGY .....	40
4.3	CALCULATED STATISTICS.....	42
4.3.1	<i>Borehole Distribution and Density</i> .....	42
4.3.2	<i>Drilling Success Rate</i> .....	45
4.3.3	<i>Borehole Yields in Various Geologies</i> .....	45
4.3.4	<i>Distribution of Boreholes per Depth</i> .....	46
4.3.5	<i>Strike Frequency and Cumulative Strike Frequency</i> .....	48
4.3.6	<i>Yield vs Strike Analysis</i> .....	48
4.3.7	<i>Yield vs Dyke Intersection</i> .....	49
4.3.8	<i>Borehole Water Levels per Geology</i> .....	50
4.3.9	<i>Borehole Chemistry per Geology</i> .....	52
<b>5</b>	<b>DELINEATION METHODOLOGY</b> .....	<b>53</b>
5.1	PREAMBLE .....	53
5.2	DELINEATION CLASS METHOD.....	53
5.3	INFERRED TRANSMISSIVITY CLASSIFICATION METHOD .....	57
5.3.1	<i>Introduction</i> .....	57
5.3.2	<i>Aquifer Parameters</i> .....	57
5.3.3	<i>Classification</i> .....	63
5.4	INTERPOLATION TECHNIQUES.....	66
5.4.1	<i>Inverse Distance Weighting (Type 1)</i> .....	68
5.4.2	<i>Splines (Type 1) (After Ikechukwu et al., 2017)</i> .....	68
5.4.3	<i>Kriging (Type 2)</i> .....	69
5.5	DELINEATION .....	70
<b>6</b>	<b>ALGORITHMS</b> .....	<b>71</b>
6.1	STATISTICAL CALCULATIONS.....	71
6.1.1	<i>Borehole Density</i> .....	71

6.1.2	<i>Borehole Parameter Histograms</i> .....	72
6.1.3	<i>Water Strike vs Yield</i> .....	74
6.1.4	<i>Water Level Correlation</i> .....	74
6.2	<i>WATER CHARACTER</i> .....	75
6.3	<i>DELINEATION</i> .....	76
6.3.1	<i>Delineation Classification Tree</i> .....	76
6.3.2	<i>Cooper-Jacob Delineation</i> .....	77
<b>7</b>	<b><i>IMPLEMENTATION</i></b> .....	<b>79</b>
7.1	<i>STUDY AREA</i> .....	79
7.2	<i>CLASSIFICATION TREE DELINEATION</i> .....	81
7.3	<i>COOPER-JACOB DELINEATION</i> .....	84
7.4	<i>DATA DENSITY</i> .....	87
7.5	<i>COMPARATIVE RESULTS</i> .....	88
7.6	<i>SOFTWARE TOOL</i> .....	88
<b>8</b>	<b><i>CONCLUSIONS AND RECCOMENDATIONS</i></b> .....	<b>90</b>
	<b><i>REFERENCES</i></b> .....	<b>93</b>
	<b><i>APPENDIX A - DRASTIC INDEX CALCULATION FOR SOUTH AFRICA</i></b> .....	<b>94</b>
	<b><i>APPENDIX B - PLOTTING PROCEDURE FOR PIPER DIAGRAM</i></b> .....	<b>105</b>
	<b><i>APPENDIX C - ROOT MEAN SQUARE ERROR</i></b> .....	<b>106</b>
	<b><i>APPENDIX D - MEAN AND STANDARD DEVIATION CALCULATIONS</i></b> .....	<b>107</b>
	<b><i>APPENDIX E - PEARSON CORRELATION</i></b> .....	<b>108</b>
	<b><i>APPENDIX F - INVERSE DISTANCE WEIGHTING</i></b> .....	<b>109</b>
	<b><i>APPENDIX G - SPLINE INTERPOLATION</i></b> .....	<b>110</b>
	<b><i>APPENDIX H - KRIGING INTERPOLATION</i></b> .....	<b>113</b>
	<b><i>APPENDIX I - USER MANUAL FOR THE SOFTWARE</i></b> .....	<b>119</b>

## LIST OF FIGURES

Figure 1 – Vegter regions .....	3
Figure 2 – Vegter's depth to water level map .....	5
Figure 3 – Vegter's depth to water level standard deviation .....	6
Figure 4 – Vegter's probability index of a successful borehole .....	7
Figure 5 – Vegter's probability of drilling a borehole with yield of 2 L/s and more .....	8
Figure 6 – Vegter's recharge map.....	9
Figure 7 – Vegter's optimal drilling depth.....	10
Figure 8 – Vegter's storage coefficient .....	10
Figure 9 – Vegter's TDS map .....	11
Figure 10 – Vegter's calcium sulphate waters.....	12
Figure 11 – Vegter's calcium bicarbonate waters.....	12
Figure 12 – Vegter's sodium bicarbonate waters.....	13
Figure 13 – Vegter's sodium chloride waters .....	13
Figure 14 – NGA Borehole distribution (DWS, 2017) .....	15
Figure 15 – Example of actual borehole coordinate and farm centroid coordinate .....	17
Figure 16 – Borehole water level vs elevation.....	18
Figure 17 – Multiple boreholes assigned to same farm centroid with offset applied .....	19
Figure 18 – Entity Relationship Diagram.....	20
Figure 19 – Database borehole distribution.....	25
Figure 20 – Database boreholes with water levels .....	26
Figure 21 – Database boreholes with EC values.....	27
Figure 22 – Database boreholes with chemistry values (2008).....	28
Figure 23 – Database boreholes with yield values .....	29
Figure 24 – Database boreholes with water strike values .....	30
Figure 25 – Database boreholes with borehole logs.....	31
Figure 26 – SRTM90 elevation grid of South Africa .....	34
Figure 27 – Simplified geology map (aquifer classification).....	35
Figure 28 – Geological and structural lineaments .....	36
Figure 29 – Groundwater occurrence map.....	37
Figure 30 – Groundwater vulnerability map .....	38
Figure 31 – Example of simplified geology .....	41

Figure 32 – Example of borehole distributions .....	43
Figure 33 – Graphical representation of borehole yield and discharge per geology .....	46
Figure 34 – Graphical representation of borehole depth and elevation per geology.....	47
Figure 35 – Example depth frequency plot .....	47
Figure 36 – Example strike frequency plot for a geological unit .....	48
Figure 37 – Example yield versus strike.....	49
Figure 38 – Example yield versus strike.....	50
Figure 39 – Example of groundwater level versus surface elevation.....	51
Figure 40 – Example of groundwater level distribution within geological unit.....	51
Figure 41 – Example Piper diagram of example study area .....	52
Figure 42 – Voronoi polygon generation with simplification.....	55
Figure 43 – Voronoi polygon generation with simplification.....	56
Figure 44 – Aquifer storativity values on regional scale .....	58
Figure 45 – Vegter’s storage coefficient map.....	59
Figure 46 – Storage coefficient vs. conversion factor.....	61
Figure 47 – Probability distributions of calculated transmissivity.....	62
Figure 48 – Probability distributions of calculated transmissivity.....	63
Figure 49 – Classification tree considering both yield and available drawdown.....	65
Figure 50 – Classification tree only considering yield.....	65
Figure 51 – Interpolation results from the same dataset (taken from GoldenSoftware).....	67
Figure 52 – Examples of Variogram models (Bohling, 2005).....	70
Figure 53 – Borehole density algorithm.....	72
Figure 54 – Generic histogram algorithm .....	73
Figure 55 – Water strike vs. yield algorithm .....	74
Figure 56 – Water level correlation algorithm .....	75
Figure 57 – Water character algorithm .....	76
Figure 58 – Inferred transmissivity delineation algorithm .....	77
Figure 59 – Inferred transmissivity delineation algorithm based on Cooper-Jacob.....	78
Figure 60 – Vegter region 7 geological map .....	79
Figure 61 – Vegter region 7 groundwater occurrence map.....	80
Figure 62 – Vegter region 7 GRAII transmissivity map.....	81
Figure 63 – Vegter region 7 boreholes with yield and transmissivity values.....	82
Figure 64 – Vegter region 7 classification tree results .....	83

*Figure 65 – Vegter region 7 boreholes with yield, strike and water level values..... 85*

*Figure 66 – Vegter region 7 Cooper-Jacob results..... 86*

*Figure 67 – Vegter region 65 delineation comparison..... 89*

## LIST OF TABLES

Table 1 – Vegter region names.....	4
Table 2 – Site table field descriptions.....	21
Table 3 – Static water level table field descriptions .....	21
Table 4 – Water strike table field descriptions .....	22
Table 5 – Chemistry table field descriptions .....	22
Table 6 – Borehole log table field descriptions .....	23
Table 7 – Database population summary.....	24
Table 8 – Summary of Vegter’s statistics .....	39
Table 9 – Supplementary statistics to Vegter’s list .....	40
Table 10 – Example of lithological description of the simplified geological map .....	42
Table 11 – Borehole summary per quaternary.....	44
Table 12 – Borehole summary per geology .....	44
Table 13 – Example summary of borehole yields and discharge .....	45
Table 14 – Example summary of borehole depth and elevation .....	46
Table 15 – Borehole summary per quaternary.....	53
Table 16 – Borehole summary per quaternary.....	54
Table 17 – Parameter assumptions .....	61
Table 18 – Classification tree probabilities .....	66
Table 19 – Classed datasets correlation with classed transmissivities .....	66
Table 20 – Region 7 classed transmissivity correlation with classification tree.....	83
Table 21 – Region 7 classed transmissivity correlation with classification tree.....	86
Table 22 – Effect of data density on correlation.....	87
Table 23 – Region 7 correlation between classification tree and Cooper-Jacob.....	88

## *ACRONYMS*

CIDA	Canadian International Development Agency
DEM	Digital Elevation Model
EC	Electrical Conductivity
GRAII	Groundwater Resource Assessment Phase II
GSA	Geohydrological Statistical Analysis
IDW	Inverse Distance Weighting
IT	Inferred Transmissivity
NGA	National Groundwater Archive
RBF	Radial Basis Functions
RMSE	Root Mean Square Error
SRTM90	Shuttle Radar Topography Mission 90m Data
STD	Standard Deviation
TDS	Total Dissolved Solids
WRC	Water Research Commission

# 1 INTRODUCTION

## 1.1 PREAMBLE

According to Vegter (2001), the aim of delineating groundwater regions is to provide guidelines for successful and cost-effective siting of boreholes. Vegter's methodology discussed in this report was derived from the following reports:

- Vegter JR (2001) Hydrogeology of groundwater: Region 1 – Makoppa Dome. WRC Report No TT 135/00, Water Research Commission, Pretoria
- Vegter JR (2001) Hydrogeology of groundwater: Region 3 – Limpopo Granulite-Gneiss Belt. WRC Report No TT 136/00, Water Research Commission, Pretoria
- Vegter JR (2003) Hydrogeology of groundwater: Region 19 – Lowveld. WRC Report No TT 208/03, Water Research Commission, Pretoria
- Vegter JR (2006) Hydrogeology of groundwater: Region 26 – Bushmanland. WRC Report No TT 285/06, Water Research Commission, Pretoria

## 1.2 VEGTER'S METHODOLOGY

The existing methodology is primarily based on existing information derived from national groundwater maps and information/data obtained from the Department of Water and Sanitation. Vegter does however state: *A comprehensive description of the occurrence of groundwater requires detailed field studies in order to identify, catalogue and map not only existing sources of supply but also those particular geomorphologic and geologic features that are indicative of favourable conditions for siting boreholes. In the event of possible exploitation of the latter confirmation by geophysical surveying may be necessary.* However, he does acknowledge that this is almost impossible to do and therefore suggests the approach followed in the above-mentioned reports – which is to perform a detailed assessment of existing data in order to set guidelines for each of the groundwater regions.

The first step in this methodology is to provide a detailed description of the geology, as this plays a predominate role in the character of the groundwater found in the area. Thereafter the occurrence of groundwater is discussed based on existing knowledge – these occurrences are then related to the geology of the area. One of the major distinctions made in all of the reports is the distinction between porous (e.g. alluvial aquifers) and secondary (e.g. hard rock aquifers).

The next step is to identify and characterize springs within the region. The temperature, water quality and rate of flow are discussed in detail.

The third step is to conduct a statistical analysis of available borehole data which is the focus of this report.

Once all these assessments have been completed borehole prospects can be defined in terms of accessibility and exploitability and various sub-regions can be delineated. Guidelines are then provided for further improvement in siting boreholes and the associated production costs of a successful borehole. Suitable geophysical methods for siting boreholes within the sub-regions are recommended and discussed. Associated drilling control procedures are stipulated.

In addition to the drilling of successful boreholes, it is important to know how much water is available as this will influence the long-term success of a borehole. Therefore, recharge and storativity have to be calculated so that a water balance can be determined for each of the sub-regions. Vegter's documents provide no guidelines on how to calculate these parameters.

Finally, water quality of available boreholes is compared to the Department of Water Affairs and Forestry 1996 drinking water guidelines. The water is classified accordingly as ideal, good, marginal, poor and unacceptable. Any harmful ion concentrations are listed.

### **1.3 VEGTER'S DELINEATION**

The Vegter delineation includes 64 regions as shown in Figure 1 and the description of each region is listed in Table 1. The sections that follow present the following Vegter parameters:

- Depth to water level
- Accessibility and probability of drilling a successful borehole
- Groundwater recharge
- Saturated interstices
- Groundwater quality (mean TDS)
- Dominant ions

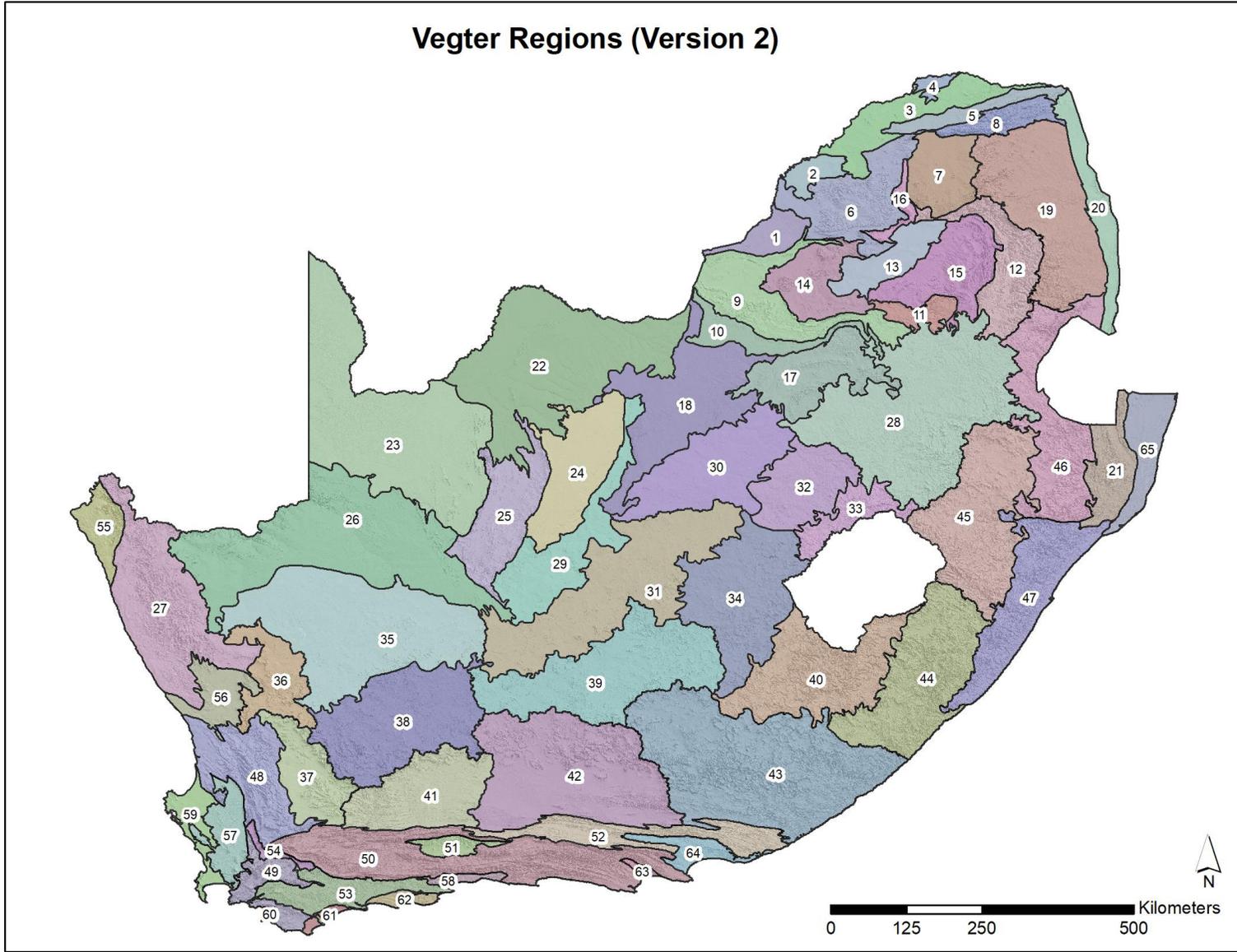


FIGURE 1 - VEGTER REGIONS

TABLE 1 – VECTER REGION NAMES

No	Region Name	No	Region Name
1	Makoppa Dome <i>Crystalline Igneous and metamorphic Basement rocks</i>	33	Northern Highland
2	Waterberg Coal Basin	34	Northeastern Upper Karoo
3	Limpopo Granulite Gneiss Belt <i>Crystalline Igneous and metamorphic Basement rocks</i>	35	Bushmanland Pan Belt
4	Limpopo Karoo Basin	36	Hantam
5	Soutpansberg Hinterland <i>Mainly Extrusive rocks</i>	37	Tanqua Karoo
6	Waterberg Plateau	38	Western Upper Karoo
7	Pietersburg Plateau <i>Crystalline Igneous and metamorphic Basement rocks</i>	39	Eastern Upper Karoo
8	Soutpansberg	40	Southeastern Highland <i>Mainly Extrusive rocks</i>
9	Western Bankeveld and Marico Bushveld	41	Western Great Karoo
10	Karst Belt	42	Eastern Great Karoo
11	Middelburg Basin	43	Ciskeian Coastal Foreland and Middleveld
12	Eastern Bankeveld	44	Transkeian Coastal Foreland and Middleveld
13	Springbok Flats <i>Mainly Extrusive rocks</i>	45	Northwestern Middleveld
14	Western Bushveld Complex <i>Mainly Extrusive rocks</i>	46	Northeastern Middleveld <i>Composite Geology</i>
15	Eastern Bushveld Complex <i>Mainly Extrusive rocks</i>	47	Kwazulu-Natal Coastal Foreland <i>Composite Geology</i>
16	Northern Bushveld Complex <i>Mainly Extrusive rocks</i>	48	Northwestern Cape Ranges
17	Central Highveld <i>Composite Geology</i>	49	Southwestern Cape Ranges
18	Western Highveld <i>Mainly Extrusive rocks</i>	50	Southern Cape Ranges
19	Lowveld <i>Crystalline Igneous and metamorphic Basement rocks</i>	51	Oudtshoorn Basin
20	Northern Lebombo <i>Mainly Extrusive rocks</i>	52	Grootrivier-Klein Winterhoek-Suur-Kaprivier Ranges
21	Southern Lebombo	53	Ruensveld
22	Eastern Kalahari	54	Intermontane Tulbagh-Ashton Valley
23	Western Kalahari	55	Richtersveld
24	Ghaap Plateau	56	Knersvlakte
25	West Griqua Land	57	Swartland
26	Bushmanland <i>Crystalline Igneous and metamorphic Basement rocks</i>	58	Outenikwa Coastal Foreland
27	Namaqualand <i>Crystalline Igneous and metamorphic Basement rocks</i>	59	Southwestern Coastal Sandveld
28	Eastern Highveld	60	Die Kelders Embayment
29	Dry Harts-Lower Vaal-Orange Lowland	61	Bredasdorp Coastal Belt
30	Northeastern Pan Belt	62	Stilbaai Coastal Belt
31	Central Pan Belt	63	Lower Gamtoos Valley
32	Southern Highveld	64	Algoa Basin
32	Makoppa Dome		

### 1.3.1 Vegter's Depth to Water Level

The depth to water level map represents the depth to water level from the ground surface and is shown in Figure 2. This depth to water level map is an inset to the Groundwater Resources of South Africa map (Seymore, 1994). In addition to the mean water level the standard deviation associated with mean value is also available as shown in Figure 3.

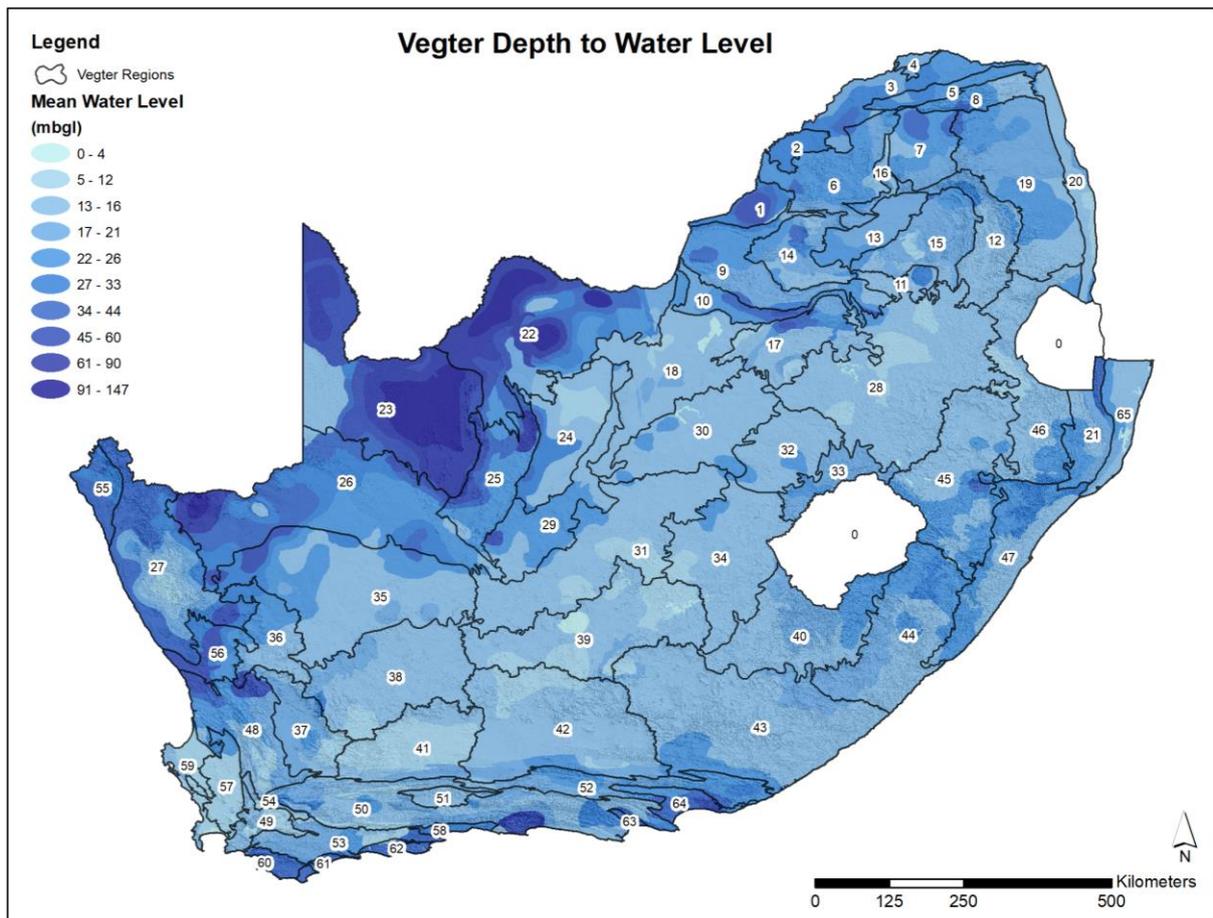


FIGURE 2 – VEGTER'S DEPTH TO WATER LEVEL MAP

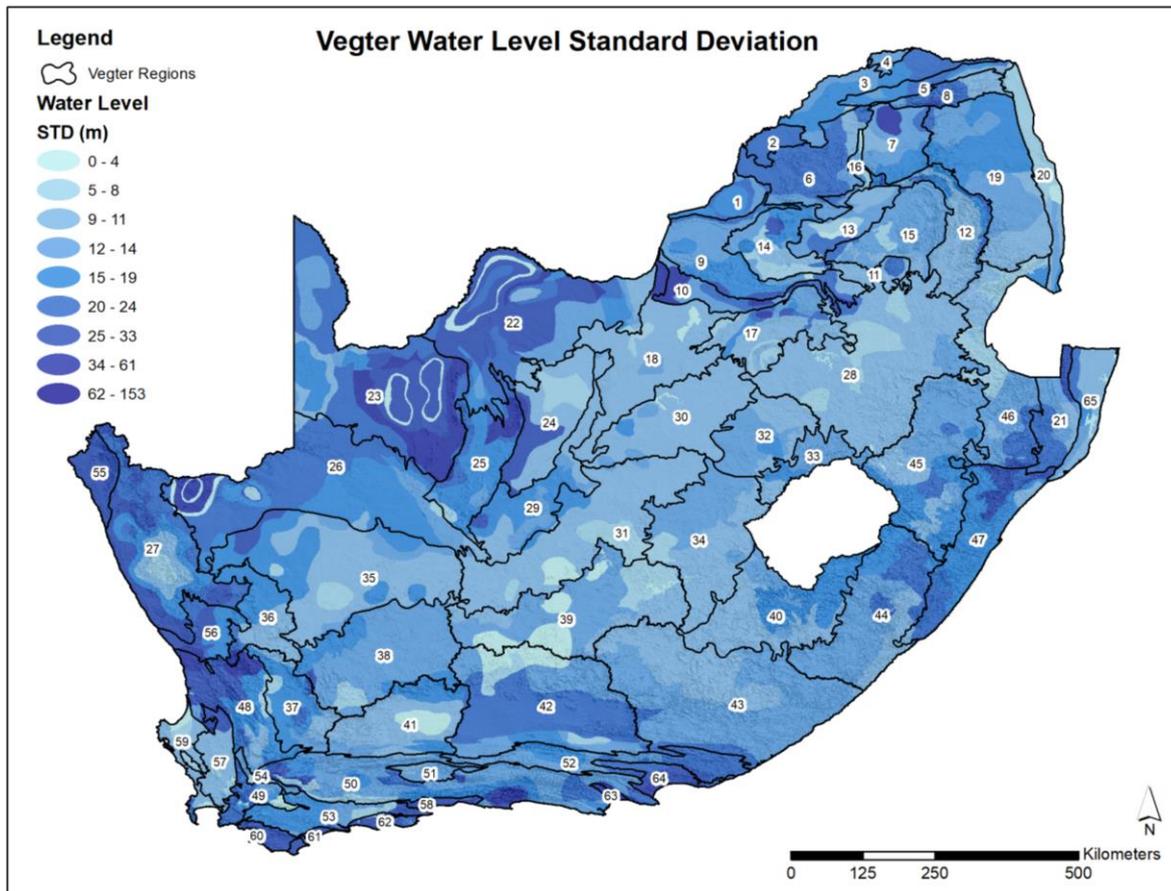


FIGURE 3 – VEGTER'S DEPTH TO WATER LEVEL STANDARD DEVIATION

### 1.3.2 Vegter's Borehole Probability Index

The probability of drilling a successful borehole is shown in Figure 4 where the probability is indicated with an index ranging between 0 and 10. The probability of drilling a successful borehole with a yield of 2 L/s and more is shown in Figure 5. The union of the aforementioned maps was the first groundwater map of South Africa (Seymore, 1994).

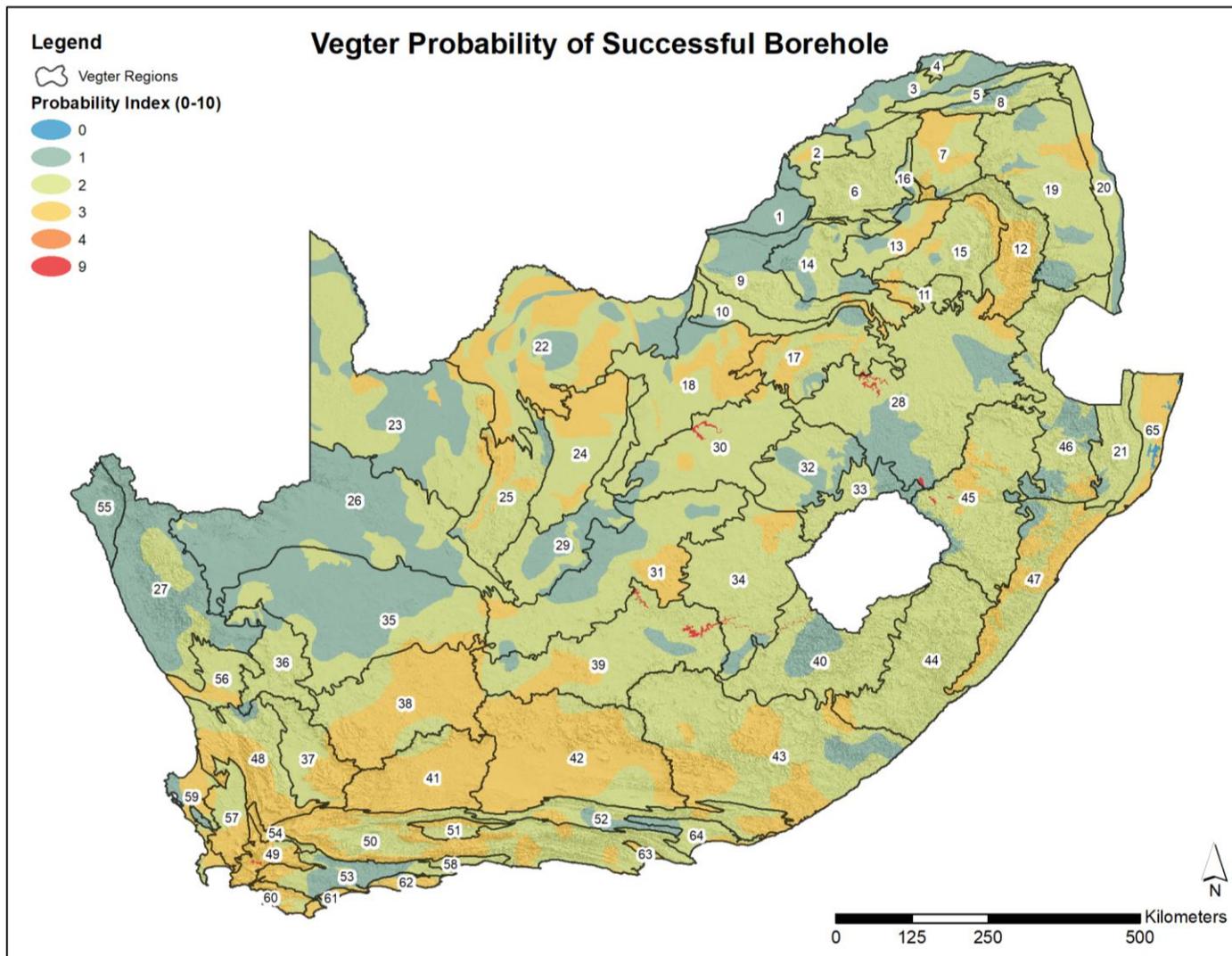


FIGURE 4 – VEGTER'S PROBABILITY INDEX OF A SUCCESSFUL BOREHOLE

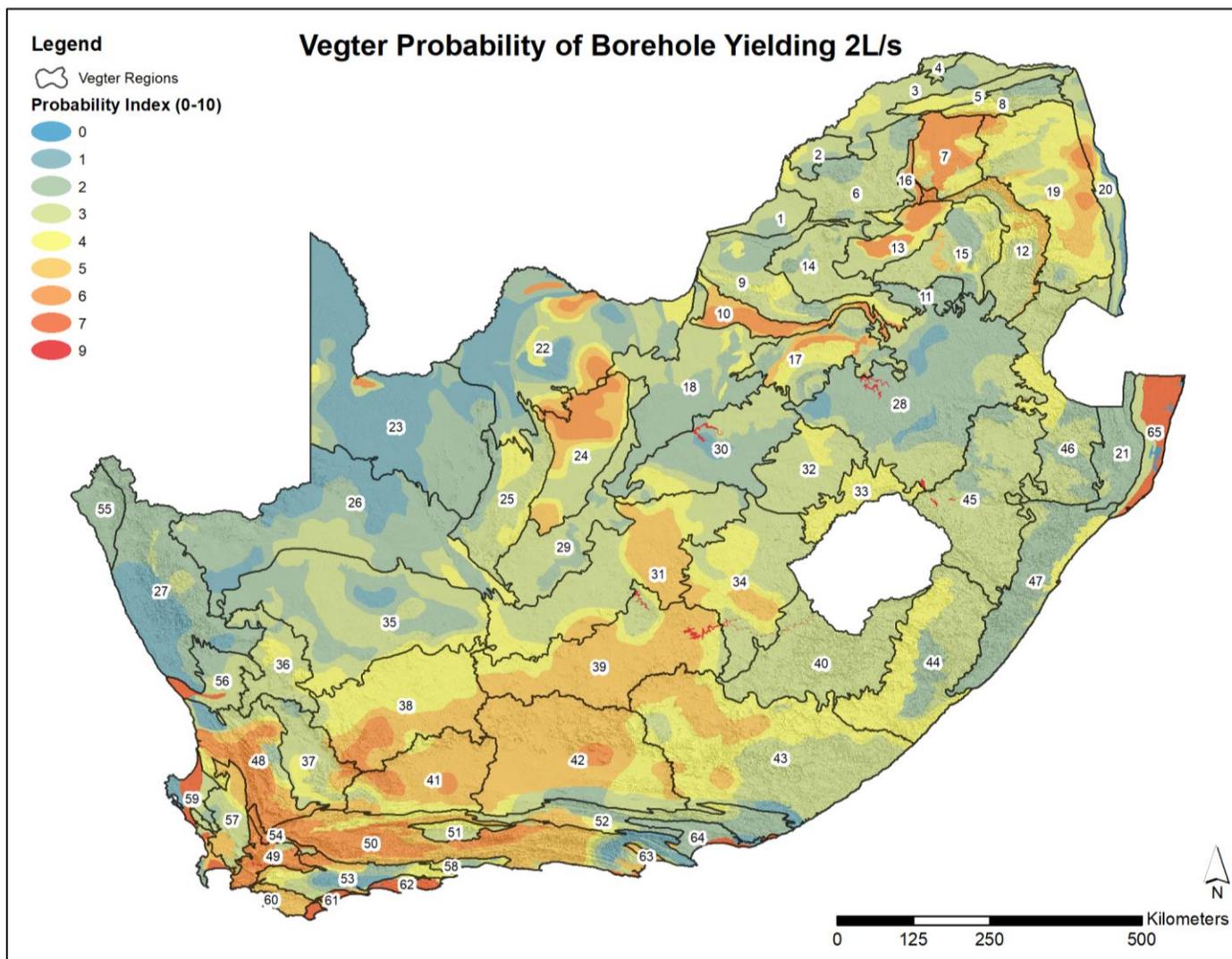


FIGURE 5 – VEGTER'S PROBABILITY OF DRILLING A BOREHOLE WITH YIELD OF 2 L/S AND MORE

### 1.3.3 Vegter's Recharge Map

The groundwater recharge map (groundwater recharge from rainfall) is an inset map on the Groundwater Resources of South Africa map and is shown in Figure 6. It should be noted that this is a generalised map (Seymore, 1994).

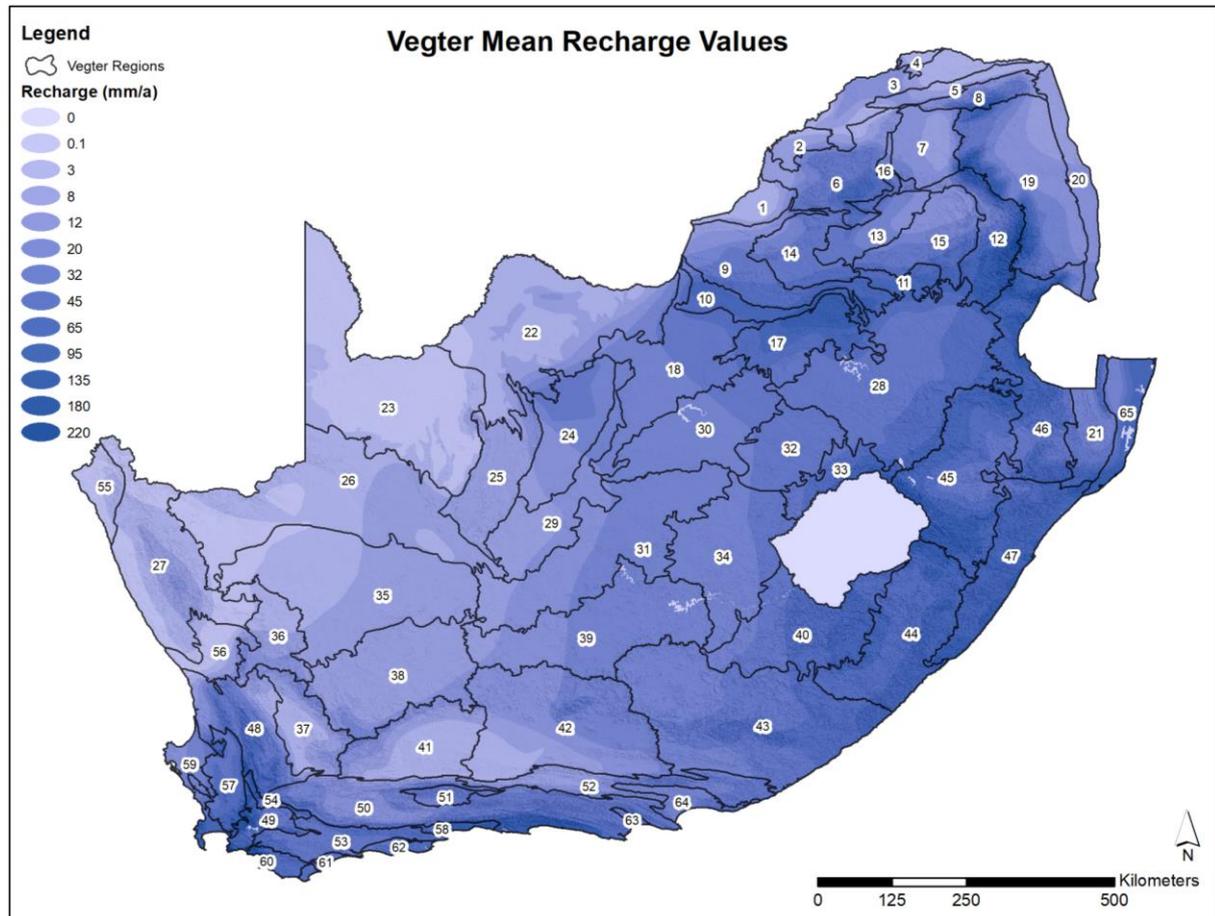


FIGURE 6 – VEGTER'S RECHARGE MAP

### 1.3.4 Vegter's Saturated Interstices

The saturated interstices is an inset map on the Groundwater Resources of South Africa map. The optimal drilling depth and indication of the storage coefficient is presented in Figure 7 and Figure 8 respectively. Only boreholes with values greater than zero for strike depth, borehole depth, yield and water depth were used (Seymore, 1994).

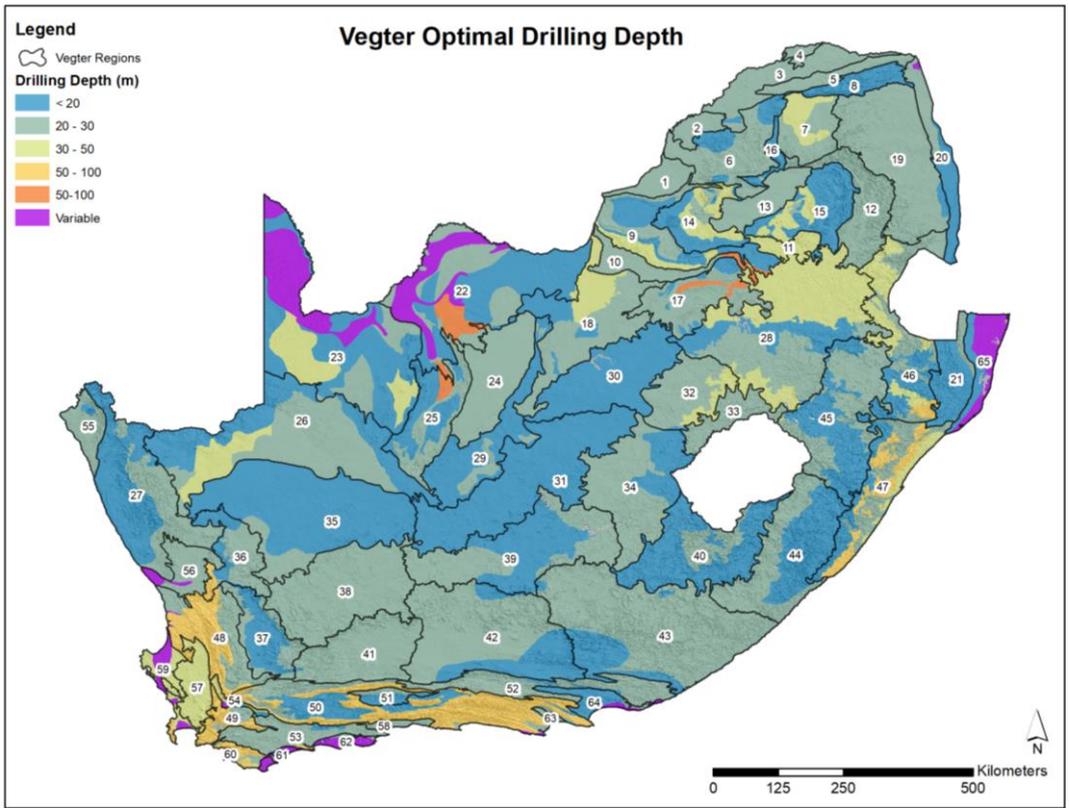


FIGURE 7 – VEGTER'S OPTIMAL DRILLING DEPTH

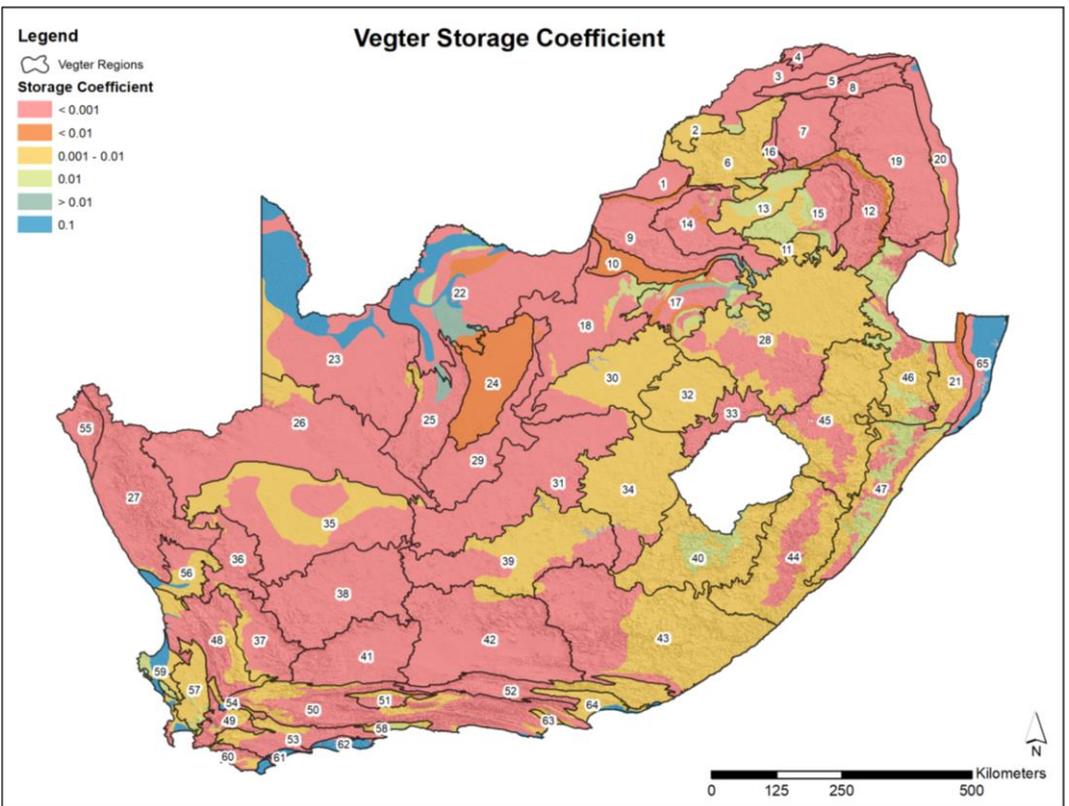


FIGURE 8 – VEGTER'S STORAGE COEFFICIENT

### 1.3.5 Vegter's water quality maps

The water quality map, represented by mean TDS, is presented in Figure 9. The TDS is used as macro indicator of water qualities. High TDS values do not necessarily refer to polluted water as the natural background can be high as well as is the case in the Kruger National Park to name one example.

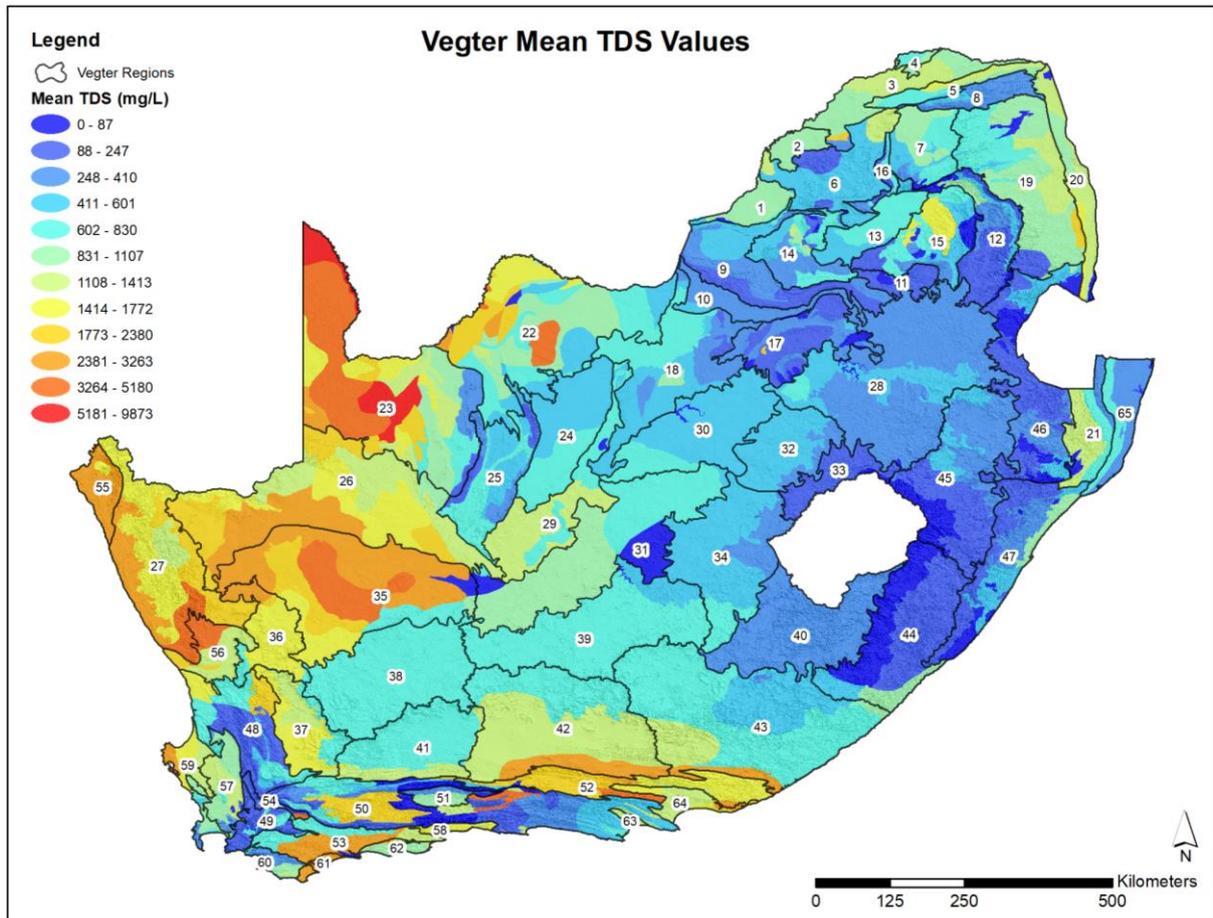


FIGURE 9 – VEGTER'S TDS MAP

The water character as determined by the major ions is presented in Figure 10 to Figure 13. This work was done by Milo Simonic and edited by JR Vegter. The four major water character types are considered and each map indicates the percentage of the relevant water character as it relates to the spatial distribution.

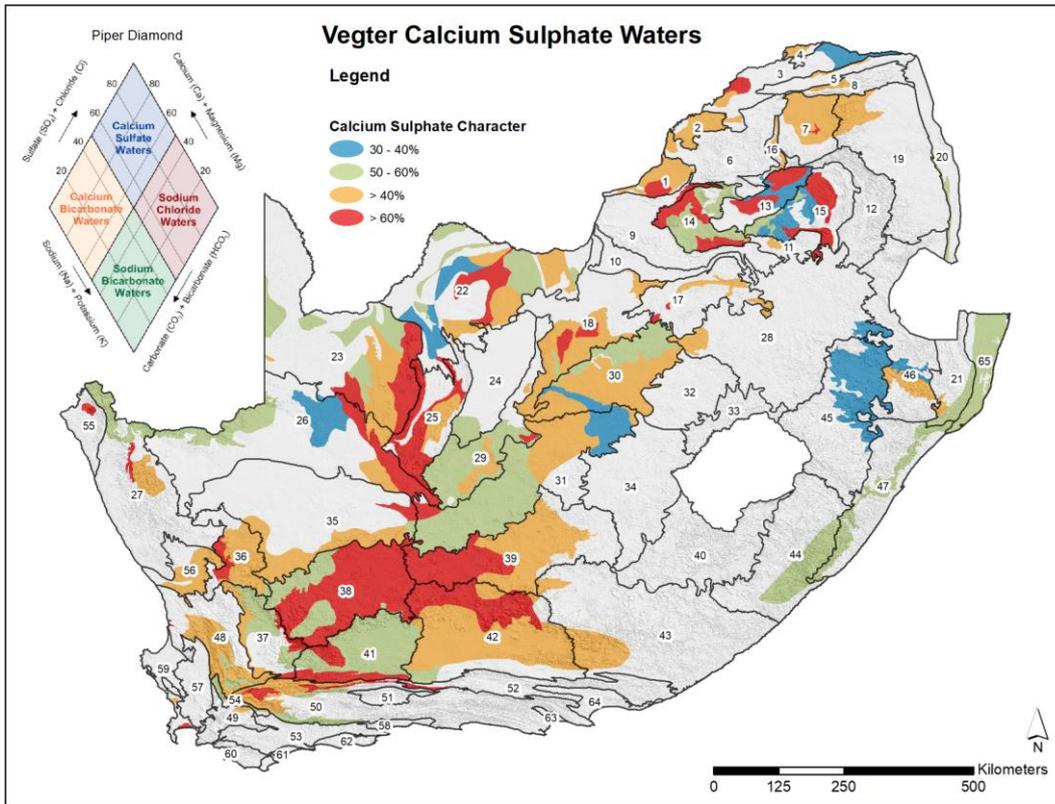


FIGURE 10 – VEGTER'S CALCIUM SULPHATE WATERS

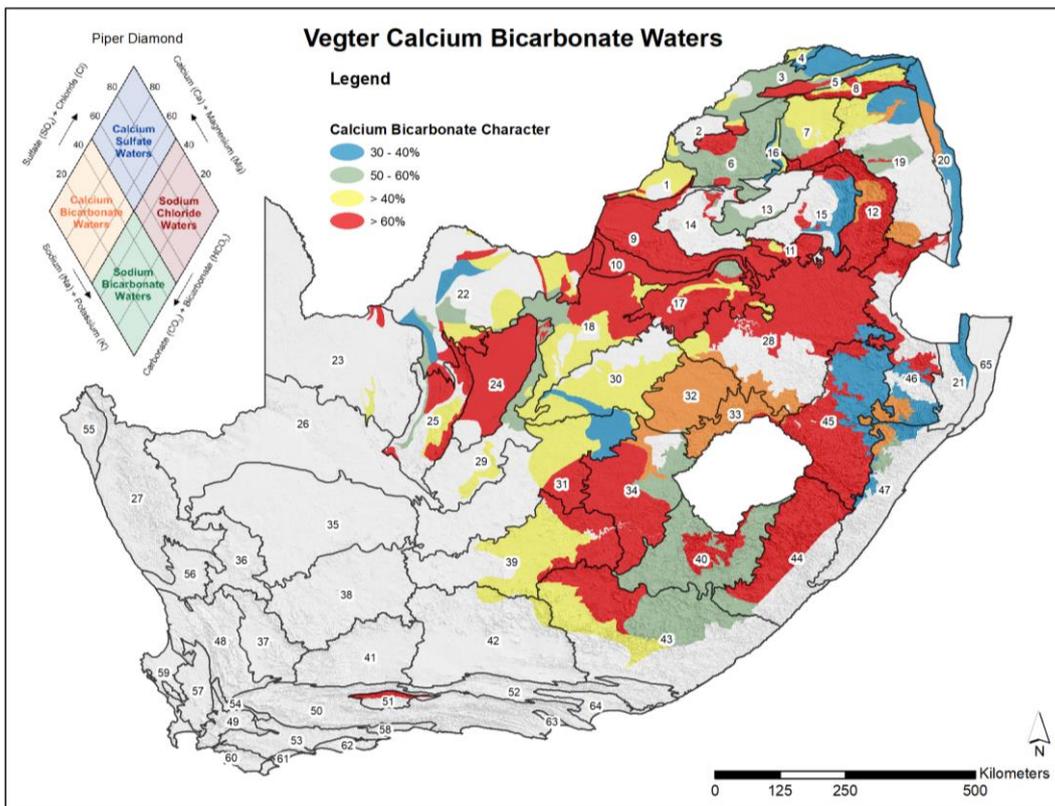


FIGURE 11 – VEGTER'S CALCIUM BICARBONATE WATERS

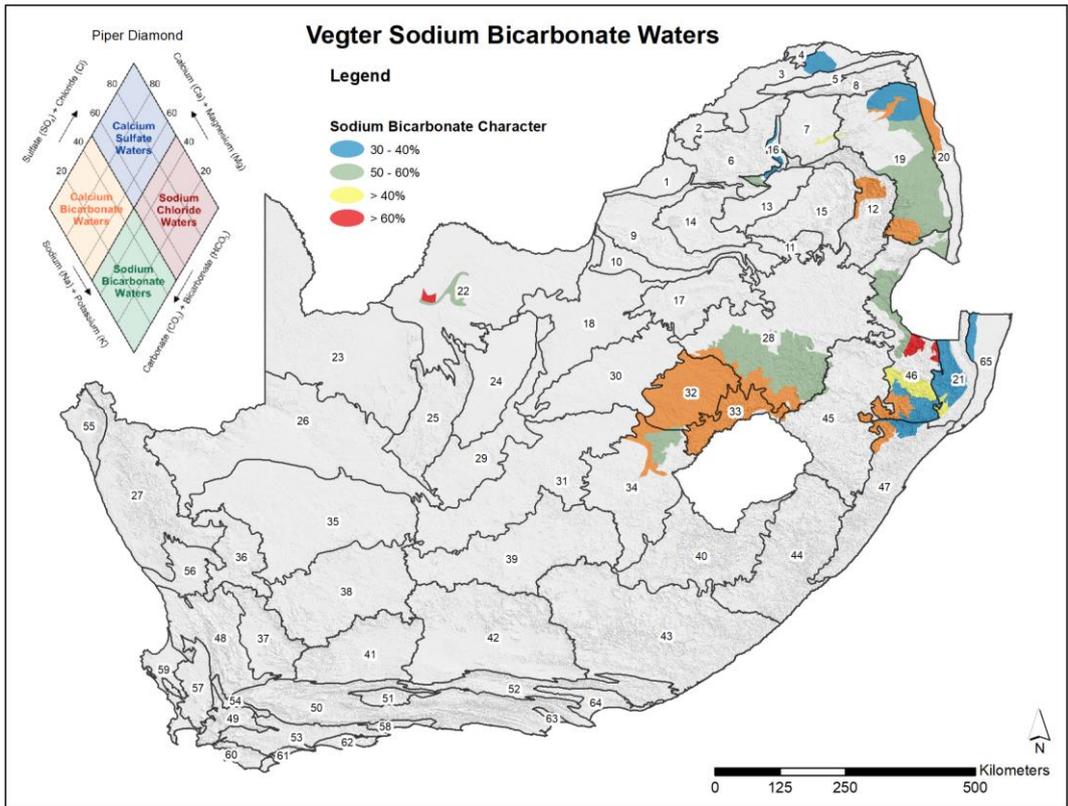


FIGURE 12 – VEGTER'S SODIUM BICARBONATE WATERS

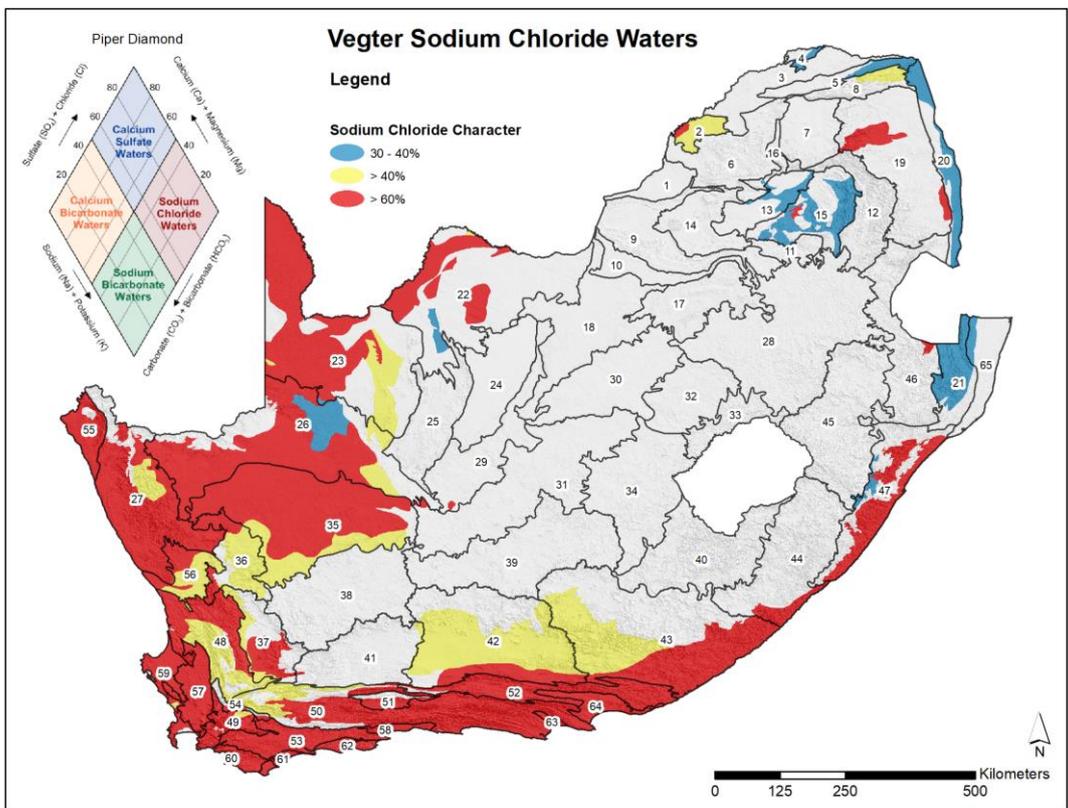


FIGURE 13 – VEGTER'S SODIUM CHLORIDE WATERS

## 2 DATA SOURCES

### 2.1 BACKGROUND

Public domain borehole information in South Africa is generally stored in the National Groundwater Archive (NGA) and the Groundwater Resources Information Project (GRIP) database of which both are centralized databases.

The NGA is a web enabled database system that allows capturing, viewing, modifying and extraction of groundwater related data by registered users. This database is maintained by Department of Water and Sanitation (DWS) and users obtain the required information by submitting a geo-request. The NGA database relies on users to upload captured borehole information to the database, but this is not happening for various reasons, e.g. data collection costs money and consultants keep borehole information to themselves as they see it as a competitive advantage and users often complain about the accessibility of data from the NGA through the web interface and therefore do not see the need to upload any of their borehole information.

The GRIP database was developed in the Limpopo province to capture borehole point data and is used for management of groundwater with emphasis on water availability and aquifer characteristics. In general, the GRIP database is regularly updated through term-contracts set up by DWS and ultimately this data should also be uploaded to the NGA. Some of the GRIP data already exist within the NGA, but to date not all GRIP data has been uploaded to the NGA.

Over and above the two mentioned public domain borehole databases, various users and consultancies maintain their own local borehole databases which are not available to the public. Often these databases also store client data which is subject to confidentiality agreements and therefore not uploaded to the NGA.

The historic NGA borehole distribution for the country is shown in Figure 14 (DWS, 2017) and it is clear that Limpopo has a higher borehole distribution when compared to the rest of the country, mainly due to the Limpopo GRIP project.

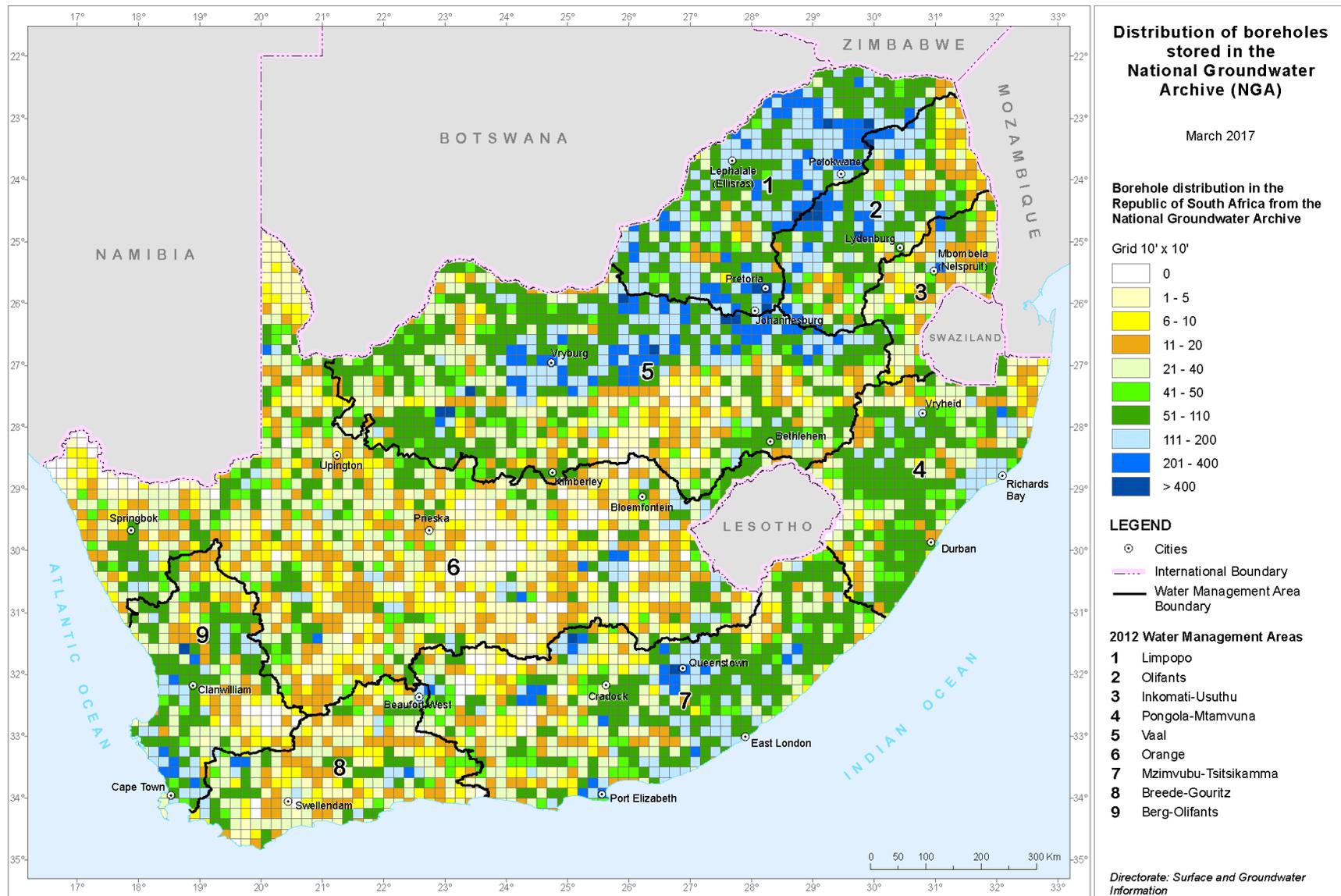


FIGURE 14 – NGA BOREHOLE DISTRIBUTION (DWS, 2017)

## 2.2 BOREHOLE INFORMATION

The NGA database structure is modelled on the Standard Descriptors for Geosites (DWAF, 2004) and therefore makes provision for various borehole related parameters to be stored. The common borehole parameters required for the Vegter analysis include:

- borehole positions (including elevation),
- borehole depths,
- water strike positions and
- borehole yields associated with the water strikes.

Additional borehole parameters are also available that can assist in the delineation of Vegter sub-regions:

- borehole logs and
- borehole chemistry (major anions and cations).

It should however be stressed that although these borehole parameters are available, both the temporal and spatial extent are limited for the majority of the country. This is partly due to the cost associated with monitoring programs and DWS only have selected boreholes that is subject to long-term monitoring. These boreholes are used to give a general indication of what is happening in the aquifer systems and can be accessed through the CHART application, although this data also resides in the NGA.

## 2.3 SPATIAL INFORMATION

Selected spatial information in the form of shape files and grid files is supplied to be used when the statistical analysis is performed. These contextual layers include:

- Surface elevation (SRTM90),
- surface geology (simplified),
- geological faults/structures,
- quaternary catchments,
- groundwater occurrence,
- aquifer rating and
- aquifer vulnerability.

The spatial information is assigned to a specific borehole when it is imported. These spatial datasets are only sampled at the specific borehole coordinates or within a specified buffer for example the case of lines representing faults and structures to obtain the required information.

## 2.4 ELEVATION DATA

Historically borehole positions were not obtained by GPS technology, simply because it did not exist at the time. Surveyors used the centroid of a farm boundary as coordinates to assign to any borehole within that particular boundary. This led to the phenomenon that some historic boreholes could have exactly the same coordinate and that these boreholes cannot be located in the field when using a GPS. This scenario is depicted in Figure 15 where the red cross indicates the actual position and the white cross represents the farm centroid.

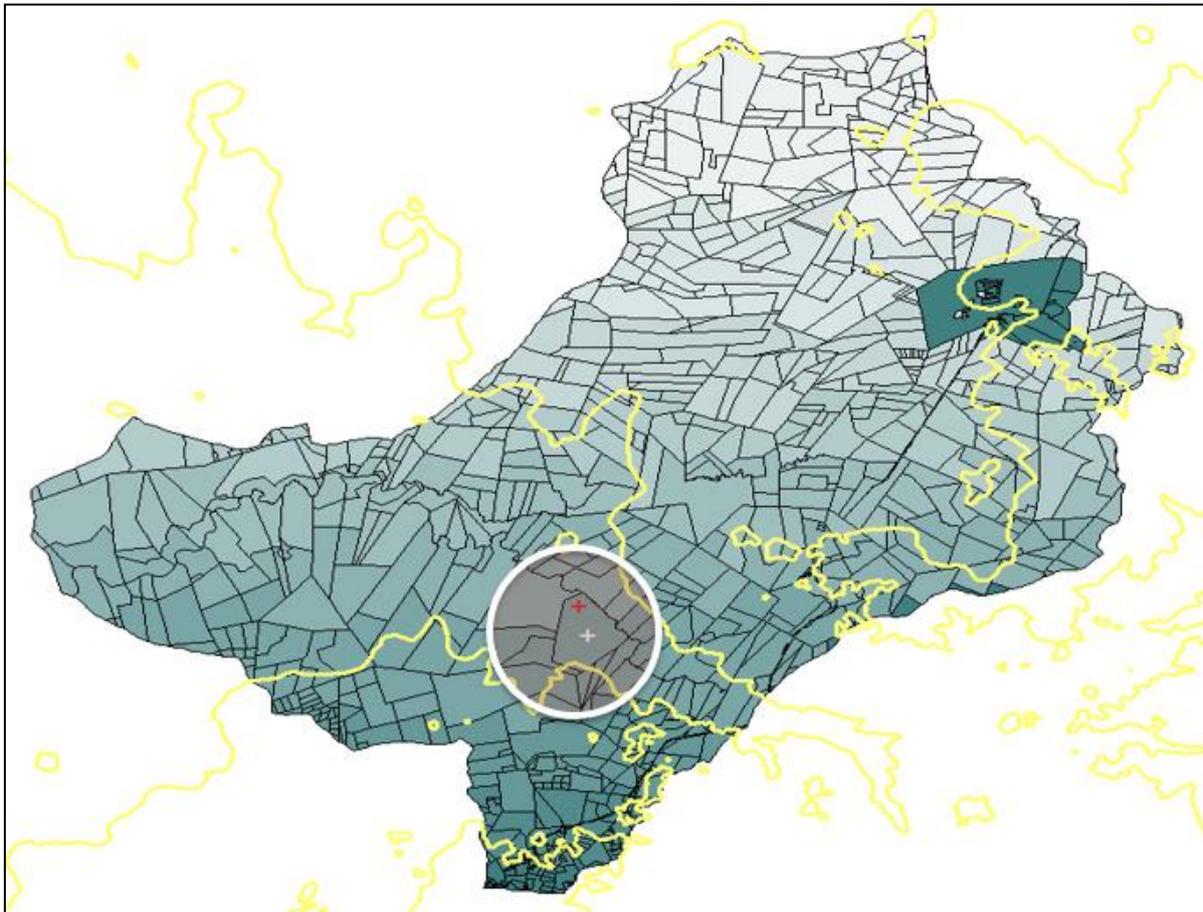


FIGURE 15 – EXAMPLE OF ACTUAL BOREHOLE COORDINATE AND FARM CENTROID COORDINATE

The elevation for these boreholes were assigned making use of an elevation contour map with 20 m intervals (yellow lines in Figure 15). Boreholes assigned to the nearest 20 m contour interval are easily identified when plotting borehole water level vs. elevation as shown in Figure 16. Variable water levels are seen at 20 m increments when considering the elevation axis.

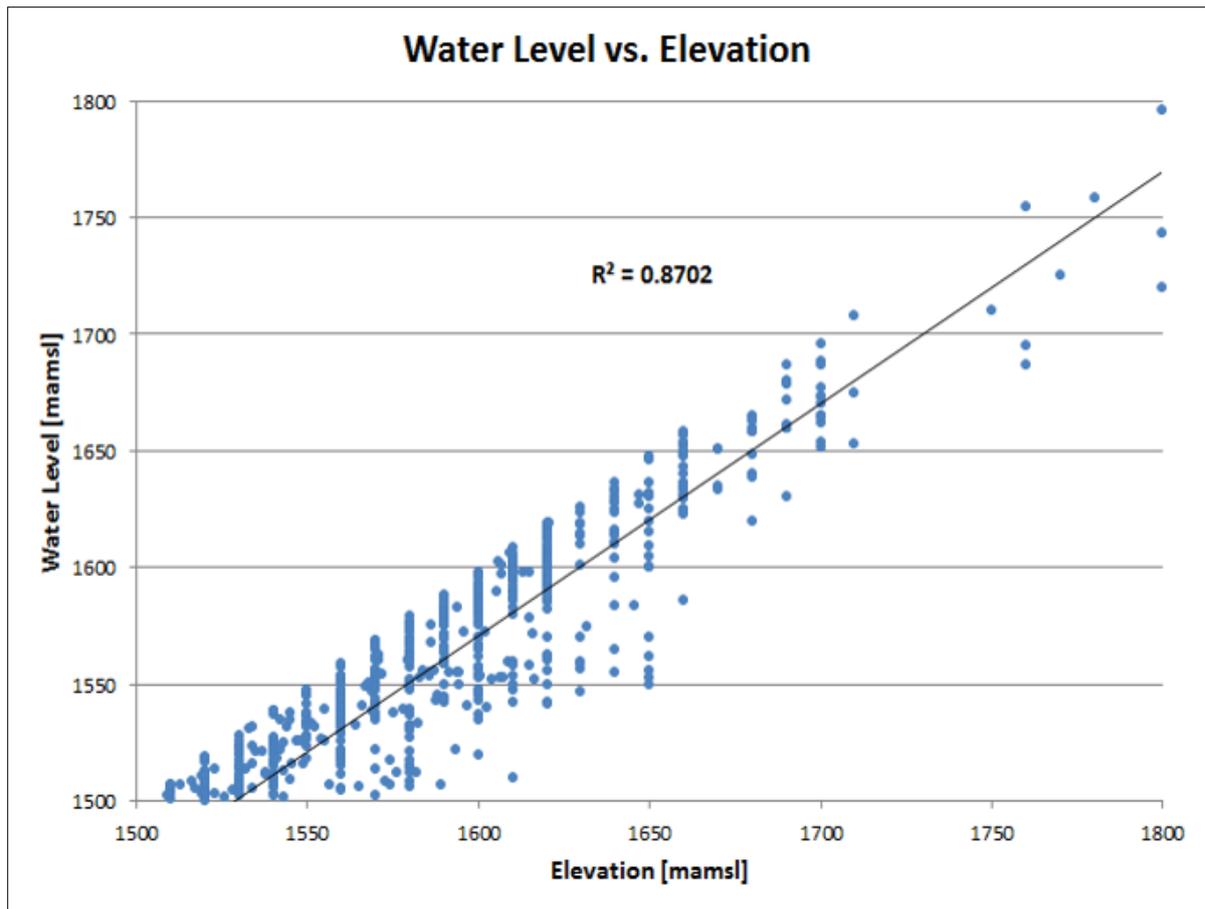


FIGURE 16 – BOREHOLE WATER LEVEL VS ELEVATION

The original 20 m contour line value will not be disregarded in the database, but the SRTM90 data will be used to update these elevations for analysis purposes. It should be noted that this cannot correct the representative position assigned to historic boreholes, but only give higher resolution in elevation. Furthermore, by making use of the STRM90 dataset as elevation reference, all boreholes used in the analysis are referenced against the same elevation source.

It should be noted that in the NGA database all boreholes assigned to the same farm centroid has been assigned a small offset of a few meters, so that the boreholes don't have exactly the same coordinate so that they are distinguishable in a GIS environment as shown in Figure 17.

It is not uncommon to come across borehole localities as depicted in Figure 17, but it should be noted that this is not the physical drill positions and that it is unlikely that one will be able to identify these boreholes in the field. Even though this is not an accurate representation of reality, it does inform geohydrological analysis at regional scale.

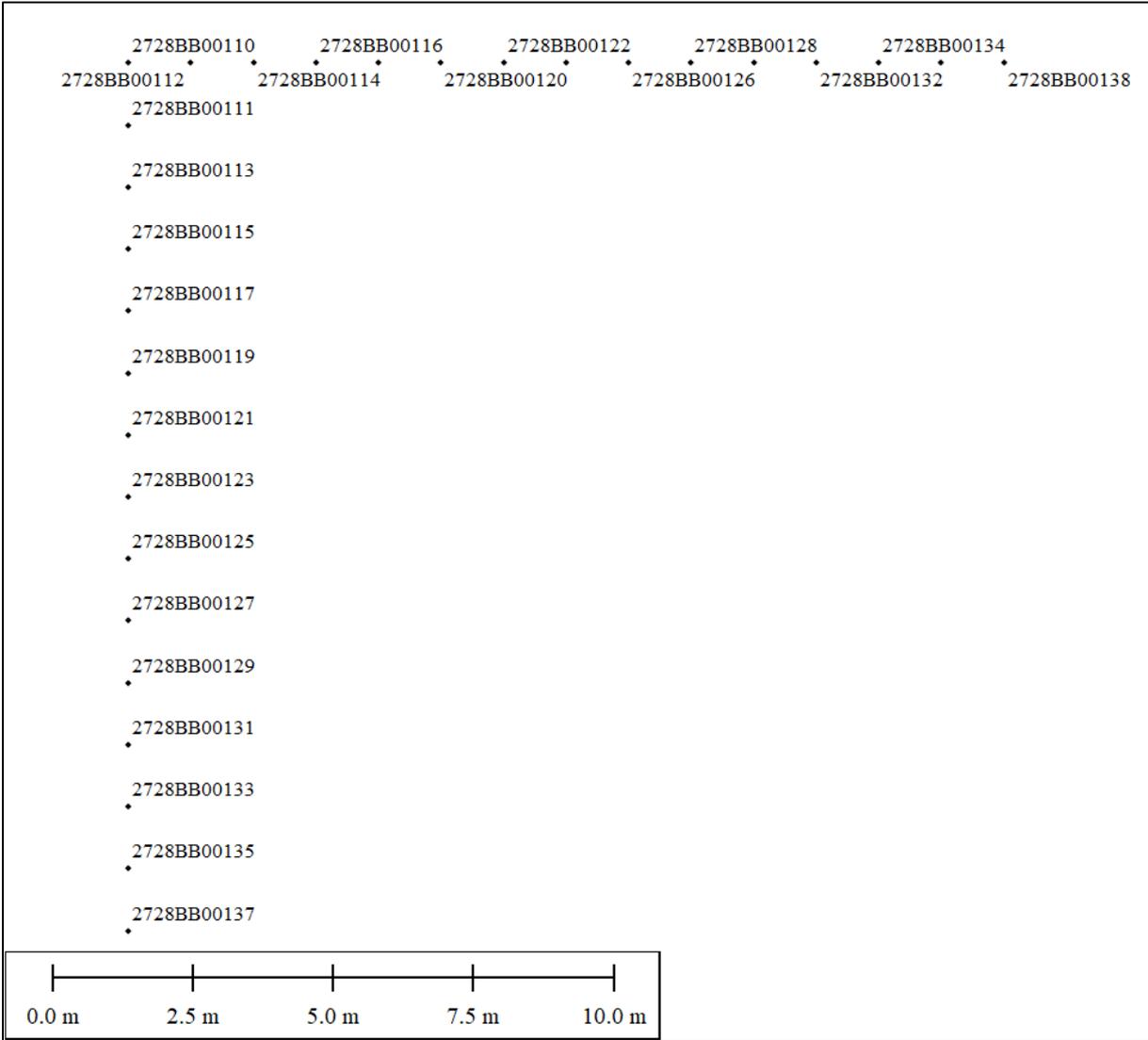


FIGURE 17 – MULTIPLE BOREHOLES ASSIGNED TO SAME FARM CENTROID WITH OFFSET APPLIED

2.5 ONGOING DATA CAPTURE

The various databases are continuously updated and some difficulty is experienced when requesting borehole information over large areas. Update functionality is offered as part of the final product to allow users to continue with these updates and also include their own borehole information, which is not part of the public domain databases.

### 3 DATABASE DESIGN

#### 3.1 DATABASE SOFTWARE

The database is developed using SQLite which is an in-process library that implements a self-contained, serverless, zero-configuration, transactional SQL database engine. The code for SQLite is in the public domain and is therefore free for use for any purpose, commercial or private. The final software package does not require the installation of software or software drivers, making it possible for the software to be run from a memory stick.

#### 3.2 RATIONAL DATABASE

The entity relationship diagram of the database is presented in Figure 18. The design consists of five tables where **SiteID** is the primary key to the *Site* table and the foreign key to the remaining tables. The *Site* table contains the borehole name, position and all common layer attributes and has a 1-to-many relationship with the other tables as shown in Figure 18. The sections that follow provides a description of fields in the tables. Note that the data types are not repeated in the table and is shown in the entity relationship diagram. The actual database contains additional tables used internally by the software tool and therefore are not discussed here as only the relevant data tables are considered for the statistical analysis.



FIGURE 18 – ENTITY RELATIONSHIP DIAGRAM

The data considered for analysis is based on the minimum data required for the Vegter analyses as well as what the general data users will have available for boreholes.

### 3.2.1 Site Table

The field descriptions of the site (*Site*) table are presented in Table 2. The **SrcID** field is used to distinguish between the various data sources to allow the Vegter analysis to be done with only selected data sources.

TABLE 2 – SITE TABLE FIELD DESCRIPTIONS

<i>Field Name</i>	<i>Description</i>
<b>SiteID</b>	Unique Site ID
<b>SiteName</b>	Descriptive site name
<b>Longitude</b>	Longitude in decimal degrees
<b>Latitude</b>	Latitude in decimal degrees
<b>Elevation</b>	Elevation (mamsl)
<b>SRTM90</b>	DEM elevation (mamsl)
<b>Depth</b>	Depth of borehole (m)
<b>Collar</b>	Collar height of borehole (m)
<b>Quat</b>	Quaternary catchment
<b>Geology</b>	Simplified geology code
<b>YieldSym</b>	Yield map symbol
<b>YieldDesc</b>	Yield map description (occurrence)
<b>Aquifer</b>	Aquifer Rating (1-10)
<b>DRASTIC</b>	Aquifer vulnerability (0-200)
<b>SrcID</b>	Data source ID

### 3.2.2 Static Level Table

The field descriptions of the static water level (*Static*) table are presented in Table 3. It is important to note that only static water levels are considered in the analysis and therefore pumped water levels should not be considered, as the natural state of the system is desired. Time series water level data is accommodated through the use of the **Date** field.

TABLE 3 – STATIC WATER LEVEL TABLE FIELD DESCRIPTIONS

<i>Field Name</i>	<i>Description</i>
<b>SiteID</b>	Unique Site ID
<b>Date</b>	The date of measurement
<b>Level</b>	Water level (mbgl)
<b>SrcID</b>	Data Source ID

### 3.2.3 Water Strike Table

The field descriptions of the water strike (*Strike*) table are presented in Table 4. Generally, the blow yield is determined when the borehole was drilled and therefore is not considered as time series data. However, boreholes can be drilled deeper or redeveloped and new blow yield values can be obtained, therefore the **Date** field is part of the table design to accommodate any new blow yield values.

TABLE 4 – WATER STRIKE TABLE FIELD DESCRIPTIONS

<i>Field Name</i>	<i>Description</i>
<b>SiteID</b>	Unique Site ID
<b>Date</b>	The date of measurement
<b>Depth</b>	Depth of water strike
<b>Yield</b>	Blow yield (L/s)
<b>SrcID</b>	Data Source ID

### 3.2.4 Chemistry Table

The field descriptions for the chemistry (*Chem*) table are presented in Table 5. Time series water chemistry data is accommodated by using the date field.

TABLE 5 – CHEMISTRY TABLE FIELD DESCRIPTIONS

<i>Field Name</i>	<i>Description</i>
<b>SiteID</b>	Unique Site ID
<b>Date</b>	The date of measurement
<b>Depth</b>	Depth of sampling (mbgl)
<b>EC</b>	Electrical conductivity (mS/m)
<b>pH</b>	pH
<b>Ca</b>	Calcium (mg/L)
<b>Mg</b>	Magnesium (mg/L)
<b>Na</b>	Sodium (mg/L)
<b>K</b>	Potassium (mg/L)
<b>Cl</b>	Chloride (mg/L)
<b>SO4</b>	Sulphate (mg/L)
<b>MAIk</b>	Methyl Orange Alkalinity (mg/L)
<b>PAIk</b>	Phenolphthalein Alkalinity (mg/L)
<b>SrcID</b>	Data Source ID

### 3.2.5 Borehole Log Table

The field descriptions of the borehole log (*Log*) table are presented in Table 6. Generally, the borehole log is determined when the borehole is drilled and therefore is not considered as time series data. However, boreholes can be drilled deeper and new borehole log values can be obtained, therefore the **Date** field is part of the table design to accommodate any new borehole values.

TABLE 6 – BOREHOLE LOG TABLE FIELD DESCRIPTIONS

<i>Field Name</i>	<i>Description</i>
<b>SiteID</b>	Unique Site ID
<b>Date</b>	The date of measurement
<b>DepthTop</b>	Start of the specified lithology measured from surface (m)
<b>DepthBot</b>	End of the specified lithology measured from surface (m)
<b>LithoCode</b>	Lithological code used for lookup of lithology details
<b>SrcID</b>	Data Source ID

### 3.3 DATABASE STATISTICS AND PREVIEWS

This section provides some basic database statistic as it relates to the populated databases as well as some spatial representation of the existing database.

#### 3.3.1 Basic Statistics

A summary for the current two databases is presented in Table 7 and the distribution of the various borehole densities are presented in Figure 19 to Figure 25. Note that the borehole counts are presented on the 1:50,000 grid to avoid clutter by plotting each borehole. The borehole count colour scale is the same for each of the presented parameters.

TABLE 7 – DATABASE POPULATION SUMMARY

<i>Count</i>	<i>Description</i>	<i>NGA</i>	<i>GRIP</i>
Borehole (Figure 19)	Total borehole count in each database. It should be noted that there exists an overlap between the two databases and that the final borehole count will not be the sum of the two. Roughly 11% of the GRIP boreholes are also present in the NGA database.	276,333	26,912
Borehole Water Level (Figure 20)	The number of boreholes that have one or more water levels associated with it.	129,615	10,333
All Water Level	The total number of water levels available in the database.	2,320,160	11,022
Borehole EC (Figure 21)	The number of boreholes that have one or more EC measurements	33,153	7,884
All EC	The total number of EC measurements available in the database.	55,110	12,057
Borehole Chemistry (Figure 22)	The number of boreholes that have one or more chemistry analysis associated with it (2008 NGA data as no new chemistry data could be obtained in 2018).	12,173	7,050
All Chemistry	The total number of chemistry analysis available in the database.	26,594	8,701
Borehole Yield (Figure 23)	The number of boreholes that have one or more yield values associated with water strikes.	75,327	10,002
All Yield	The total number of water strikes with yield available in the database.	2,894,751	10,425
Water Strikes (Figure 24)	The number of boreholes that have one or more water strike associated with it.	112,880	2,386
All Strikes	The total number of water strikes available in the database.	4,029,258	2,680
Borehole Logs (Figure 25)	Total number of boreholes that have a borehole log.	147,837	3,403

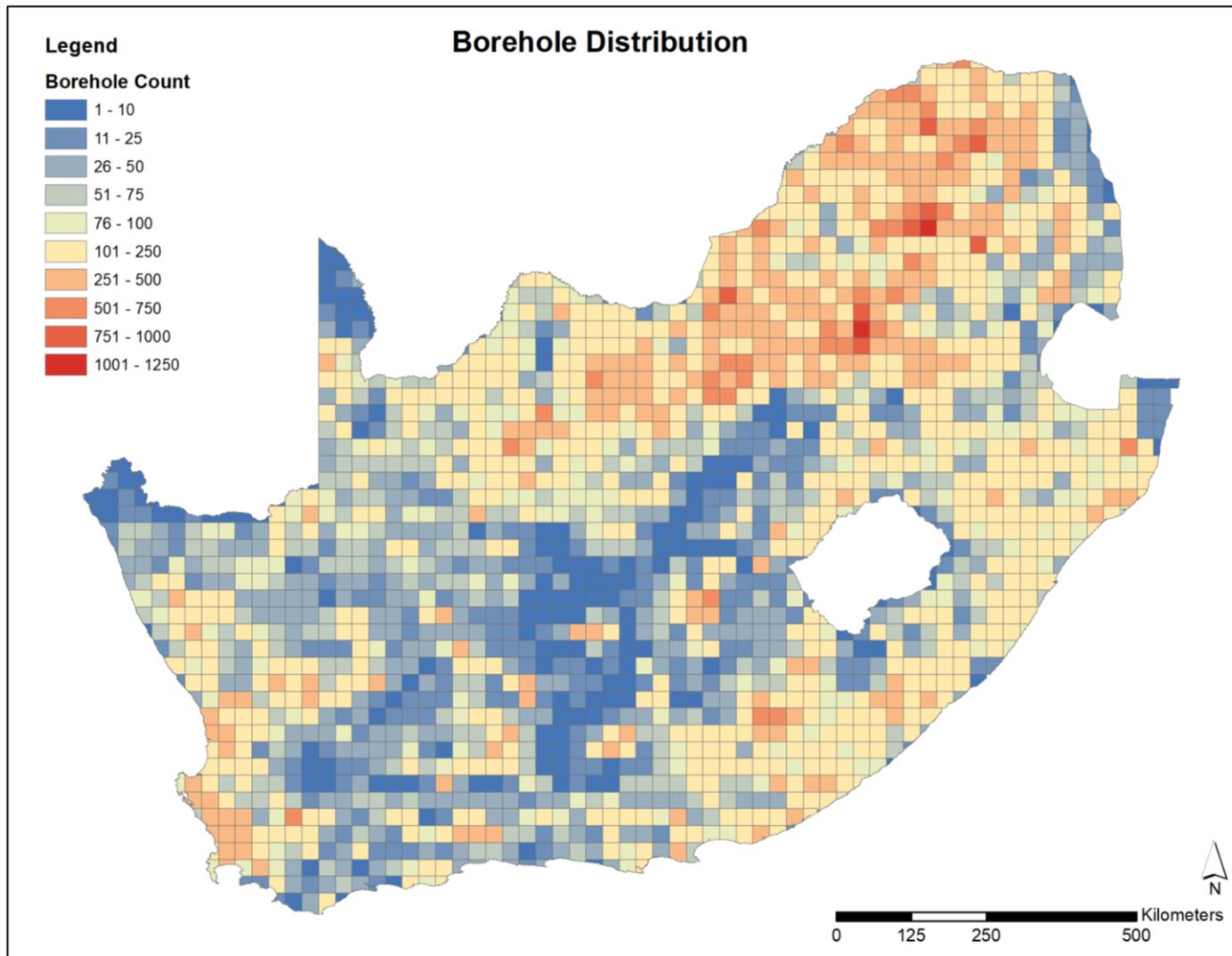


FIGURE 19 – DATABASE BOREHOLE DISTRIBUTION

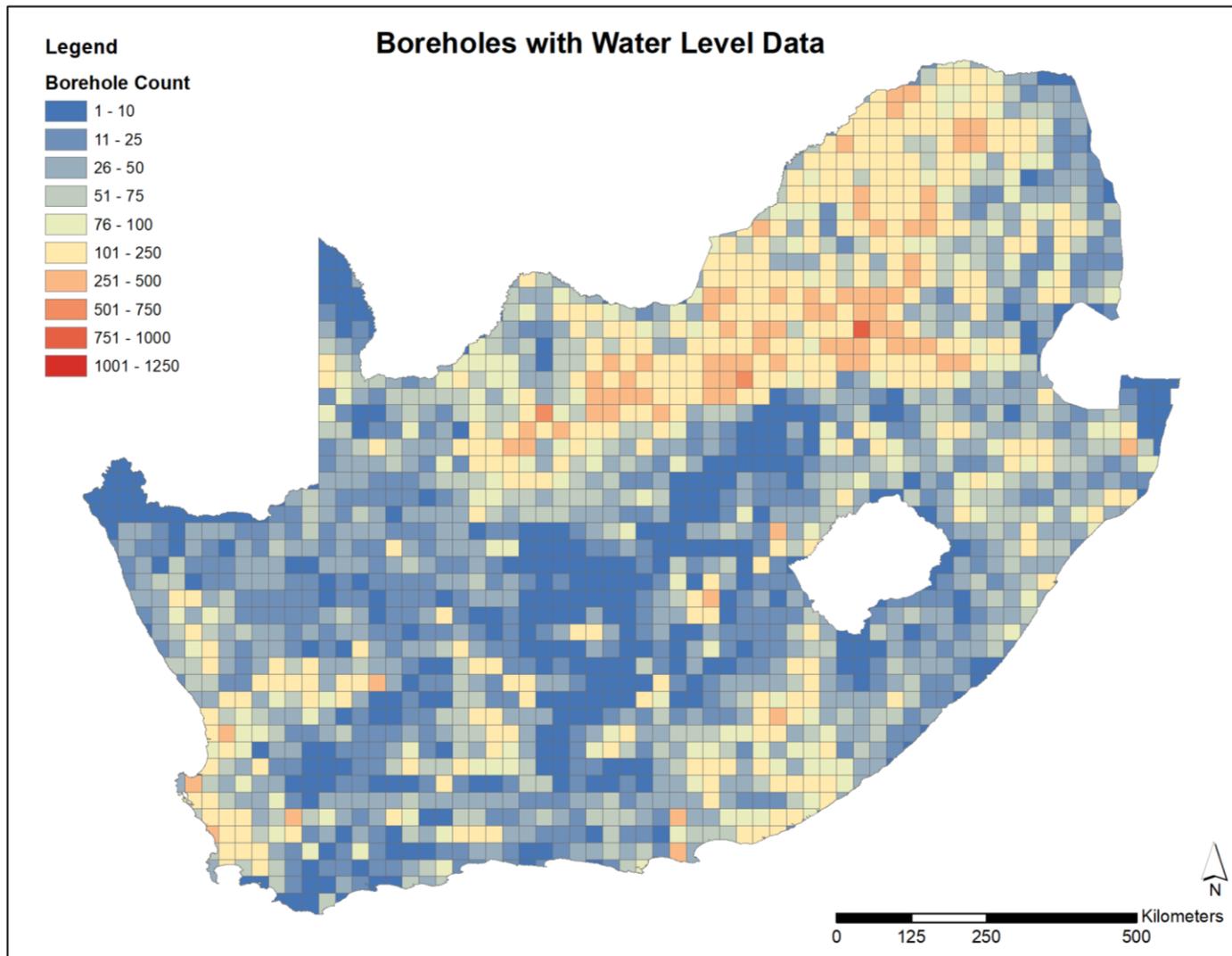


FIGURE 20 – DATABASE BOREHOLES WITH WATER LEVELS

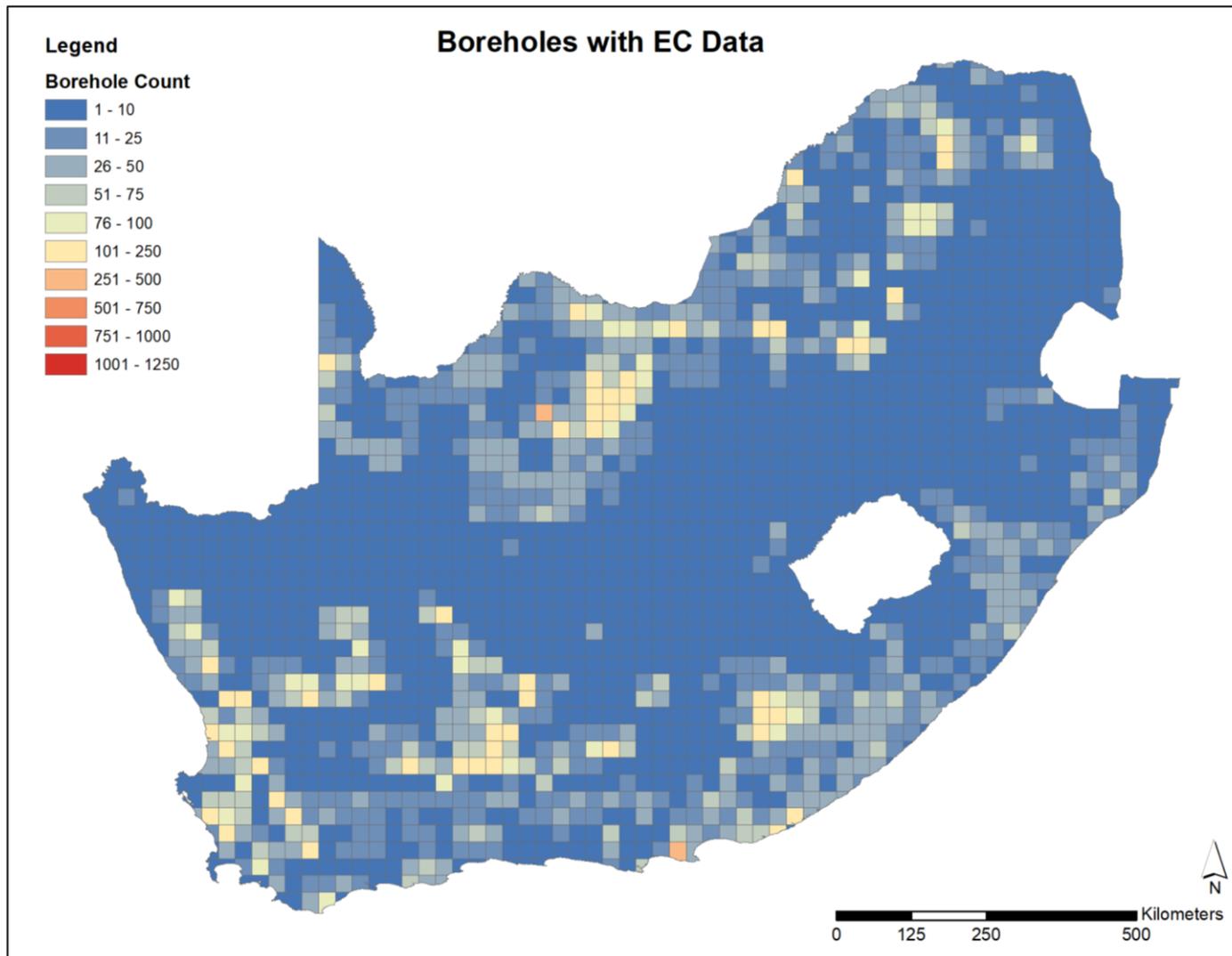


FIGURE 21 – DATABASE BOREHOLES WITH EC VALUES

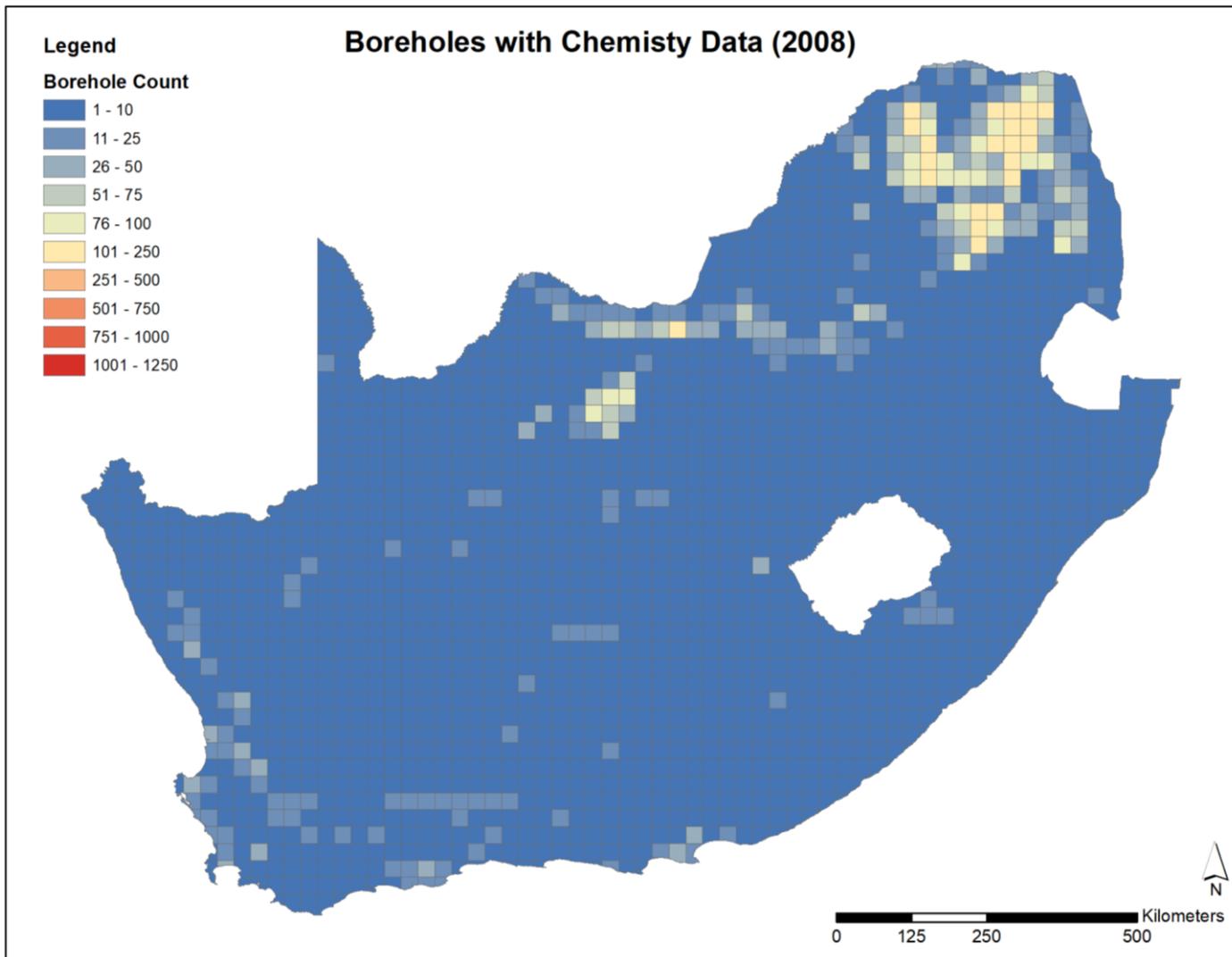


FIGURE 22 – DATABASE BOREHOLES WITH CHEMISTRY VALUES (2008)

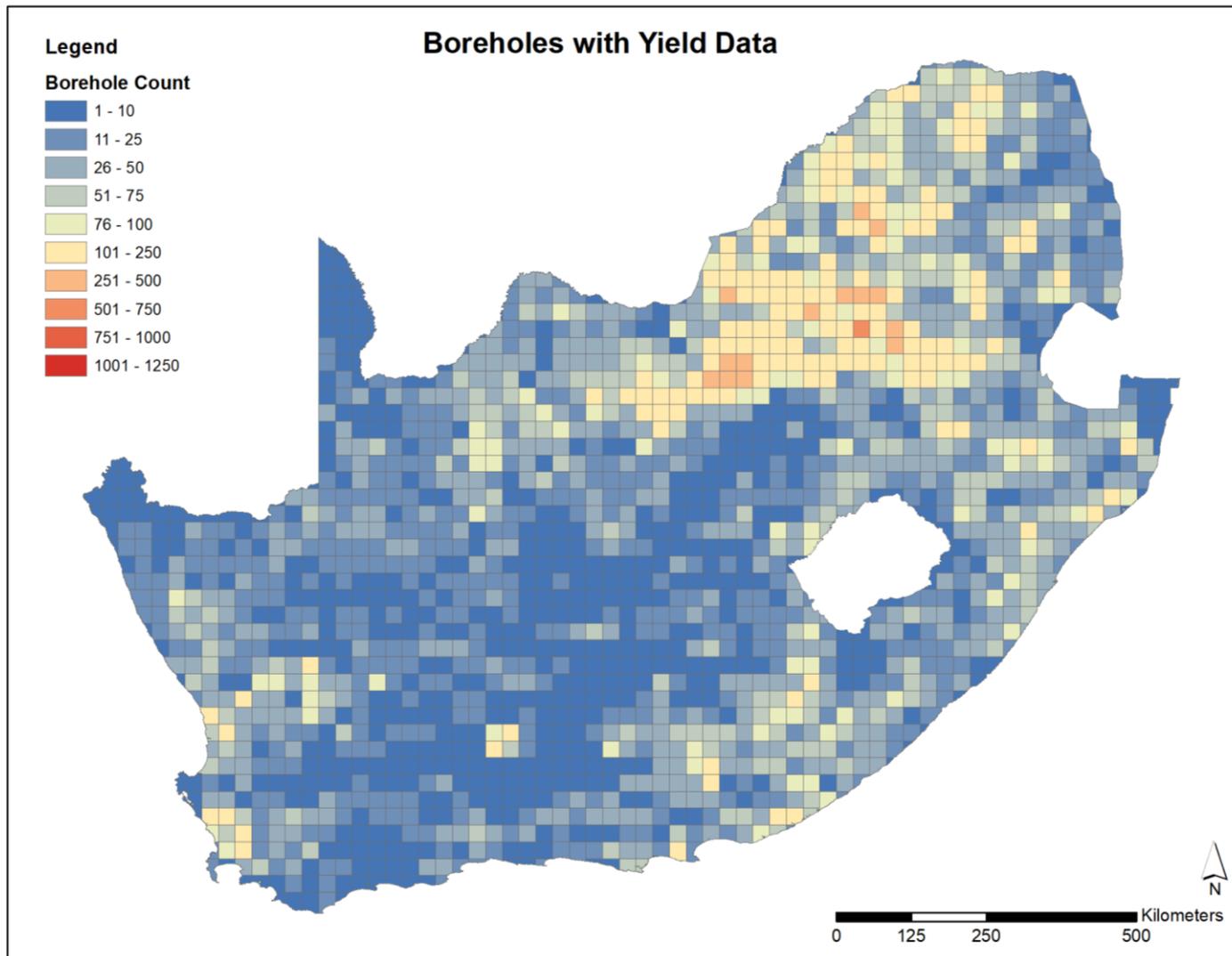


FIGURE 23 – DATABASE BOREHOLES WITH YIELD VALUES

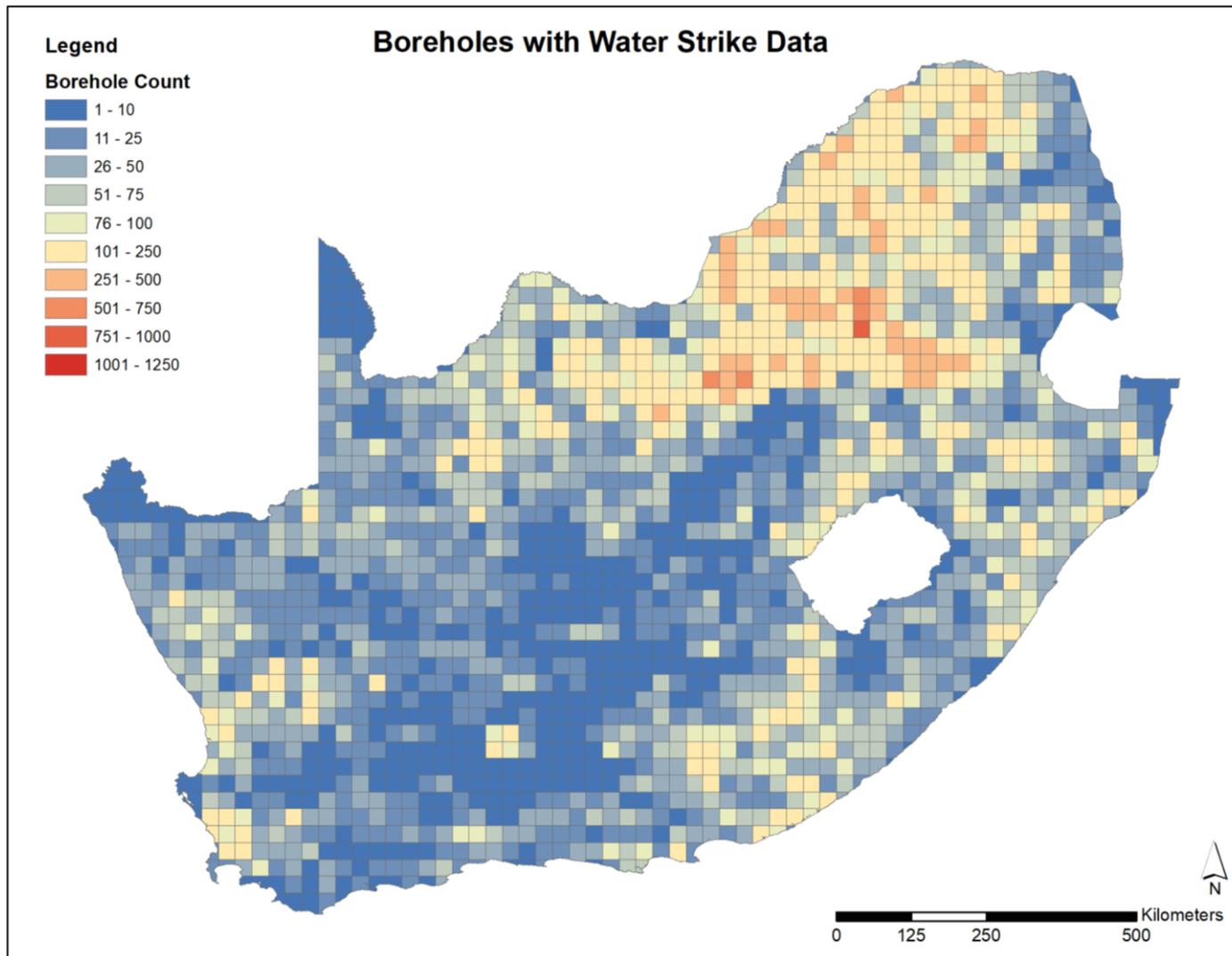


FIGURE 24 – DATABASE BOREHOLES WITH WATER STRIKE VALUES

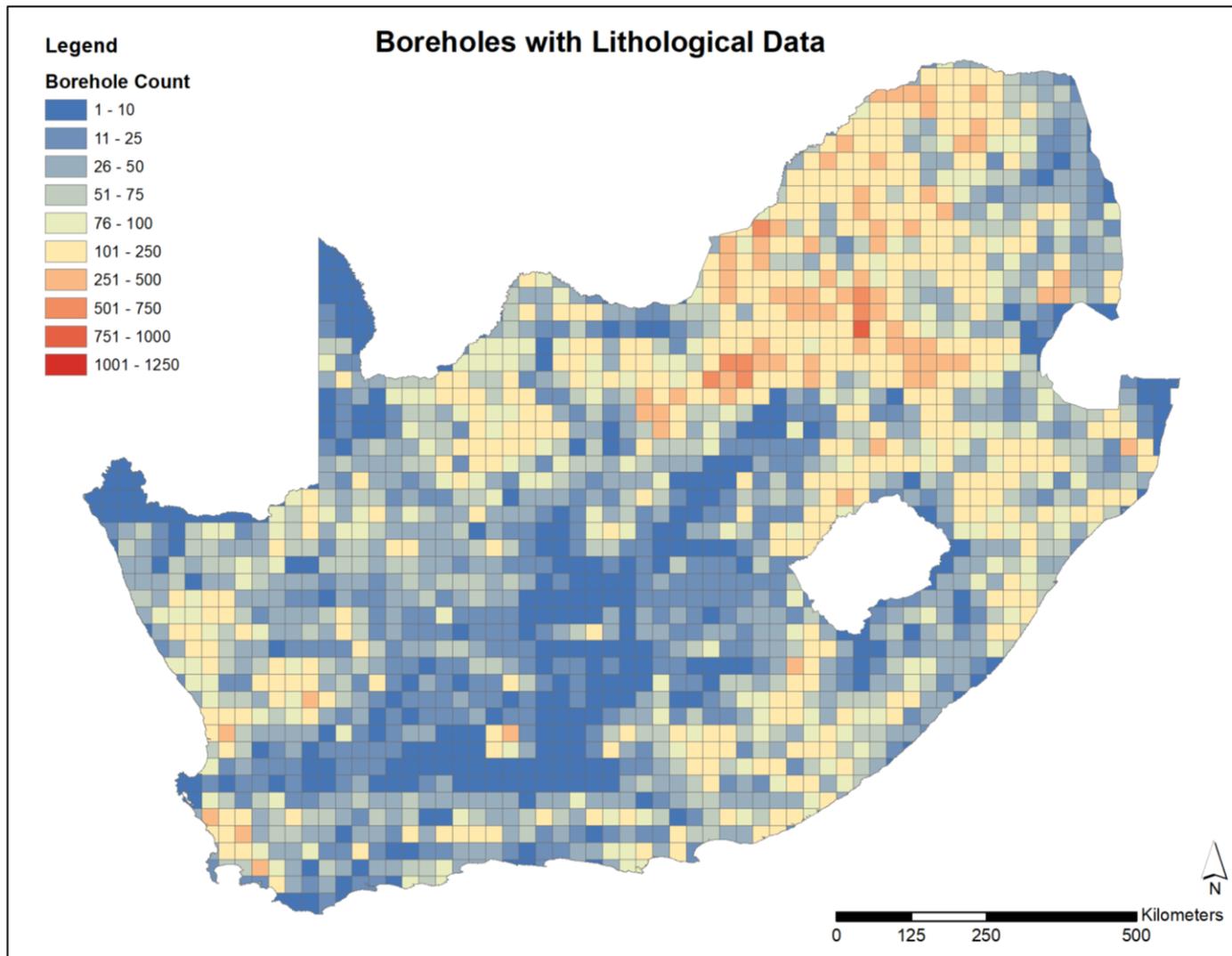


FIGURE 25 – DATABASE BOREHOLES WITH BOREHOLE LOGS

### 3.3.2 *Spatial Datasets*

This section presents a selection of the available spatial datasets included in the database.

#### 3.3.2.1 *Surface Elevation (SRTM90)*

The SRTM90 dataset is used to refine the borehole elevations based on the 20 m elevation contour intervals used historically as discussed earlier in this document. The SRTM90 data is presented in Figure 26.

#### 3.3.2.2 *Simplified surface geology and geological and structural lineaments*

A simplified surface geology map of South Africa is used in the Vegter analysis. The aquifer classification based on the aforementioned map is shown in Figure 27. The geological and structural lineaments are shown in Figure 28 which were generated from the 1:50,000 geological maps.

#### 3.3.2.3 *Groundwater occurrence*

The groundwater occurrence map based on the geohydrological map of South Africa is presented in Figure 29. The basis for this map is also the simplified geological map of South Africa.

#### 3.3.2.4 *Aquifer vulnerability*

Aquifer vulnerability refers to the tendency or likelihood for contamination to reach a specified position in the groundwater system after introduction at some location above the uppermost aquifer. The DRASTIC aquifer vulnerability method makes use of seven (7) factors to calculate the vulnerability index value (Appendix A):

- Depth to groundwater (D) – determines the maximum distance contaminants travel before reaching the aquifer;
- Net recharge (R) – the amount of water that is able to travel from ground surface to the water table;
- Aquifer (A) – the composition of the aquifer material;
- Soil media (S) – the uppermost portion of the unsaturated zone;
- Topography (T) – the slope of the ground surface;

- Impact of vadose zone (I) – the type of material present between the bottom of the soil zone and water table; and
- Hydraulic conductivity of the aquifer (C) – indicates the aquifer’s ability to allow for the flow of water to occur.

This vulnerability index is used to determine the aquifer’s vulnerability to pollution and the index range from 1 to 200, where 200 represents the theoretical maximum aquifer vulnerability.

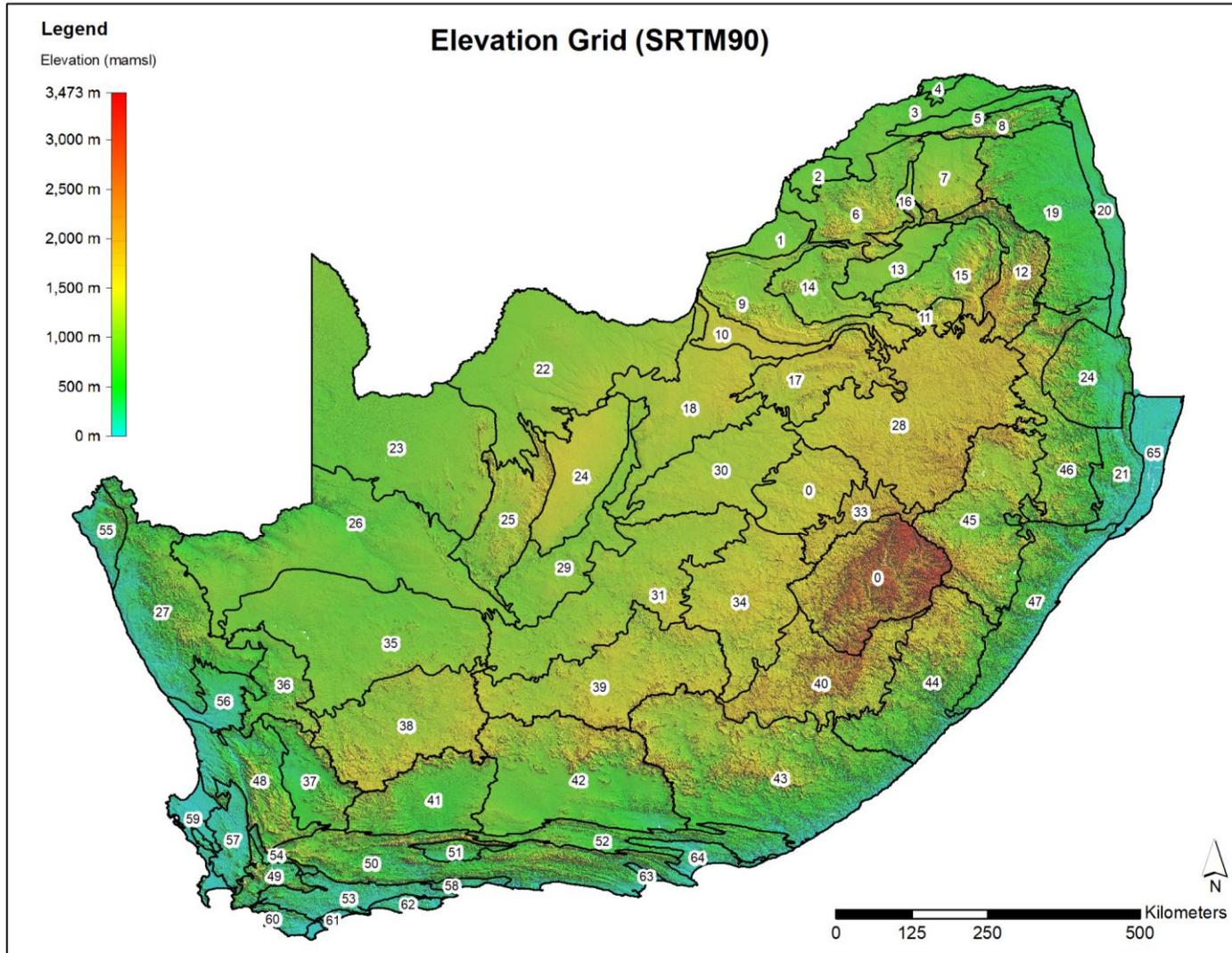


FIGURE 26 – SRTM90 ELEVATION GRID OF SOUTH AFRICA

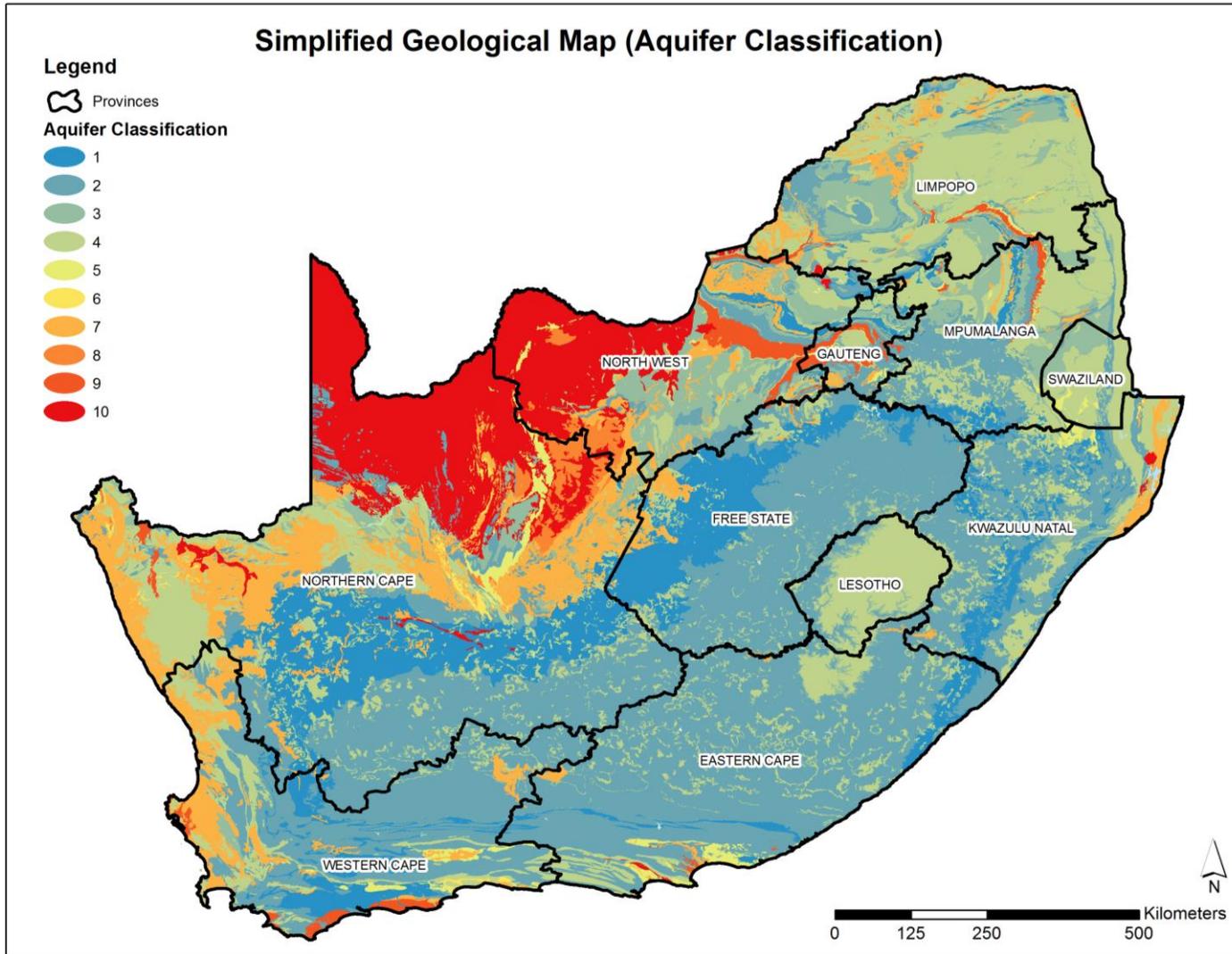


FIGURE 27 – SIMPLIFIED GEOLOGY MAP (AQUIFER CLASSIFICATION)

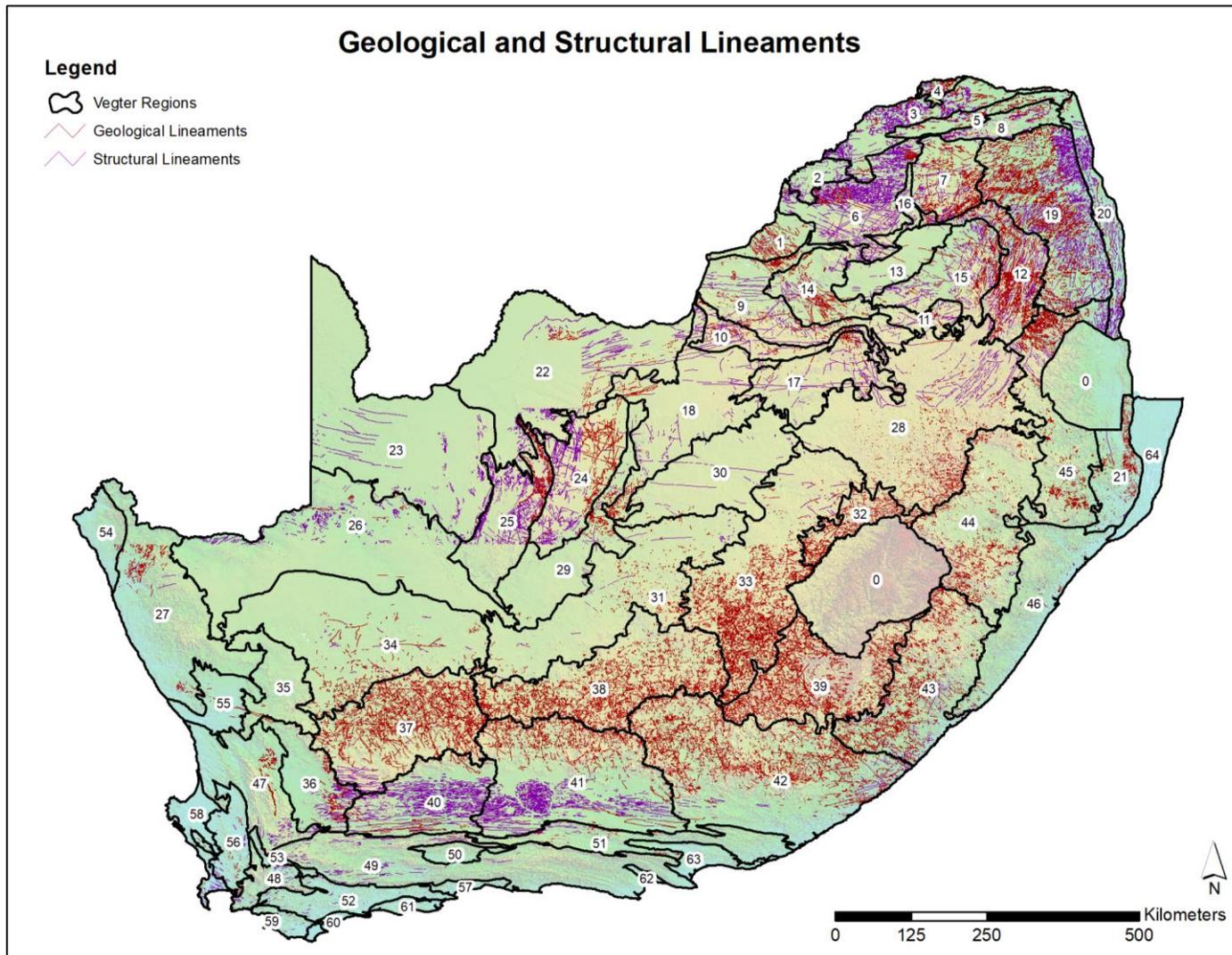


FIGURE 28 – GEOLOGICAL AND STRUCTURAL LINEAMENTS

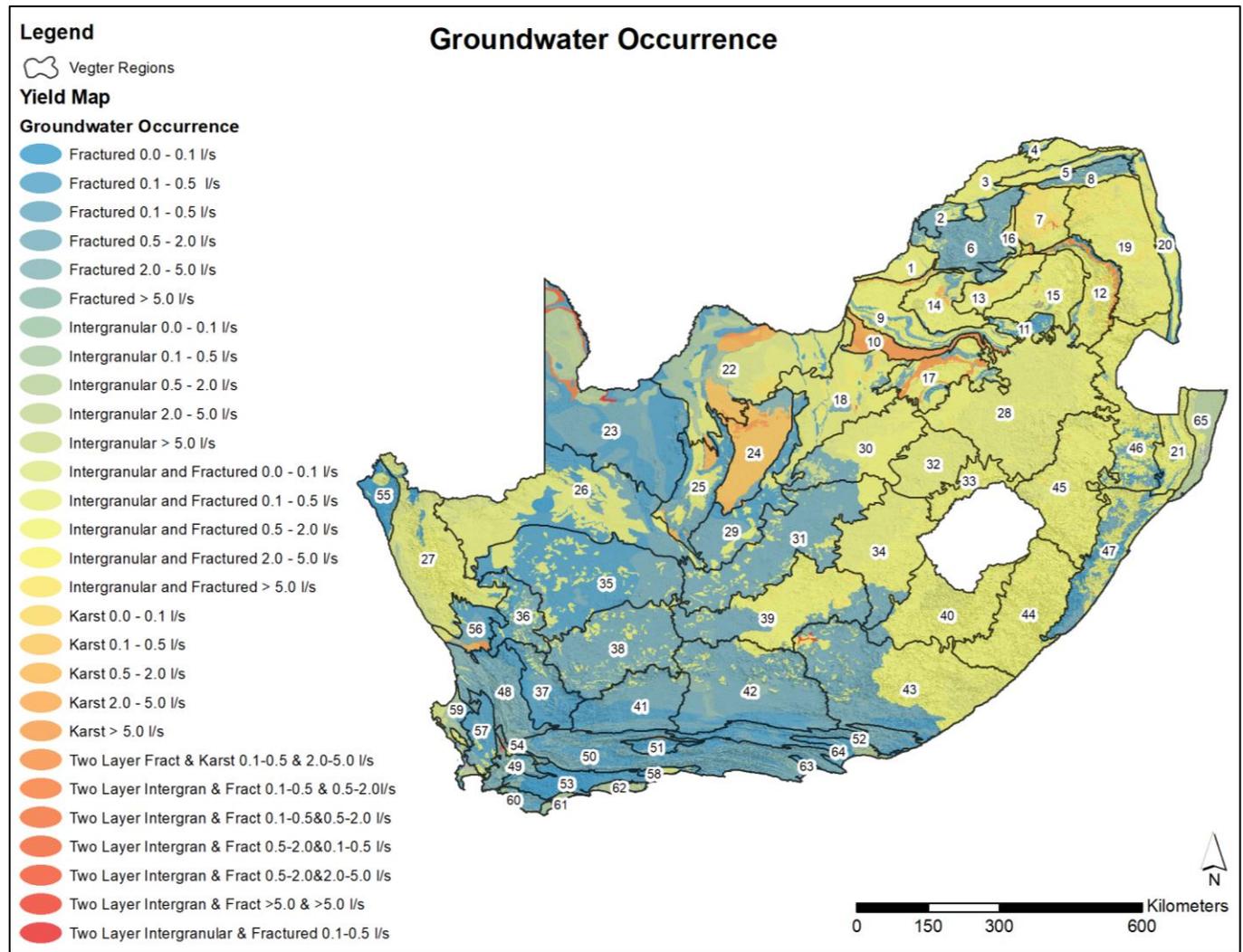


FIGURE 29 – GROUNDWATER OCCURRENCE MAP

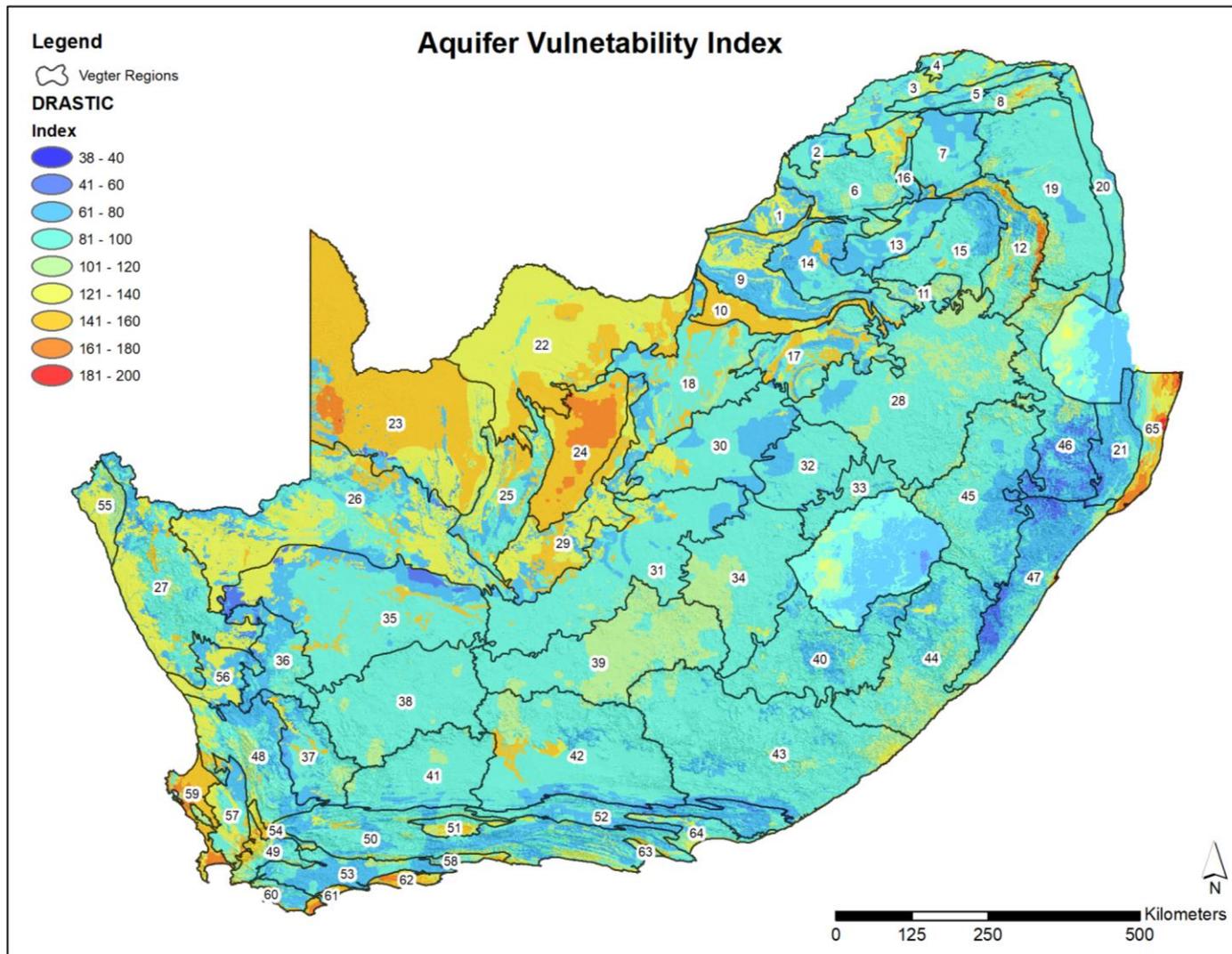


FIGURE 30 – GROUNDWATER VULNERABILITY MAP

## 4 STATISTICAL ANALYSIS

Note: This section of the report makes use of the results of the Vegter analysis of Region 65 (Project K5/2251/1) for illustration purposes only.

### 4.1 INTRODUCTION

Statistical methods associated with Vegter's analysis method are summarized in Table 8. The spatial dependency arises from the fact that the physical study area boundary selection dictates the borehole selection.

TABLE 8 – SUMMARY OF VEGTER'S STATISTICS

<i>Description of Statistic</i>	<i>Spatial Dependency</i>
1 The drilling success rate	X
2 Borehole yields in the various geologies	X
3 Distribution of boreholes per depth	X
4 Distribution of weathering and fracturing per depth	X
5 Strike frequency and cumulative strike frequency	X
6 Yield versus strike analysis	X
7 Yield versus dyke intersection	X

Considering national borehole databases and typical borehole information available, certain statistics are not readily available, e.g. the drilling success rate as generally only high yielding boreholes will be logged and captured in the database.

In addition to the statistics as described by the Vegter methodology, additional statistics are available for the purpose of delineation of areas that exhibit a similar geohydrological response. The concept of borehole spatial context is introduced here. The data attributes associated with each borehole as described in Section 2.3 can be used as the context for filtering the boreholes. As an example, if geology is chosen the results will be grouped by geology and the calculated statistics will be available per geological unit. The Vegter methodology reported most statistics per geology and this is now extended to the specific context selection as summarised in Table 9.

The remainder Section 4 will illustrate the statistics in the context of geology as specified by Vegter, but all statistics is also available by selected context.

TABLE 9 – SUPPLEMENTARY STATISTICS TO VEGTER’S LIST

	<i>Description of Statistic</i>	<i>Temporal Dependency</i>	<i>Spatial Dependency</i>
1	Borehole distribution and density		X
2	Water levels in context	X	X
3	Water strike in context		X
4	Yield in context		X
5	EC in context	X	X
6	Borehole chemistry in context	X	X

It is important to note that some of the borehole statistics can have temporal and spatial dependencies. It is a well-known fact that the NGA has poor temporal data with respect to water level and chemistry, due to the fact that the majority of boreholes are not being monitored. Boreholes generally only have a single entry in the NGA for water level and roughly 10% of the boreholes have a chemistry analysis associated with it. The temporal dependency arises from the fact that different results will be obtained for different time frames and in this case, have a large impact on the resultant spatial distribution of boreholes.

#### 4.2 SIMPLIFIED GEOLOGY

The 1:1 000 000 simplified geological map available from the Council for Geoscience is used in the statistical analyses, where the parameter of interest is related to geology. This is done to reduce the complexity of the data available in the 1:50 000 geological maps in relation to the available borehole data. It should be noted that the simplified geology only represents the surface geology and that detailed lithological information is contained in the borehole logs where data is available. This dataset also contains the aquifer rating (1-10) as it relates to the geology.

An example of the simplified geological map shown in Figure 31. The lithological layers and associated aquifer rating of Figure 31 is documented in Table 10.

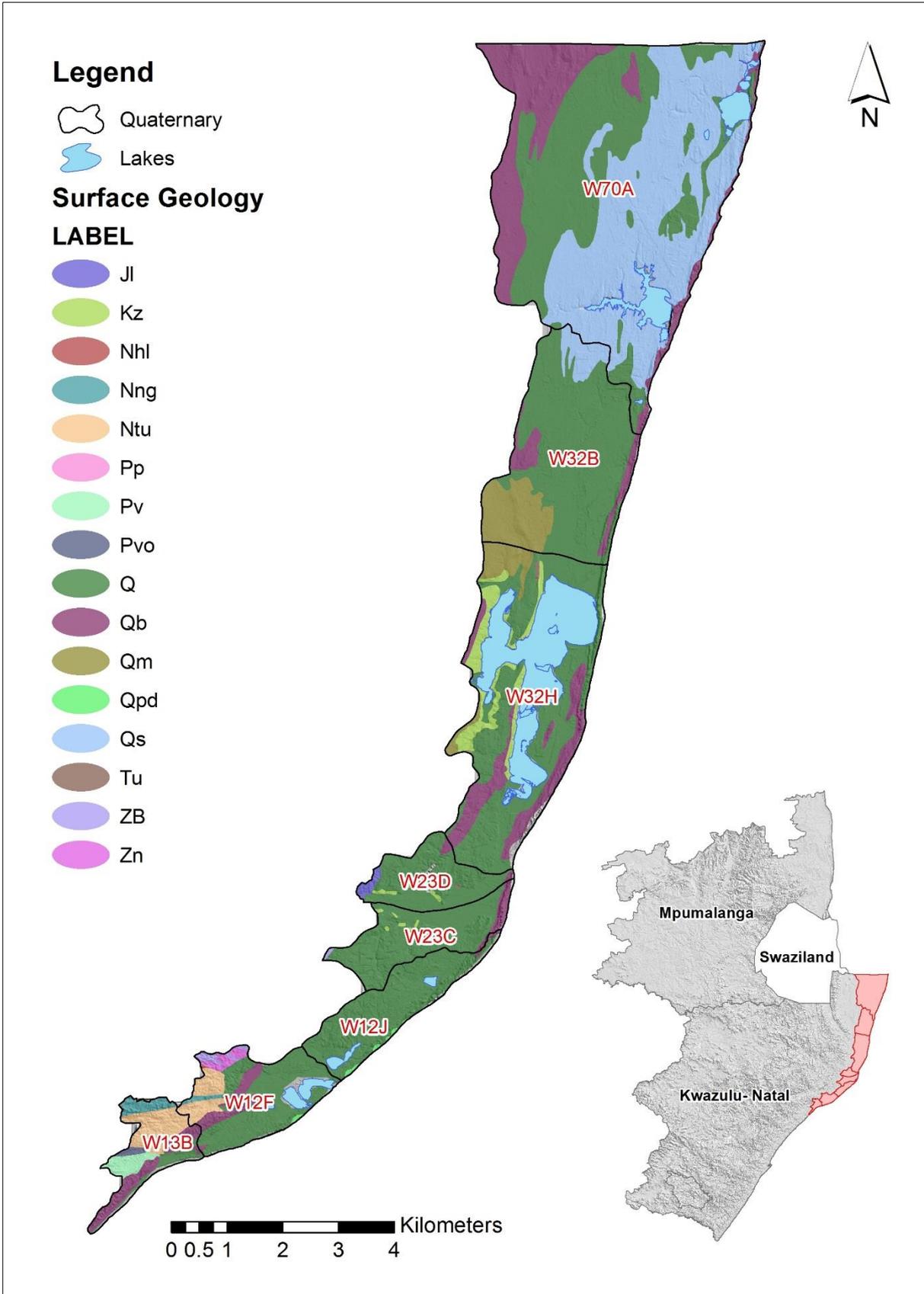


FIGURE 31 – EXAMPLE OF SIMPLIFIED GEOLOGY

TABLE 10 – EXAMPLE OF LITHOLOGICAL DESCRIPTION OF THE SIMPLIFIED GEOLOGICAL MAP

<i>Label</i>	<i>Lithology 1</i>	<i>Lithology 2</i>	<i>Lithology 3</i>	<i>Aquifer Rating (1-10)</i>
Jl	BASALT			4
Kz	SILTSTONE	ARENITE	CONGLOMERATE	2
Nhl	OLIVINE GABBRO	GABBRO		3
Nng	GNEISS			4
Ntu	AMPHIBOLITE	GNEISS	SCHIST	3
Pv	ARENITE	SHALE	COAL	2
Pvo	SHALE			1
Q	SEDIMENTARY	SAND	CALCRETE	7
Qb	ARENITE			4
Qm	SAND			10
Qpd	ARENITE	MUDSTONE	LIGNITE	2
Qs	ARENITE			4
Tu	SILTSTONE	LIMESTONE	LIMESTONE	5
ZB	GRANITE			4
Zn	GREENSTONE	AMPHIBOLITE	GRANULITE	3

It is worth mentioning that for structural and geological lineaments, the 1:50 000 data are used instead as a high-resolution data set is required to apply proximity sensing.

### 4.3 CALCULATED STATISTICS

#### 4.3.1 Borehole Distribution and Density

Borehole distribution and associated density is related to the specific geometry selected. Typically, the borehole distribution will be reported in the context selected. Figure 32 shows the borehole distribution for the example study area and Table 11 and Table 12 present the calculated borehole statistics for both the associated quaternary catchments and the associated simplified geology.

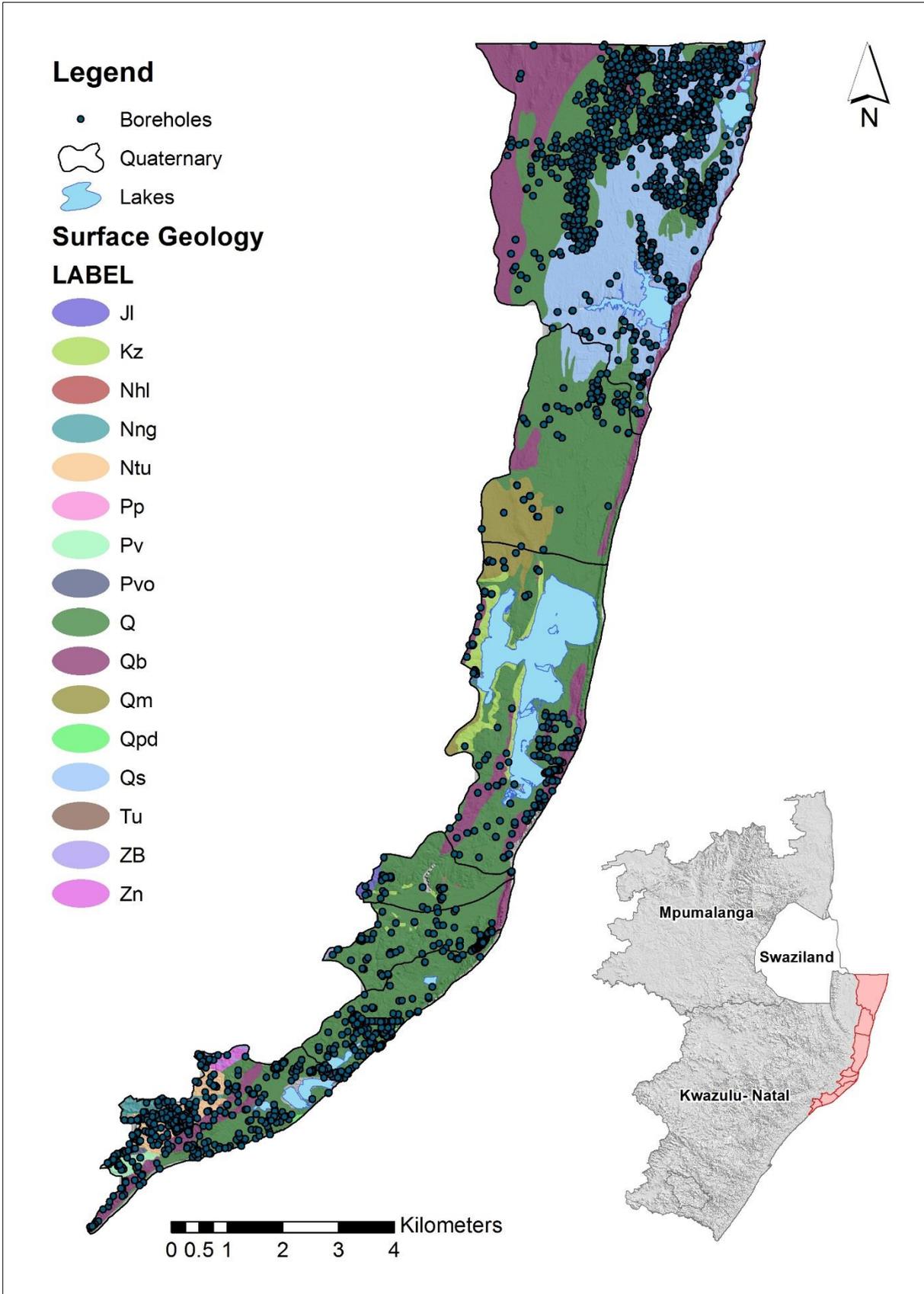


FIGURE 32 – EXAMPLE OF BOREHOLE DISTRIBUTIONS

TABLE 11 – BOREHOLE SUMMARY PER QUATERNARY

<i>Quaternary</i>	<i>Area (km<sup>2</sup>)</i>	<i>Boreholes</i>	<i>Density (bh/km<sup>2</sup>)</i>
W12F	399.0	268	0.67
W12J	332.1	257	0.77
W13B	222.4	257	1.16
W23C	312.6	206	0.66
W23D	247.9	61	0.25
W32B	192.8	88	0.46
W32H	1275.1	284	0.22
W70A	2589.0	1827	0.71
<i>Total</i>	5570.9	3248	

TABLE 12 – BOREHOLE SUMMARY PER GEOLOGY

<i>Geology</i>	<i>Area (km<sup>2</sup>)</i>	<i>Boreholes</i>	<i>Density (bh/km<sup>2</sup>)</i>
Jl	12.2	8	0.65
Kz	134.3	21	0.16
Nhl	0.7	2	2.90
Nng	30.3	21	0.69
Ntu	116.8	141	1.21
Pv	27.8	35	1.26
Pvo	7.0	14	2.00
Q	3016.8	1513	0.50
Qb	796.1	315	0.40
Qm	214.5	27	0.13
Qpd	7.1	31	4.35
Qs	1174.7	1095	0.93
Tu	3.9	2	0.51
ZB	11.6	8	0.69
Zn	17.0	15	0.88
<i>Total</i>	5570.9	3248	

The borehole density is simply calculated using Equation 1 presented below. The borehole count is recorded per context selection, but since the user can make use of custom polygons representing the study area, he/she will have to cut out contextual areas intersecting the study area provided to obtain the resultant areas to complete the calculation.

$$Density(borehole/km^2) = \frac{Borehole\ Count}{Area\ (km^2)} \quad (1)$$

### 4.3.2 Drilling Success Rate

The drilling success rate is dependent on the number of dry boreholes and the total number of boreholes drilled. Since this information is not readily available over the extent of South Africa, this statistic cannot be calculated making use of the national databases and hence is excluded from the calculated statistics.

### 4.3.3 Borehole Yields in Various Geologies

The borehole yield refers to the blow yield at the time of drilling and the discharge is the production abstraction for each borehole. Both the borehole yield and discharge per surface geology unit is presented as averages as well as the standard deviations. The results of the example study area are presented in Table 13 and a graphical representation is shown in Figure 33.

TABLE 13 – EXAMPLE SUMMARY OF BOREHOLE YIELDS AND DISCHARGE

<i>Geology</i>	<i>Borehole Yield (L/s)</i>			<i>Borehole Discharge (L/s)</i>		
	<i>Average</i>	<i>Std. Dev</i>	<i>Count</i>	<i>Average</i>	<i>Std. Dev</i>	<i>Count</i>
Kz	0.02	0.04	6	0.10	-	1
Ntu	0.60	0.90	12	0.52	0.56	17
Pv	2.21	1.41	24	1.80	1.70	4
Pvo	1.85	1.45	5	0.42	0.22	6
Q	16.91	10.18	696	1.69	2.87	81
Qb	11.46	10.76	75	0.77	1.43	28
Qm	7.35	8.40	16	1.47	1.25	7
Qs	18.92	7.69	937	5.92	6.82	3
Zn	1.64	3.19	9	1.83	3.35	8
		<i>Total</i>	1780		<i>Total</i>	155

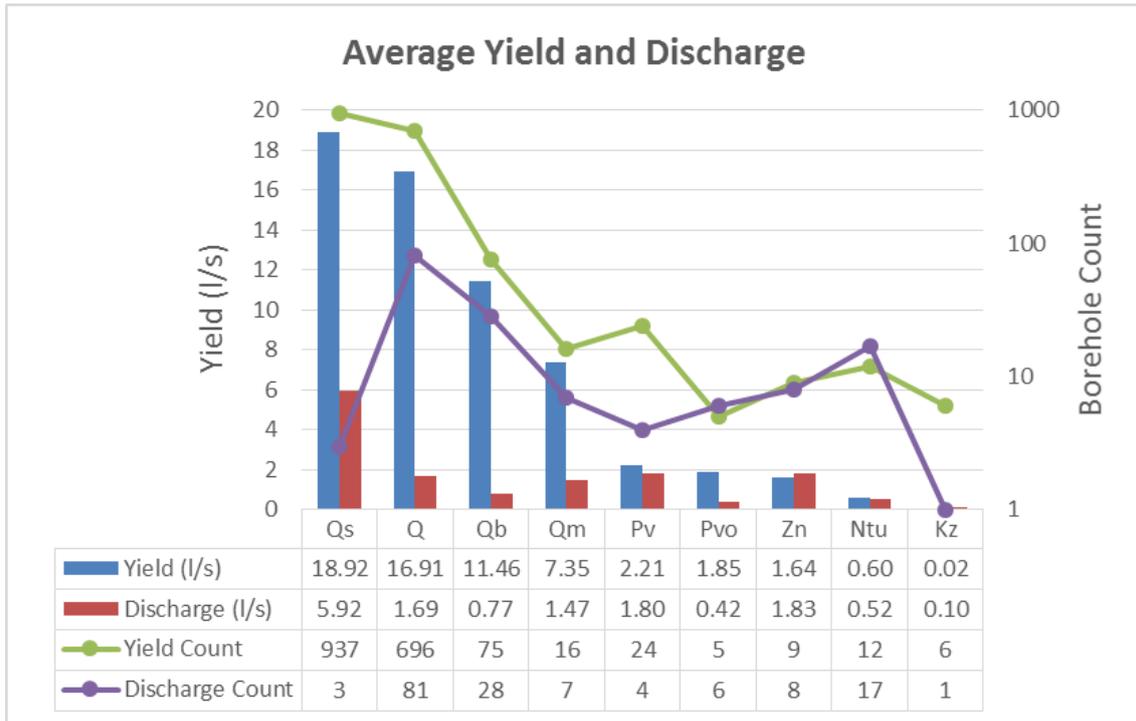


FIGURE 33 – GRAPHICAL REPRESENTATION OF BOREHOLE YIELD AND DISCHARGE PER GEOLOGY

#### 4.3.4 Distribution of Boreholes per Depth

The borehole distribution with depth as well as the surface elevation is also calculated per geology and both the average and standard deviation is reported. The results of the example study area are shown in Table 14 and Figure 34 respectively.

TABLE 14 – EXAMPLE SUMMARY OF BOREHOLE DEPTH AND ELEVATION

Geology	Borehole Depth			Elevation		
	Average	Std. Dev	Count	Average	Std. Dev	Count
Kz	47.7	37.5	13	41.0	16.0	21
Ntu	73.3	38.3	23	85.2	34.9	141
Pv	41.5	14.1	28	46.9	15.8	35
Pvo	39.0	23.8	10	29.4	8.3	14
Q	22.0	33.0	807	50.1	36.0	1513
Qb	37.3	37.1	103	51.9	26.4	315
Qm	19.4	11.3	21	33.9	21.9	27
Qs	7.8	6.9	940	51.9	21.1	1095
Zn	56.9	30.5	14	119.6	14.37	15
			<i>Total</i>			<i>3176</i>
			1959			

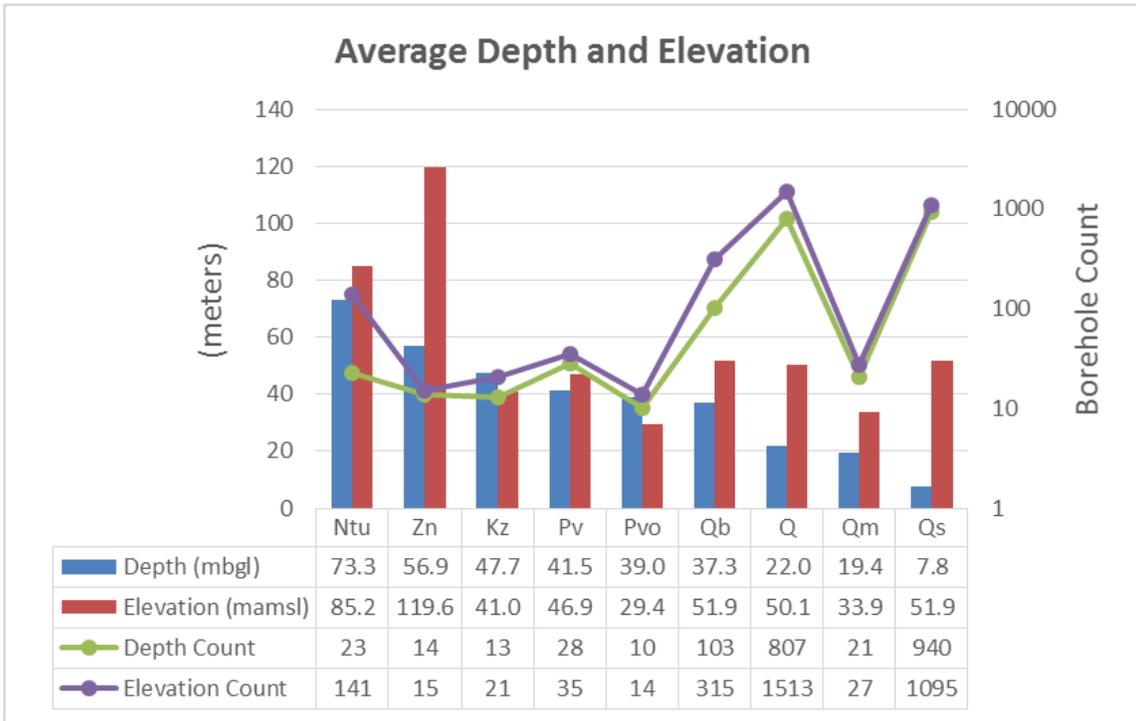


FIGURE 34 – GRAPHICAL REPRESENTATION OF BOREHOLE DEPTH AND ELEVATION PER GEOLOGY

In addition to the summary statistics provided above the depth frequency per simplified geological unit is also calculated. A depth frequency analysis of a geological unit in the example study area is shown in Figure 35.

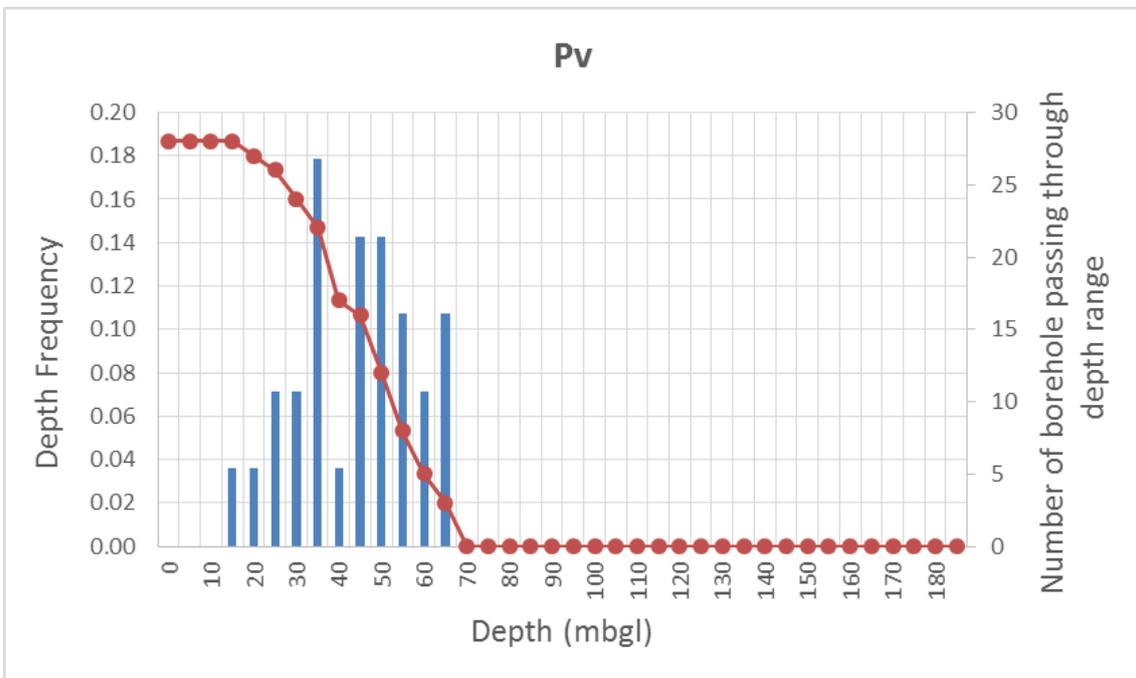


FIGURE 35 – EXAMPLE DEPTH FREQUENCY PLOT

The depth frequency is represented by a normalised histogram of borehole depth and the total number of boreholes passing through each depth range is displayed on the secondary axis to emphasise the number of boreholes used in the analysis.

#### 4.3.5 Strike Frequency and Cumulative Strike Frequency

The strike frequency expresses the frequency of water strikes with depth for a specific geological unit. This calculated frequency is presented in Equation 2.

$$\text{Strike Frequency} = \frac{\text{Number of strikes per depth range below surface}}{\text{Number of boreholes passing through depth range}} \quad (2)$$

An example of the strike frequency plot of the example study area is shown in Figure 36. The number of boreholes passing through the specific depth range associated with the strike is also shown on the secondary axis to emphasise the number of boreholes used in the analysis.

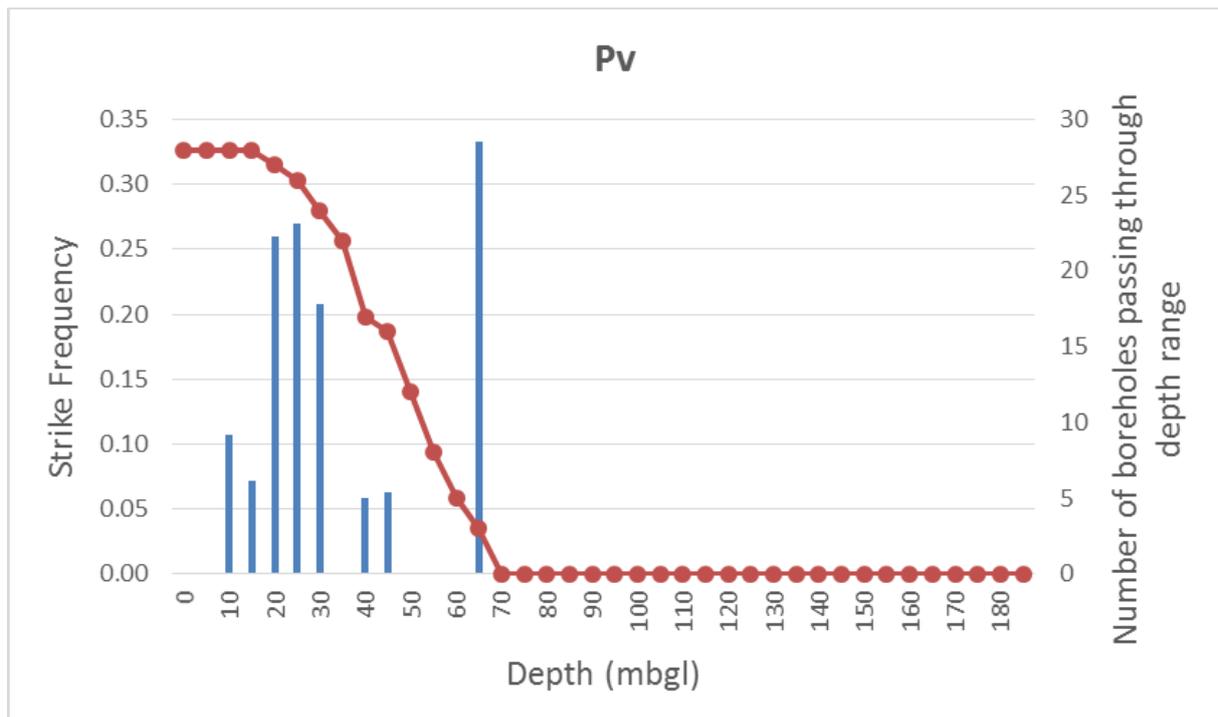


FIGURE 36 – EXAMPLE STRIKE FREQUENCY PLOT FOR A GEOLOGICAL UNIT

#### 4.3.6 Yield vs Strike Analysis

The yield versus strike analysis portrays the average yields for the different geologies, associated with strike depth. The average yield per strike depth based on the various geologies are also presented. The yield versus strike analysis for the example study area is shown in Figure 37.

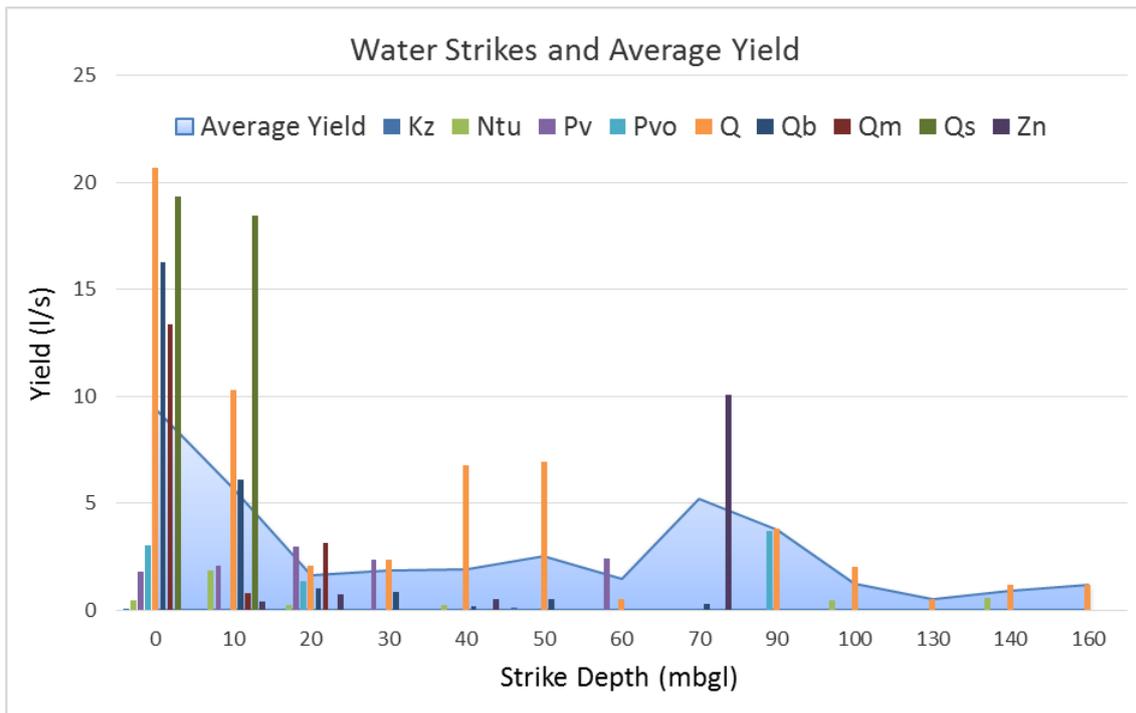


FIGURE 37 – EXAMPLE YIELD VERSUS STRIKE

#### 4.3.7 Yield vs Dyke Intersection

The general conception is that when a borehole intersects a dyke it is likely to have good yields. Often this is not the case since the reason dykes are targeted is only because they are easily detected by employing geophysical techniques. The yield versus dyke intersection analysis aims to show if boreholes intersecting dykes will have better yields compared to those with no dyke intersections for the specific area in question.

The dyke intersections are determined by analysing the existing borehole logs and analysing the associated yields. The yield versus dyke intersection analysis for the example study area is shown in Figure 38. The number of boreholes used in the analysis are displayed on the secondary axis, as it is an important indicator in the analysis.

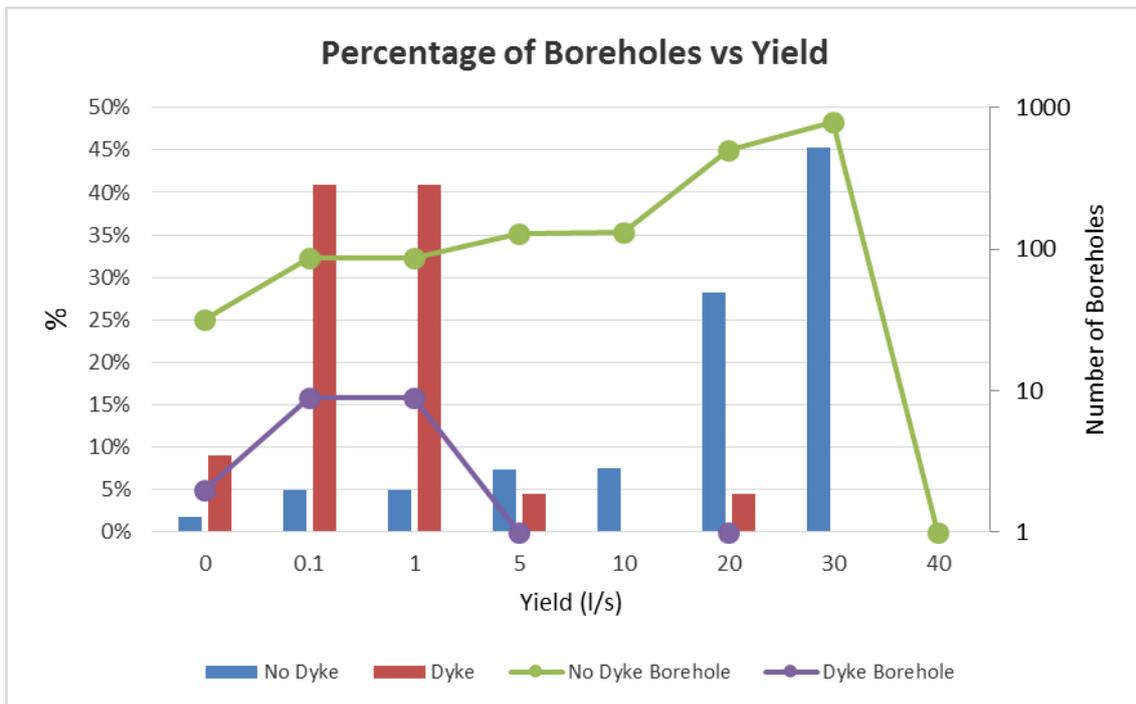


FIGURE 38 – EXAMPLE YIELD VERSUS STRIKE

Since boreholes intersecting dykes are mainly determined by means of borehole logs and often borehole log data is sparse in certain study areas, alternative spatial algorithms are also employed where boreholes are in close proximity to a dyke structure are also assumed to intersect the dyke. The aforementioned algorithms are an option to the user performing the analysis.

#### 4.3.8 Borehole Water Levels per Geology

Water level analysis per geology allows the investigation of how much the geology governs the regional water level response. Average water levels are used in these investigations. Pumping boreholes and those affected by pumping should typically be ignored.

When plotting water level (*mamsl*) versus surface elevation (*mamsl*), systems can be identified that exhibit a close correlation to surface topography, typically the shallow unconfined aquifers (Figure 39). The spatial distribution of these boreholes representing the different systems must also be taken into account in the delineation process.

A normalised water level histogram is used to depict the distribution of groundwater levels within a geological unit (Figure 40). Similar to previous frequency analyses, the number of boreholes passing through the associated depth is also displayed on the secondary axis.

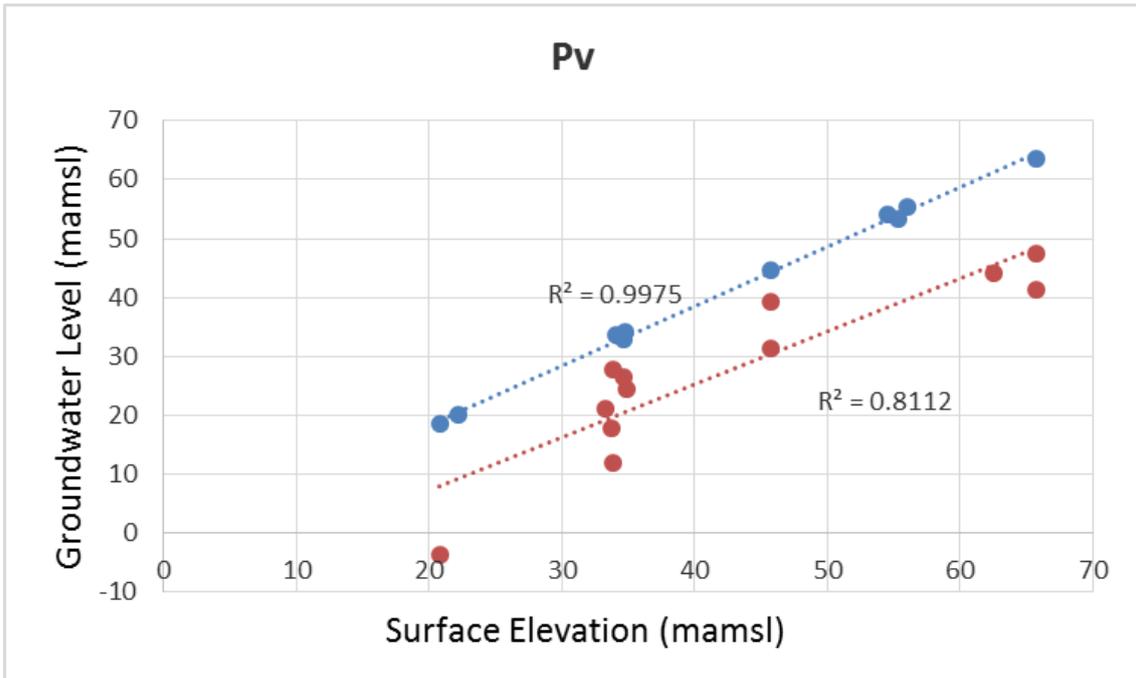


FIGURE 39 – EXAMPLE OF GROUNDWATER LEVEL VERSUS SURFACE ELEVATION

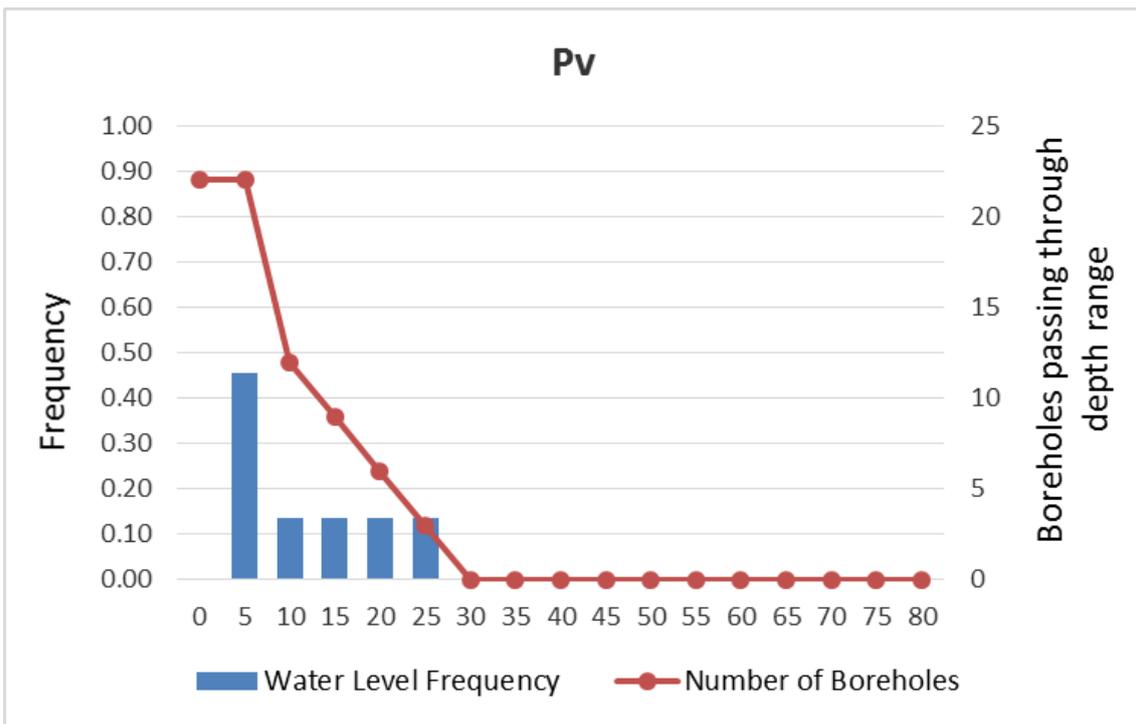


FIGURE 40 – EXAMPLE OF GROUNDWATER LEVEL DISTRIBUTION WITHIN GEOLOGICAL UNIT

### 4.3.9 Borehole Chemistry per Geology

The analysis of borehole chemistry per geological unit can further aid in the delineation of regions with similar geohydrological character. The analysis makes use of a Piper diagram as shown in Figure 41 from the example study area.

The plotting procedure for the aforementioned diagrams is presented in Appendix B. It is important to only consider the borehole chemistries not affected by pollution sources. This requires manual intervention in the borehole selection process as well as prior knowledge regarding the study area. A first step is to filter out boreholes with excessive EC values.

The challenge with this type of analysis is that the chemistry data in the national data bases are very sparse and the diagrams require data for all the major anions and cations. Generally, the EC and pH are more readily available.

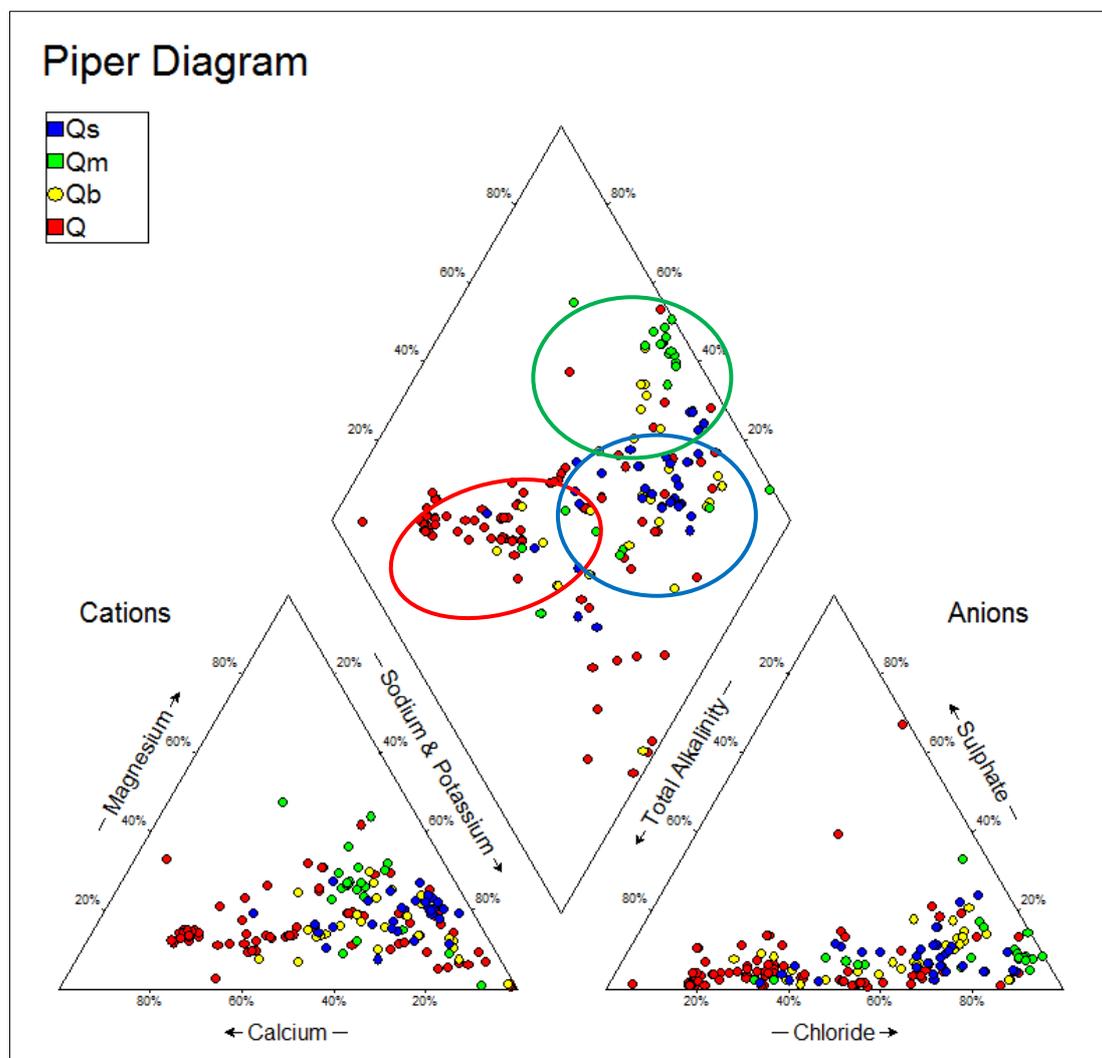


FIGURE 41 – EXAMPLE PIPER DIAGRAM OF EXAMPLE STUDY AREA

## 5 DELINEATION METHODOLOGY

### 5.1 PREAMBLE

This section discusses the delineation methodology developed during the course of the project. The purpose of the delineation is to allow users to sub-delineate the Vegter regions based on available borehole parameters. The resulting delineation should be areas that are similar in terms of the geohydrological character.

The major challenge is automating the delineation result so that a consistent delineation is obtained that is continuous across boundaries and study areas, with minimal user intervention.

### 5.2 DELINEATION CLASS METHOD

The delineation methodology applied in Vegter analysis of Region 65 (Project K5/2251/1) was based on the frequency analysis of available borehole parameters. This frequency analysis was conducted on the scale of the study area and the borehole parameters chosen (based on availability) is presented in Table 15.

TABLE 15 – BOREHOLE SUMMARY PER QUATERNARY

<i>Parameter</i>	<i>Comment</i>
Water Level	Average borehole water level measured from surface to water level position.
Blow Yield	During the drilling of a borehole, water strikes are encountered and the “blow yield” measured gives an indication of the sustainable yield of the borehole.
Water Strike Depth	During the process of drilling, water strikes are encountered at specific depths. For the purposes of delineation, the deepest water strike for each borehole was used.
Water Quality (EC)	The EC (Electrical Conductivity) of a borehole is an indication of the salinity of the water. Very saline water is associated with deep stagnant water, but can also be due to the mineralogy of the host rock.

A frequency analysis of the aforementioned parameters (Table 15) was done in such a way, so that the histogram bin selection resulted in the majority of the dataset (> 90%) residing in the first four bins. This is done to combine the aforementioned parameters in a linear combination to calculate a delineation class. A summary of this frequency analysis of the example study area is shown in Table 16.

TABLE 16 – BOREHOLE SUMMARY PER QUATERNARY

<i>Bin No</i>	<i>Water Level</i>		<i>Blow Yield</i>		<i>Water Strike</i>		<i>EC</i>	
	<i>Bin</i>	<i>Count</i>	<i>Bin</i>	<i>Count</i>	<i>Bin</i>	<i>Count</i>	<i>Bin</i>	<i>Count</i>
1	5	1498	8	307	5	194	60	165
2	10	250	16	240	10	1216	120	49
3	15	54	24	474	15	85	180	17
4	20	28	32	656	20	26	240	14
		1830		1677		1521		245

A delineation class was formulated based on the water level, blow yield, water strike and EC parameter frequency analysis. Since the bin selection was done to ensure the bulk of the dataset are contained within the first four bins, only the bin number and associated range was of importance for the delineation process. The delineation class calculation applied is presented in Equation 3.

$$Delineation\ Class = \frac{Level + (5 - Yield) + Depth + EC}{4} \quad (3)$$

Each of the parameters in Equation 3 correspond to the physical bin number ranging from 1 to 4 which implies the minimum delineation class is 1 and the maximum delineation class is 4. Note that the yield bin was reversed to group high yields with shallow water levels and shallow water strikes, as this was the trend observed in the dataset for the example study area.

The visualisation of the delineation class is required to transform point data to representative polygon regions. The principle of Voronoi polygons involves generating a Voronoi polygon around each borehole and assigning the calculated delineation class to the generated polygon. Once this is done adjacent polygons with the same delineation class values are merged and this represents the simplification part of this operation. This process is shown for the example study area in Figure 42 and it is evident that a high borehole density is present in the northern part of the study area, which results in more detailed areas being generated.

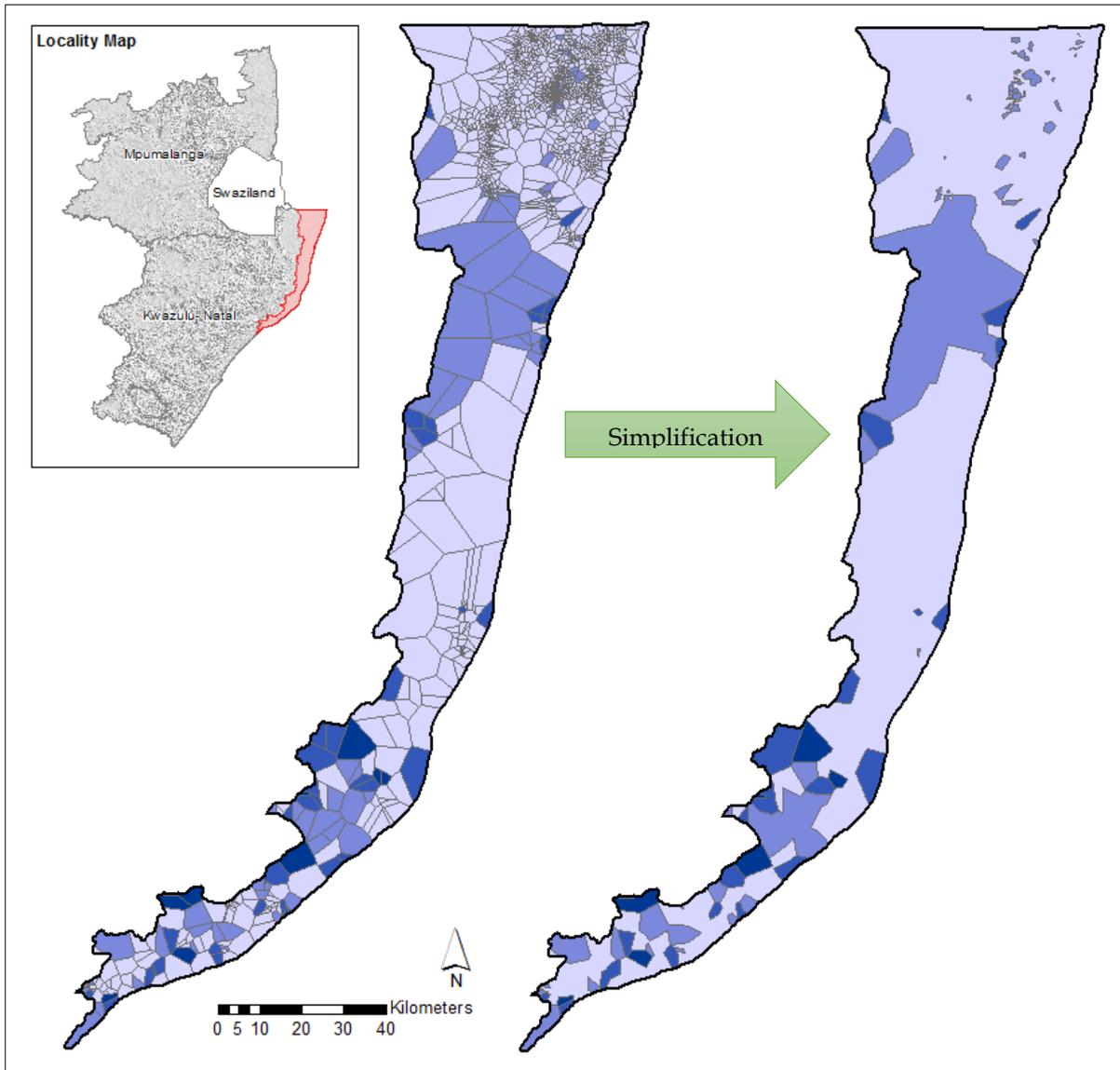
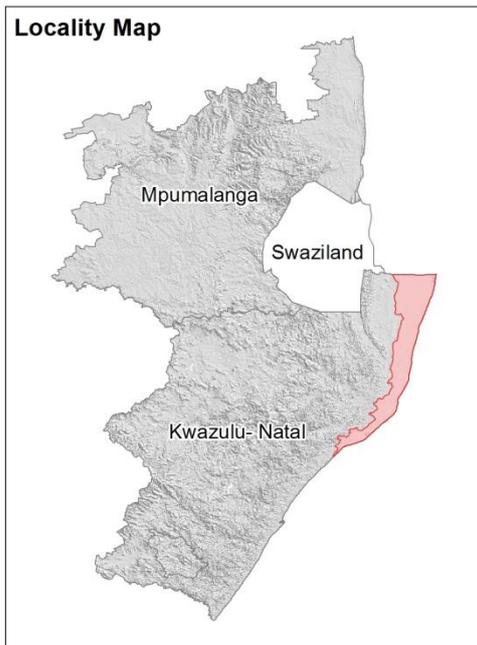


FIGURE 42 – VORONOI POLYGON GENERATION WITH SIMPLIFICATION

The final delineation map of the example study area is shown in Figure 43 together with the water level distribution and the importance of good data distribution is evident in this map. The final result is a delineation represented by four classes where each class is representative of the following four parameters and typical ranges as identified in the frequency analysis:

- Water level
- Blow yield
- Water strike
- EC

# VEGTER REGION: 65 Zululand Coastal Plain Groundwater Units with groundwater levels



## Legend



Most probable borehole water level

**Water Level (mbgl)**

- 5
- 10
- 15
- 20

**Groundwater Units**

**Most probable borehole profile**

- Shallow water level (<5 mbgl); high blow yield (~32 L/s); shallow water strike (5-10 mbgl); low EC (<60 mS/m)
- Shallow water level (~5 mbgl); low blow yields (~24 L/s); water strikes (10-15 mbgl); moderate EC (~180 mS/m)
- Intermediate water level (10-15 mbgl), low blow yield (~8 L/s), avg water strike (15-20 mbgl); moderate EC (~180 mS/m)
- Deep water level (>20 mbgl); low blow yield (<8 L/s); deep water strike (>20 mbgl); high EC (>240 mS/m)

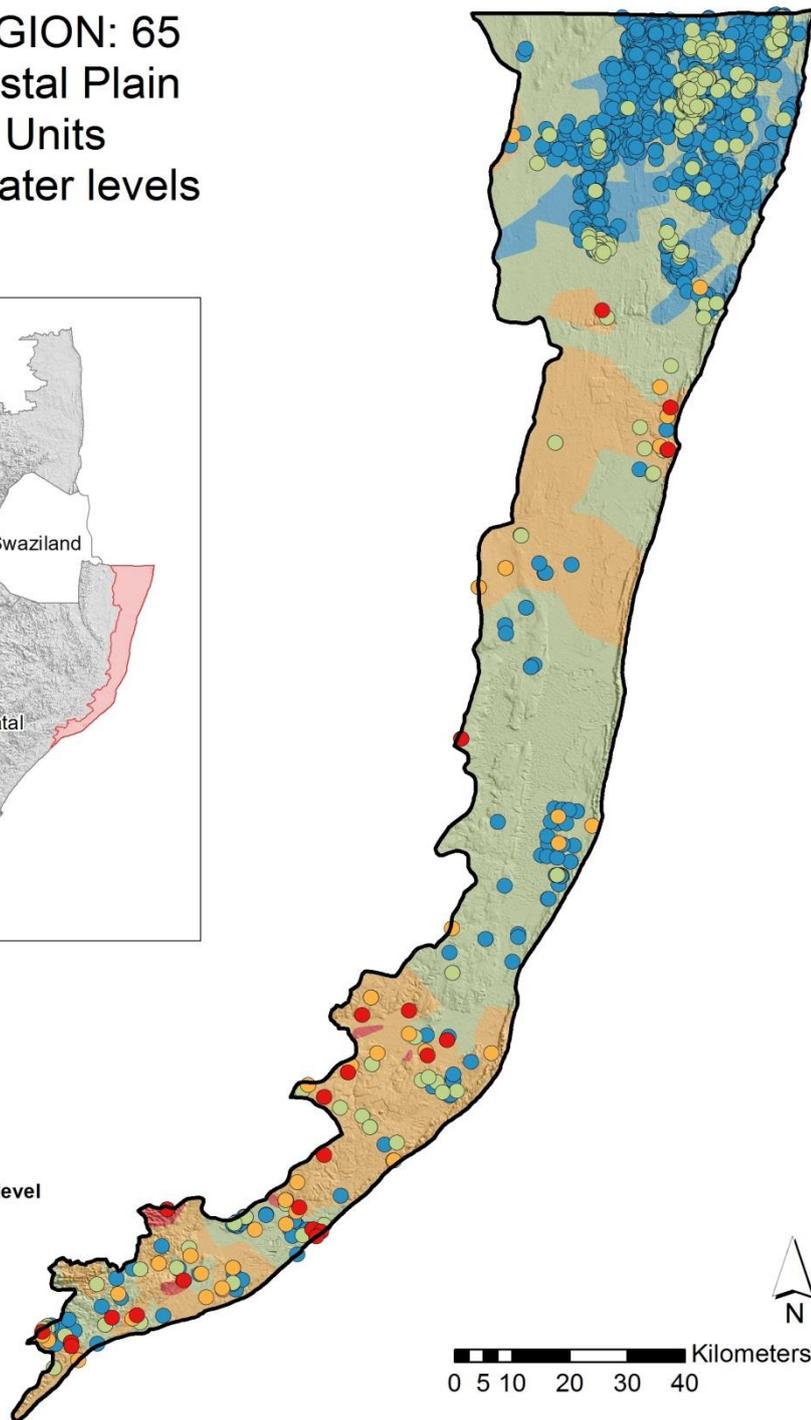


FIGURE 43 – VORONOI POLYGON GENERATION WITH SIMPLIFICATION

### 5.3 INFERRED TRANSMISSIVITY CLASSIFICATION METHOD

#### 5.3.1 Introduction

Although the delineation class methodology presented in the previous section, seems to work for the example study area presented, it has the following flaws when applied over large areas:

- It is not always the case that the deeper water levels are associated with the higher EC values or that the shallow water levels are associated with higher blow yields. Therefore, it becomes increasingly difficult to formulate four unique classes in terms of the parameter combinations.
- The delineation class method considers the whole range of values for each parameter and transforms them to a delineation class based on the results of the frequency analysis. Therefore, when applied to large areas with a wide range of parameter values, delineation detail is lost in certain areas and will only be represented by a single class.
- Continuity is lost in delineation when considering adjacent areas which exhibit a contrast in parameter values, since the delineation classes for each area is formulated in terms of the frequency analysis of the total range of values.

To address these aforementioned issues, the delineation class should be replaced with a calculated value that relates to a physical property. Since in the geohydrological response or character of the study area is of interest it makes sense to look at aquifer parameters.

#### 5.3.2 Aquifer Parameters

When considering the available parameters, water level and yield in the Cooper-Jacob equation comes to mind as this equation determines the drawdown in a borehole based on abstraction rate, duration of pumping and the associated aquifer parameters as presented in Equation 4.

$$s = \frac{2.3Q}{4\pi T} \log \left( \frac{2.25Tt}{r^2S} \right) \quad (4)$$

where

- $s$  = Drawdown from static water level (m)
- $Q$  = Abstraction rate ( $m^3/d$ )
- $T$  = Transmissivity ( $m^2/d$ )
- $S$  = Storativity or Storage Coefficient
- $t$  = Time of pumping (days)
- $r$  = Borehole radius

The Cooper-Jacob equation is less sensitive to uncertainty in the storativity due to the fact that the storativity only appears in the log-term of the equation and some estimations of storativity already exist in map form.

Dennis and Dennis (2012) assigned ranges for storativity according to the aquifer type as shown in Figure 44.

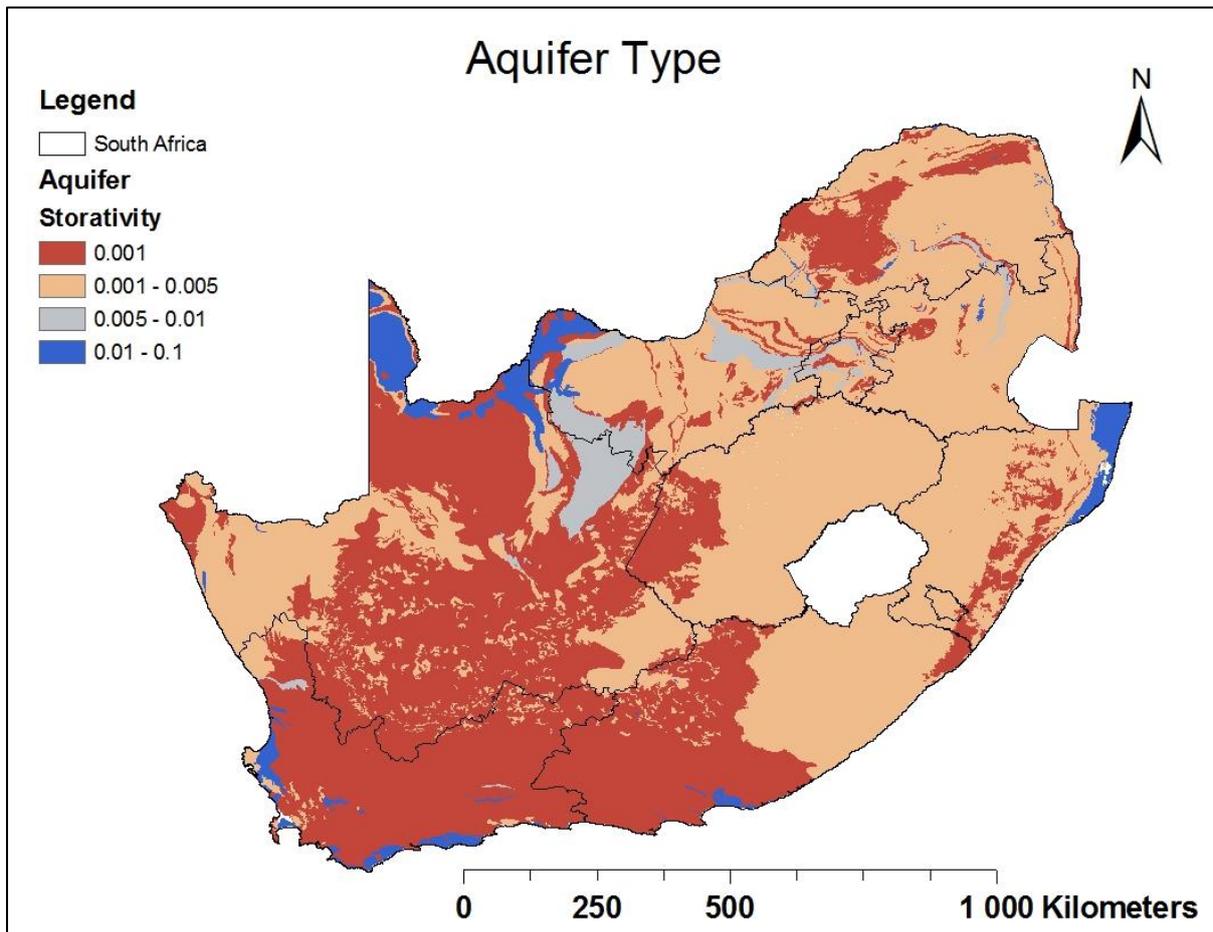


FIGURE 44 – AQUIFER STORATIVITY VALUES ON REGIONAL SCALE

The Vegter map of storage coefficient or storativity is shown again in Figure 45 and it is evident that some correlation exists between this map and the one presented in Figure 44.

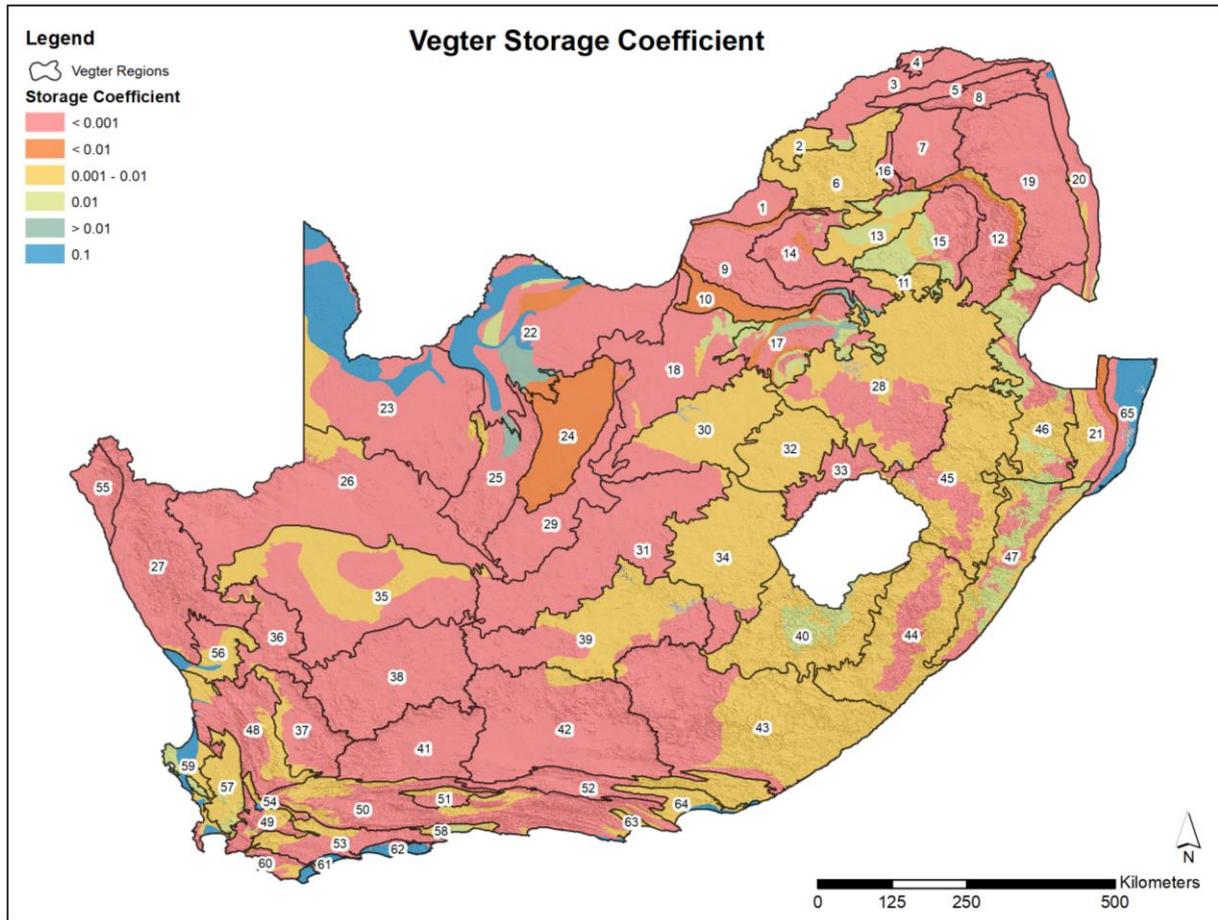


FIGURE 45 – VEGTER’S STORAGE COEFFICIENT MAP

It is clear from Equation 4 that if the storativity is estimated and the borehole radius is taken as a constant, then the transmissivity can be calculated if the abstraction rate, drawdown and time of pumping is known.

A method of calculating the sustainable yield is described in unpublished groundwater course notes by the Canadian International Development Agency (CIDA), and has been used by the Geological Survey in Swaziland. Swaziland’s Geological Survey carries out twenty-four-hour constant discharge tests on boreholes to be equipped with motorised pumps and eight-hour tests on boreholes to be equipped with handpumps or windmills. An approximate daily production yield is calculated using Equation 5.

$$Q = 0.068Ts \tag{5}$$

where

- $s$  = Available drawdown from static water level (m)
- $Q$  = Abstraction rate ( $m^3/d$ )
- $T$  = Transmissivity ( $m^2/d$ )

Equation 5 can be rearranged to make transmissivity the subject of the equation and the resultant expression is presented in Equation 6.

$$T = \frac{14.7Q}{s} \quad (6)$$

When analysing a pump test making use of the Cooper-Jacob equation (Equation 4), the transmissivity can be calculated making use of the drawdown that occurs over 1 log-cycle of time, therefore resulting in Equation 7 where  $\Delta s$  represents the drawdown over 1 log-cycle.

$$T = \frac{2.3Q}{4\pi\Delta s} = \frac{0.183Q}{\Delta s} \quad (7)$$

Equation 6 and Equation 7 are similar in form, but differ with a factor of 80. This is attributed to the fact that in Equation 7 typically a constant pump rate over 24 hours is used and in Equation 6 a sustainable pump rate is used.

If Equation 4 is directly applied to calculate the transmissivity, by using an estimation for storativity, using a borehole radius of 0.08 m and assuming the abstraction rate is over a period of a year (aquifer will have been subjected to seasonal effects), the following relationship (Figure 46) is obtained when compared to Equation 7 and assuming  $\Delta s = s$  :

$$F = -0.434 \ln(S) + 5.1 \quad (8)$$

where

- $S$  = Available drawdown from static water level (m)
- $F$  = Conversion factor between transmissivity calculated by Equation 4 and Equation 7

$$T_{Eq4} = F(T_{Eq7}) \quad (9)$$

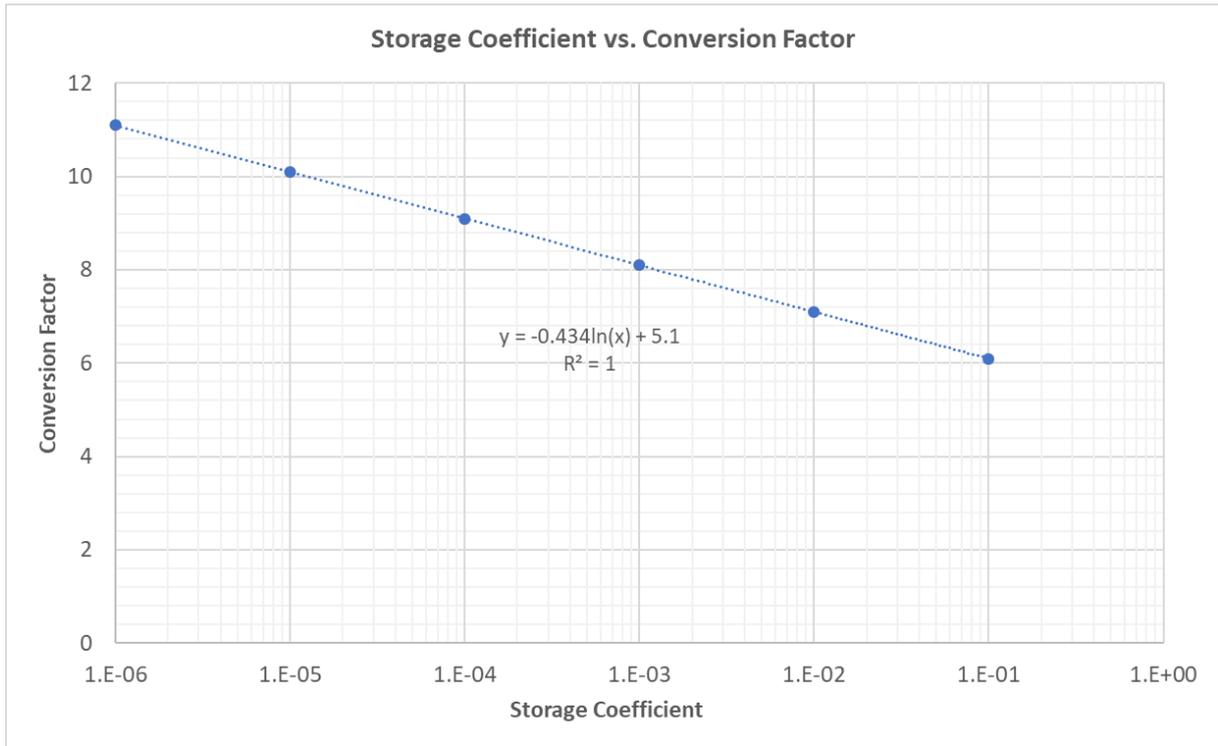


FIGURE 46 – STORAGE COEFFICIENT VS. CONVERSION FACTOR

The magnitude of the conversion factor is related to the abstraction duration (365 days), since Equation 7 is not explicitly a function of time (1 log-cycle of time assumed).

The goal is to calculate a parameter that relates to a physical parameter to be used for delineation purposes. Table 17 lists the assumptions made with regard to the available borehole parameters and the parameters of Equation 4:

TABLE 17 – PARAMETER ASSUMPTIONS

<i>Cooper-Jacob</i>	<i>Assumptions</i>
Q	The blow yield is representative of Q
s	The drawdown can be calculated as $s = \text{Strike} - \text{Level}$
S	Estimated from aquifer type (Figure 44 or Figure 45)

The only country wide database of transmissivity values is that of the GRAII dataset. According to DWAF (2006), as no regional aquifer transmissivity information are available, borehole yields may be used in a qualitative way as a proxy for transmissivity and the following relationship is applied, where  $T$  is in  $m^2/d$  and  $Q$  expressed in  $L/s$ :

$$T = 10Q \tag{10}$$

Since the Limpopo GRIP database has recorded transmissivity values of pump tests conducted, this database was used to select boreholes for which transmissivity values were recorded. In addition to the recorded transmissivity values, the parameters as described in Table 17 were obtained for the same dataset, which resulted in 502 boreholes being used in the analysis. Transmissivity values based on Equation 4 and Equation 7 were computed to compare to that of the transmissivity values obtained from pumping tests and the GRAII dataset. The probability distribution of the results is shown in Figure 47.

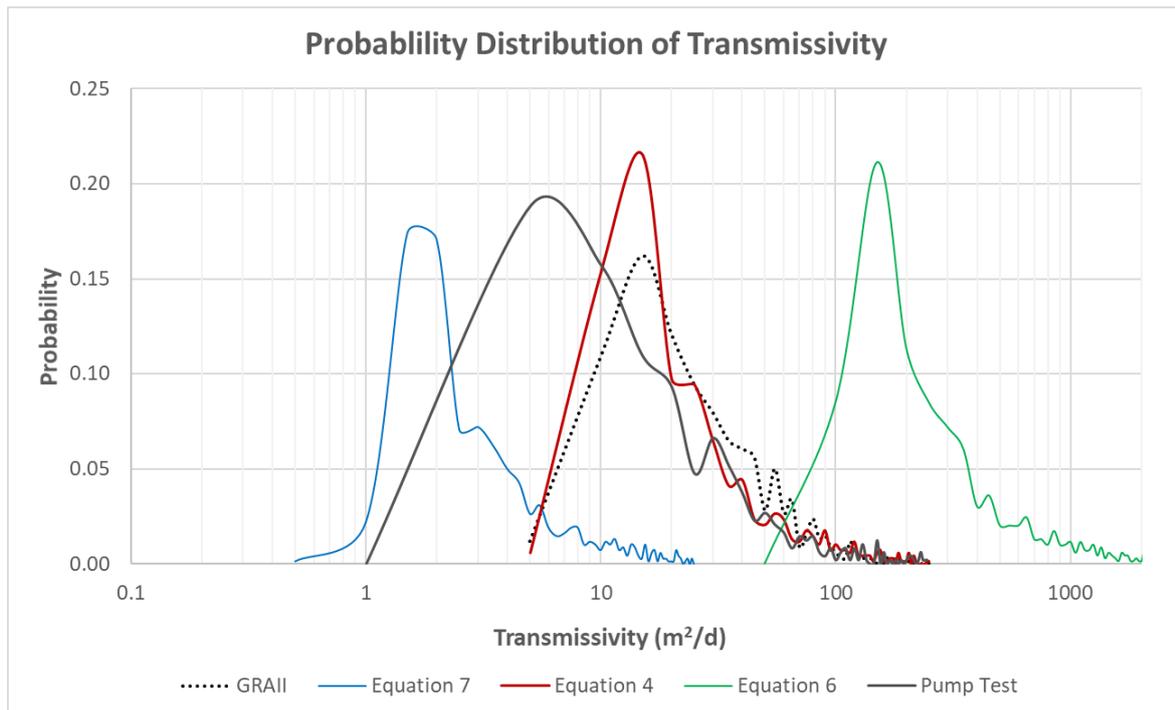


FIGURE 47 – PROBABILITY DISTRIBUTIONS OF CALCULATED TRANSMISSIVITY

It is clear that the distribution of transmissivity values obtained from Equation 4 and that of the GRAII dataset compare well. The distribution of the higher transmissivity values from the pump test data also compare well with that of Equation 4 and the GRAII dataset although the mean value from the pump test data is almost an order of magnitude lower. It is clear that Equation 6 and Equation 7 represent transmissivity distributions that are much higher and lower respectively when compared to the rest.

Due to the difference in mean values between the pump test data and Equation 4, it is not expected that a high correlation will exist between these two datasets. The correlation between the transmissivity obtained from the pump test data and that of Equation 4, Equation 6, Equation 7 and the GRAII dataset is shown in Figure 48. The poor correlation is evident between the datasets, but as expected, Equation 4 and the GRAII data shows the closest values to that of the pump test data analysis.

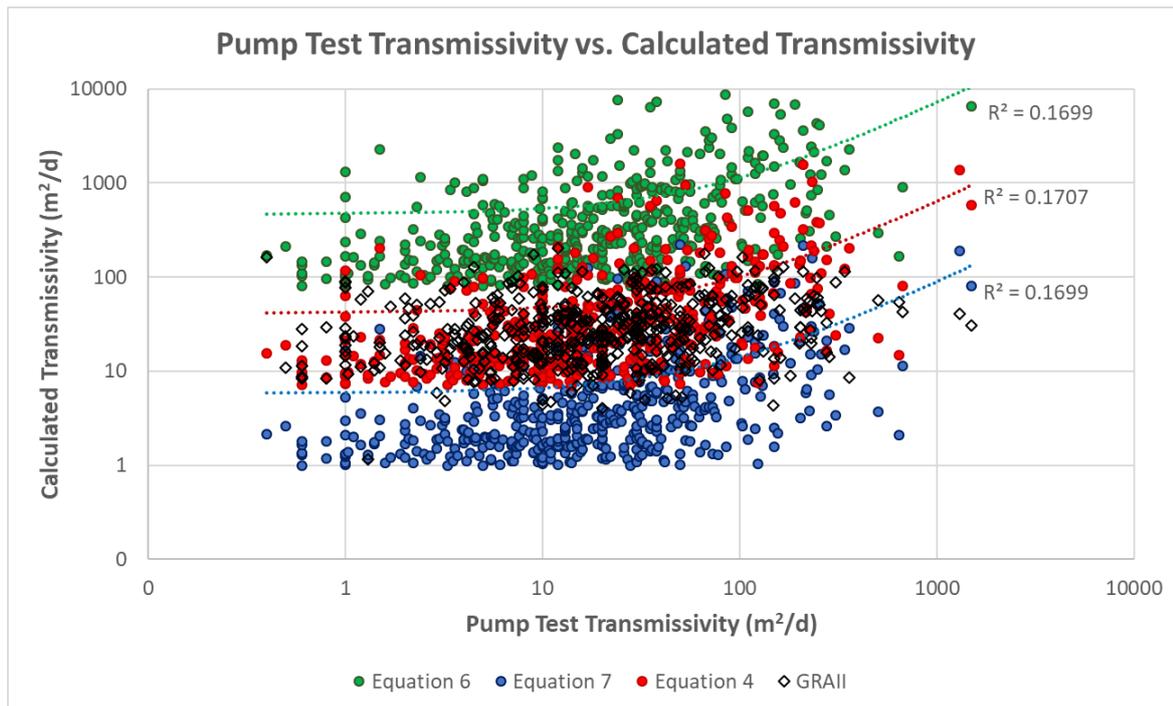


FIGURE 48 – PROBABILITY DISTRIBUTIONS OF CALCULATED TRANSMISSIVITY

Since transmissivity values are not readily available across the country, Equation 4 will be used to calculate an inferred transmissivity for delineation purposes of groundwater units that exhibit a similar response. The term inferred is used due to the fact that the transmissivity calculated using Equation 4 does not reflect the actual transmissivity that is obtained through analysing a pumping test.

### 5.3.3 Classification

In the previous section it was shown that it is possible to calculate an inferred transmissivity by means of an analytical equation (Equation 4) based on parameter values available in the database. The transmissivity was inferred using representative parameters and the said equation, based on some assumptions.

If it is difficult to derive parameters that would describe the problem of estimating transmissivity, as the level of knowledge regarding this specific field is not yet advanced enough to compute analytically, classification algorithms can be used to compute results. Since there is some knowledge to compute a result, classification is used to verify the existing analytical relationship proposed (Equation 4 with parameters from Table 17).

Classification algorithms are able to compute many probable results across many parameters, if there are sufficient examples with a known outcome – this is called inference. It is not

needed to know what the actual analytical relation between the parameters and the outcome is as classification algorithms will internally detect this relation, although they will not be able to express it. If an algorithm is able to express the detected relation, it is called a machine learning algorithm. Because not all relations can always be expressed with signs and words, the classification algorithms have higher classification accuracy than machine learning algorithms.

Considering various parameter combinations available from the database, the two governing parameters identified by various classification schemes was *yield* and *available drawdown* which is in line with two of the parameters presented in Table 17. Using a trial and error approach the following relationship was determined for transmissivity classes considered in the classification algorithm:

$$TC_n = 5^n \quad (11)$$

where

$$n = \text{Enumerator: } 0 \leq n \leq 3$$

$$TC = \text{Transmissivity Class}$$

The first classification tree makes use of both *yield* and *available drawdown (AD)* to predict the transmissivity is shown in Figure 49. The four transmissivity classes specified by Equation 11 and the associated classification probabilities is also shown in Figure 49.

On closer inspection it is clear that the *AD* is only used to classify two different instances of a transmissivity of 25 m<sup>2</sup>/d, which is one class. Therefore, *AD* can be omitted as classification criteria, resulting in the classification tree presented in Figure 50.

Note the yield values in the classification trees refers to the unit m<sup>3</sup>/d.

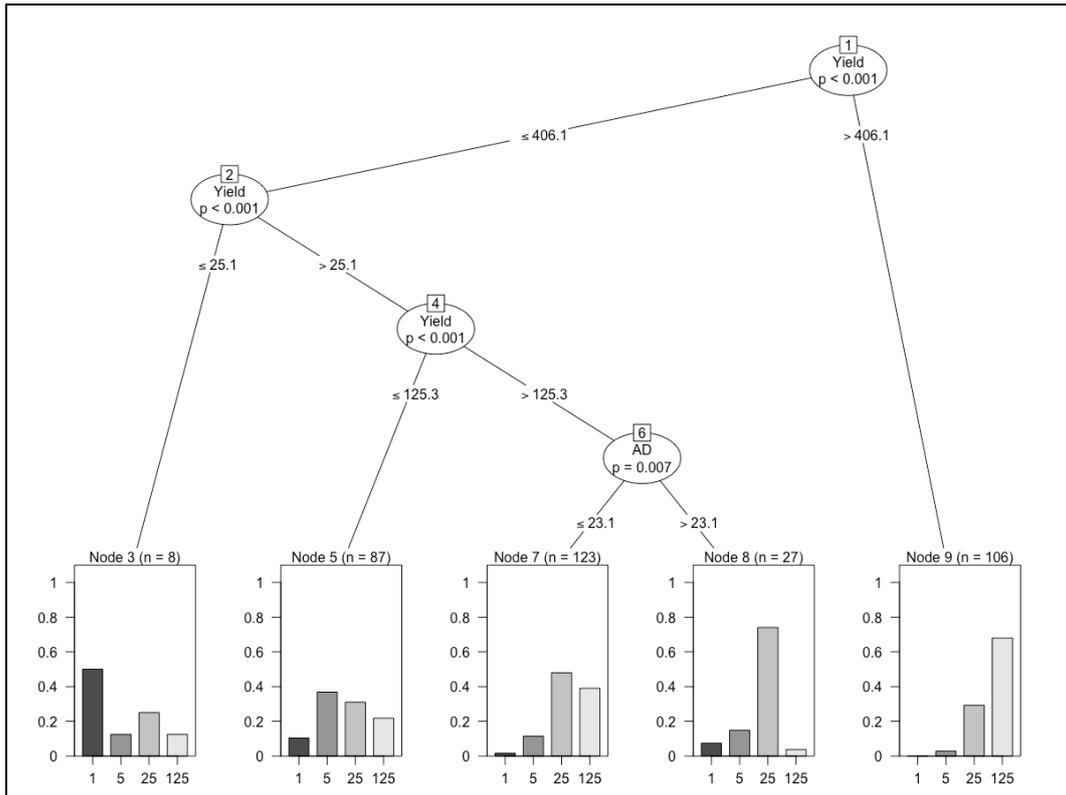


FIGURE 49 – CLASSIFICATION TREE CONSIDERING BOTH YIELD AND AVAILABLE DRAWDOWN

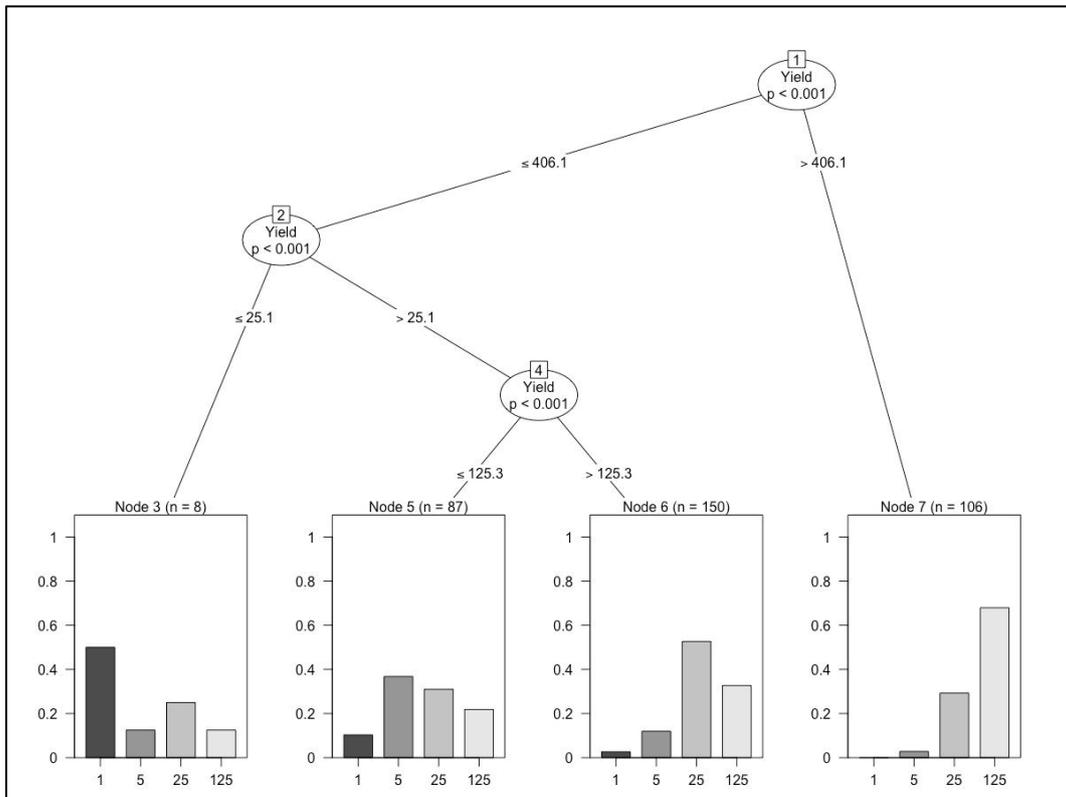


FIGURE 50 – CLASSIFICATION TREE ONLY CONSIDERING YIELD

The indicated probabilities refer to the probability of predicting the correct transmissivity class based on the transmissivities obtained from the pumping test data and is summarised in Table 18. Note that the last transmissivity class represents all transmissivities greater than 25 even though the maximum is indicated as 125.

TABLE 18 – CLASSIFICATION TREE PROBABILITIES

<i>Transmissivity Class (m<sup>2</sup>/d)</i>	<i>Prediction Probability (%)</i>
0-1	50%
1-5	39%
5-25	52%
> 25 (25-125)	70%

When classing all the datasets according to Equation 10 the following correlations were obtained with respect to the actual classed transmissivities:

TABLE 19 – CLASSED DATASETS CORRELATION WITH CLASSED TRANSMISSIVITIES

<i>Classed Dataset</i>	<i>Pearson Correlation</i>
Classification Tree	0.42
Equation 4	0.38
GRAII	0.24

Considering Table 19 it seems that the classification tree outperform Equation 4 and based on this result, the classification tree (Figure 50) is recommended for delineation purposes. As more data becomes available in terms of actual transmissivities across the country, the existing classification tree should be updated if required.

#### 5.4 INTERPOLATION TECHNIQUES

It is a well-known fact that different interpolation techniques can produce different results from the same data set as shown in Figure 51. Data density plays a major role when interpolating and few interpolation algorithms are suited to extrapolate to areas where data distribution is sparse. The aforementioned can lead to large errors in the estimations.

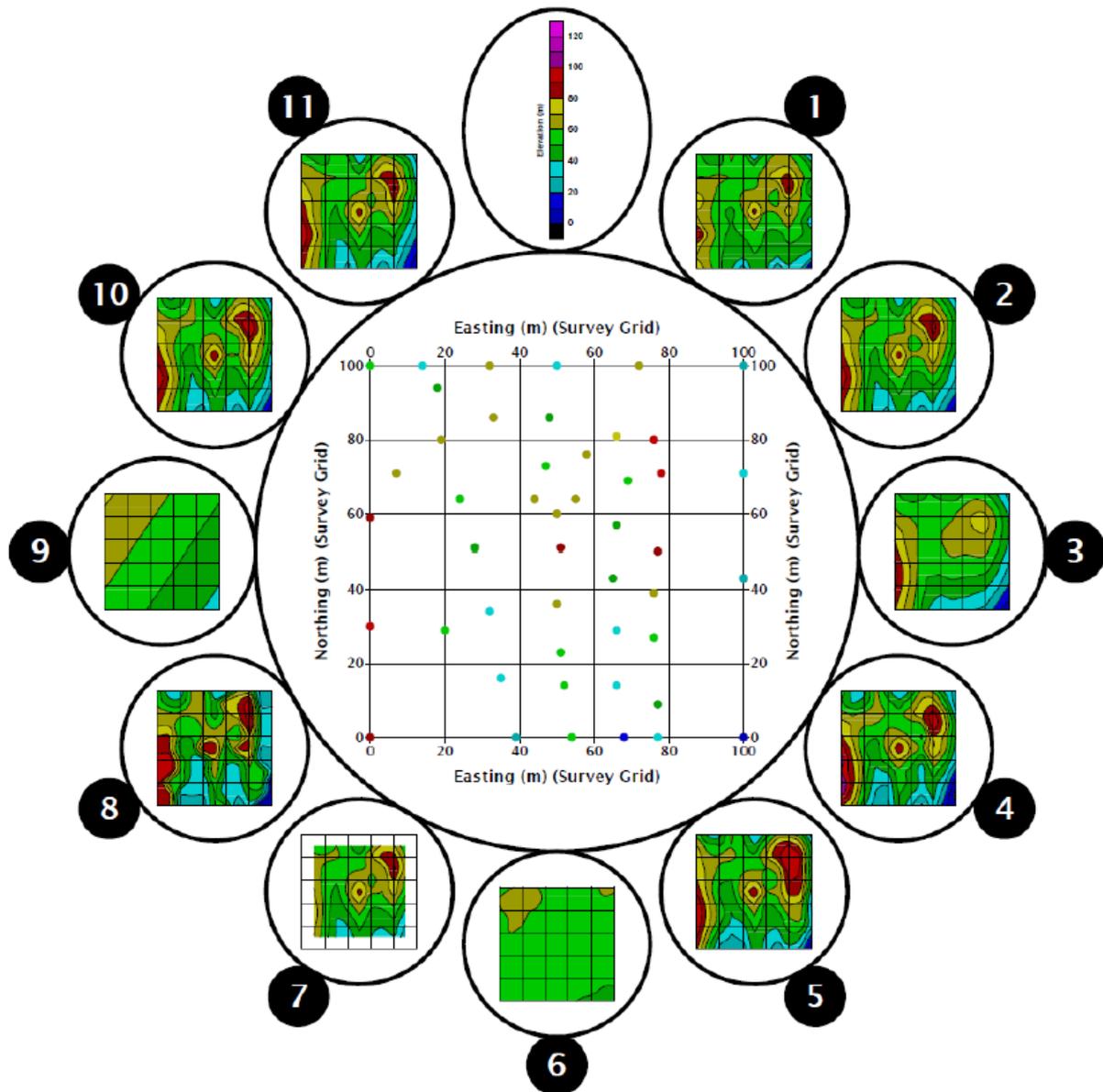


FIGURE 51 – INTERPOLATION RESULTS FROM THE SAME DATASET (TAKEN FROM GOLDENSOFTWARE)

It is important to understand that different interpolation methods have their strength and weaknesses when using different datasets. It is not correct to generalise that a given interpolation method is better than the other without taking into account, the type and nature of the dataset and phenomenon involved. Therefore, it is important to select an appropriate interpolation technique when delineating.

Different spatial interpolation methods have been developed in different domains for different applications. Tobler's first Law of Geography, states everything is related to everything else, but near things are more related than distant things. This forms the general principle of many interpolation methods. Spatial interpolation methods are classified into two major groups (Ikechukwu *et al.*, 2017):

*Type 1* – Mechanical/ deterministic/non-geostatistical

*Type 2* – Linear/stochastic/geostatistical

Three interpolation methods are considered for the purpose of delineation and is discussed in the following sections.

#### **5.4.1 Inverse Distance Weighting (Type 1)**

The IDW function should be used when the set of points is dense enough to capture the extent of local surface variation needed for analysis. IDW determines cell values using a linear-weighted combination set of sample points. The weight assigned is a function of the distance of an input point from the output cell location. The greater the distance, the less influence the cell has on the output value.

#### **5.4.2 Splines (Type 1) (After Ikechukwu et al., 2017)**

Splines belong to a group of interpolators called Radial Basis Functions (RBF). Methods in this group include Thin-Plate Spline, Regularized Spline with Tension, and Inverse Multi-Quadratic Spline. These models use mathematical functions to connect the sampled data points. They produce continuous surfaces while limiting the bending of the surface produced to a minimum. RBF models are best employed in smooth surfaces for which the available sample data size is large as their performance is less than optimum for surfaces with appreciable variations spanning short ranges. RBF does not force estimates to maintain the range of the sampled data in these models.

Spline functions are the mathematical equivalents of the flexible ruler cartographers used, called splines, to fit smooth curves through several fixed points. It is a piecewise polynomial consisting of several sections, each of which is fitted to a small number of points in such a way that each of the sections join up at points referred to as break points. Splines are normally fitted using low order polynomials (*i.e.* second or third order) constrained to join up. The smoothing spline function also assumes the presence of a measurement error in the data that needs to be smoothed locally. Among the many versions and modifications of spline interpolators, the most widely used technique is the thin-plate splines as well as the regularized spline with tension and smoothing.

### 5.4.3 Kriging (Type 2)

Kriging is one of several methods that use a limited set of sampled data points to estimate the value of a variable over a continuous spatial field. Kriging also generates estimates of the uncertainty surrounding each interpolated value.

In a general sense, the kriging weights are calculated such that points nearby to the location of interest are given more weight than those farther away. Clustering of points is also taken into account, so that clusters of points are weighted less heavily (in effect, they contain less information than single points). This helps to reduce bias in the predictions.

Kriging will in general not be more effective than simpler methods of interpolation if there is little spatial autocorrelation among the sampled data points (that is, if the values do not co-vary in space). If there is at least moderate spatial autocorrelation, however, kriging can be a helpful method to preserve spatial variability that would be lost using a simpler method.

Kriging can be understood as a two-step process: first, the spatial covariance structure of the sampled points is determined by fitting a variogram; and second, weights derived from this covariance structure are used to interpolate values for unsampled points or blocks across the spatial field.

#### *The Variogram*

A variogram (sometimes called a “semi-variogram”) is a visual depiction of the covariance exhibited between each pair of points in the sampled data. For each pair of points in the sampled data, the gamma-value or “semi-variance” (a measure of the half mean-squared difference between their values) is plotted against the distance, or “lag”, between them. The “experimental” variogram is the plot of observed values, while the “theoretical” or “model” variogram is the distributional model that best fits the data. Variogram models are drawn from a limited number of “authorized” functions, including spherical, exponential, and Gaussian models as shown in Figure 52.

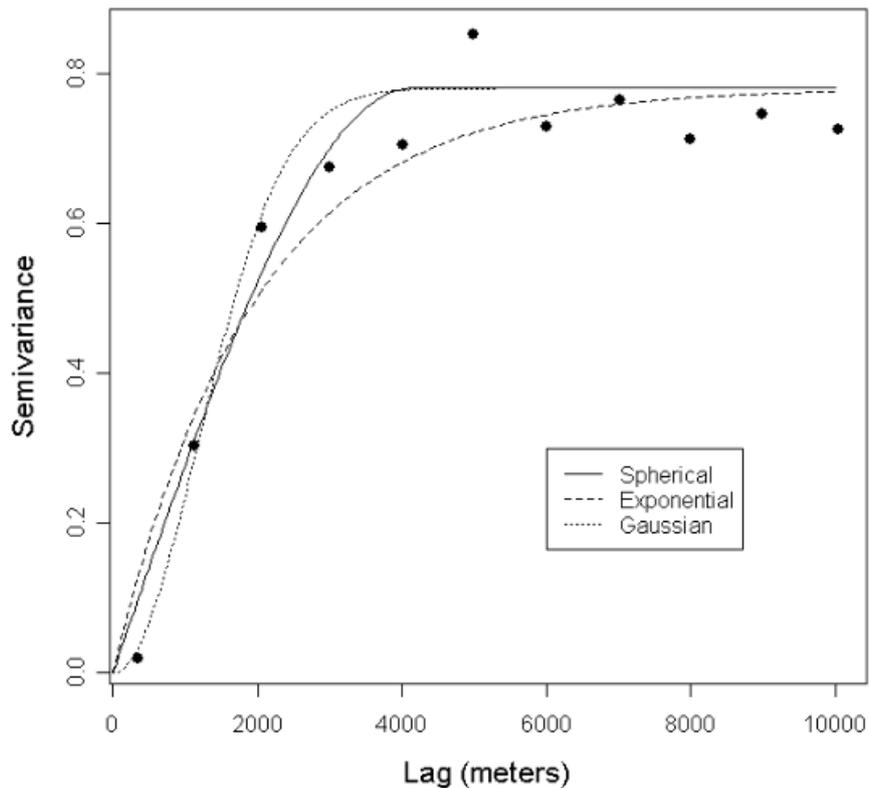


FIGURE 52 – EXAMPLES OF VARIOGRAM MODELS (BOHLING, 2005)

## 5.5 DELINEATION

Finally, applying the appropriate interpolation technique on the inferred transmissivity class will then represent the delineated groundwater units. This delineation is purely dependent on yield values although it has been shown that similar results are obtained when applying Equation 4 which requires additional parameters.

The recommended interpolation technique to apply, would be the one which results in the smallest RMSE (Appendix C) between observed and simulated values. The interpolation technique may vary from one study area to another depending on the yield data distribution.

Since water quality can be highly variable due to various factors, it is not explicitly incorporated into the inferred transmissivity classification delineation, but it is recommended to rather use as secondary delineation criteria based on specific user requirements.

Although it is recommended that the interpolation technique with the smallest RMSE is used, visual inspection of the result is important. In areas tested the spline interpolation method with minimum curvature gave visually good results even though the RMSE for this method was not always the smallest.

## 6 ALGORITHMS

The algorithms applied to the compiled Vegter database is presented in this section in the form of flow charts or where appropriate explicit mathematical formulations. The diagrams refer to the GSA database which is the Vegter database in the context of the software tool that was developed.

### 6.1 STATISTICAL CALCULATIONS

The statistical calculations as described by the Vegter methodology is presented in Chapter 4 where the majority of the analysis are performed in the context of geology. This section presents the statistical algorithms used in the developed software tool. The statistical calculations are basically the same as presented in Chapter 4, with the main difference being that the context of the analysis is not fixed to that of the geology, but can be a selection of the following list:

- Quaternary catchment
- Geology
- Groundwater occurrence
- Aquifer type
- Borehole depth
- Borehole elevation
- Aquifer vulnerability

#### 6.1.1 Borehole Density

The borehole density is calculated making use of the flow chart presented in Figure 53 and is the implementation of Equation 1. This algorithm is applied to the study area as the area is represented by a polygon with known area. The algorithm can be applied to any of the contextual features as long as a polygon or area exists for the required feature within the boundary of the defined study area.

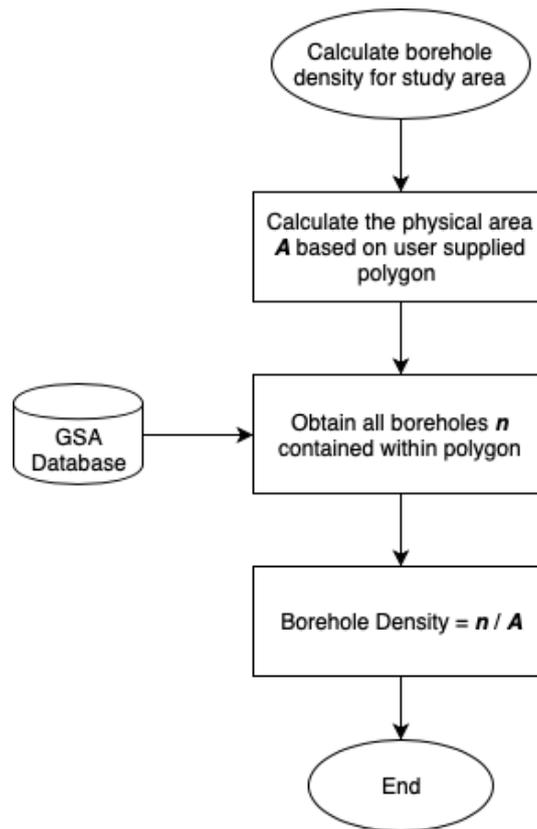


FIGURE 53 – BOREHOLE DENSITY ALGORITHM

### 6.1.2 Borehole Parameter Histograms

The list of borehole parameters that are expressed as histograms based on the context selection include:

- Water level
- Water strike and borehole depth
- Blow yield
- Borehole yield (boreholes with and without dyke intersections)
- EC

A generic algorithm is presented in Figure 54 to generate a histogram based on a specific parameter  $p$  when a list of parameter values is available. The minimum value of the parameters is always considered to be zero. The standard deviation (Appendix D) is also displayed for each histogram.

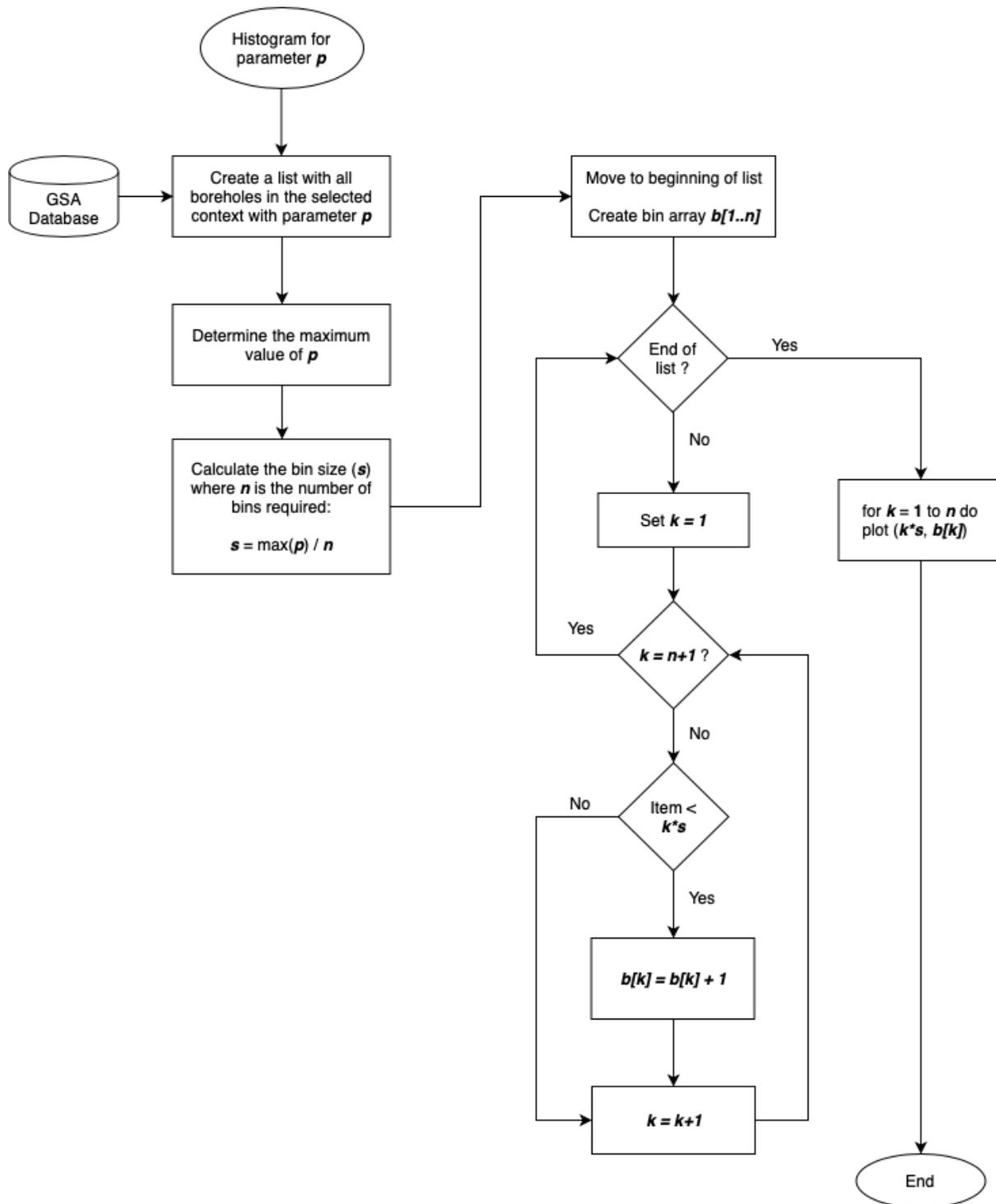


FIGURE 54 – GENERIC HISTOGRAM ALGORITHM

### 6.1.3 Water Strike vs Yield

The water strike vs. yield algorithm only uses the water strike and associated yield values for each borehole within the selected context and plots these two parameters against each other in a scatter plot. The Pearson correlation coefficient is displayed (Appendix E). The high-level flow chart of this algorithm is presented in Figure 55.

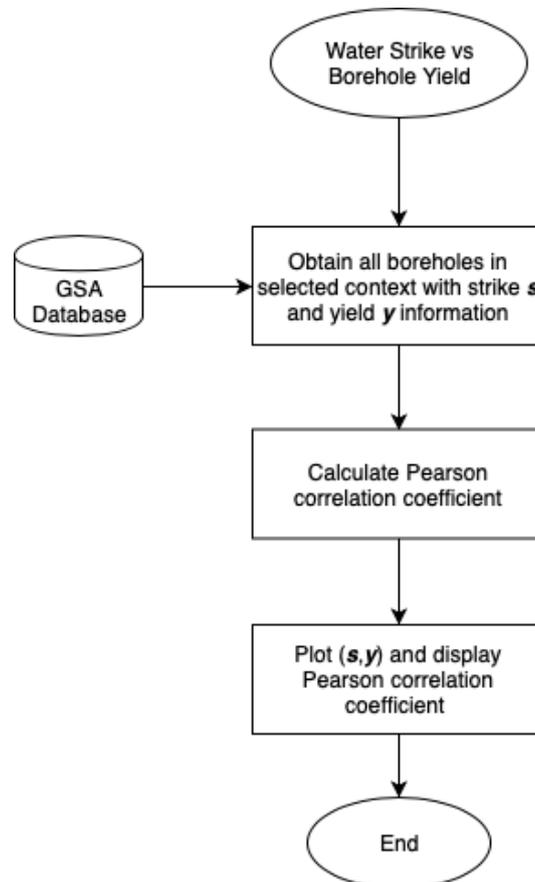


FIGURE 55 – WATER STRIKE VS. YIELD ALGORITHM

### 6.1.4 Water Level Correlation

The water level correlation algorithm uses the water level and associated elevation values for each borehole within the selected context and plots these two parameters against each other in a scatter plot. The Pearson correlation coefficient is displayed (Appendix E). The high-level flow chart of this algorithm is presented in Figure 56.

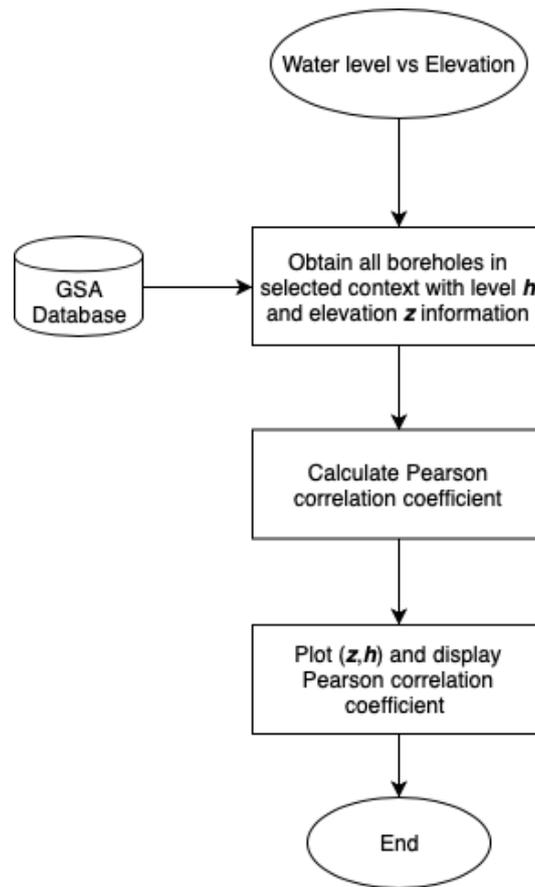


FIGURE 56 – WATER LEVEL CORRELATION ALGORITHM

## 6.2 WATER CHARACTER

The water character is depicted by a Piper plot and the plotting procedure is outlined in Appendix B. The high-level flow chart for this algorithm is presented in Figure 57.

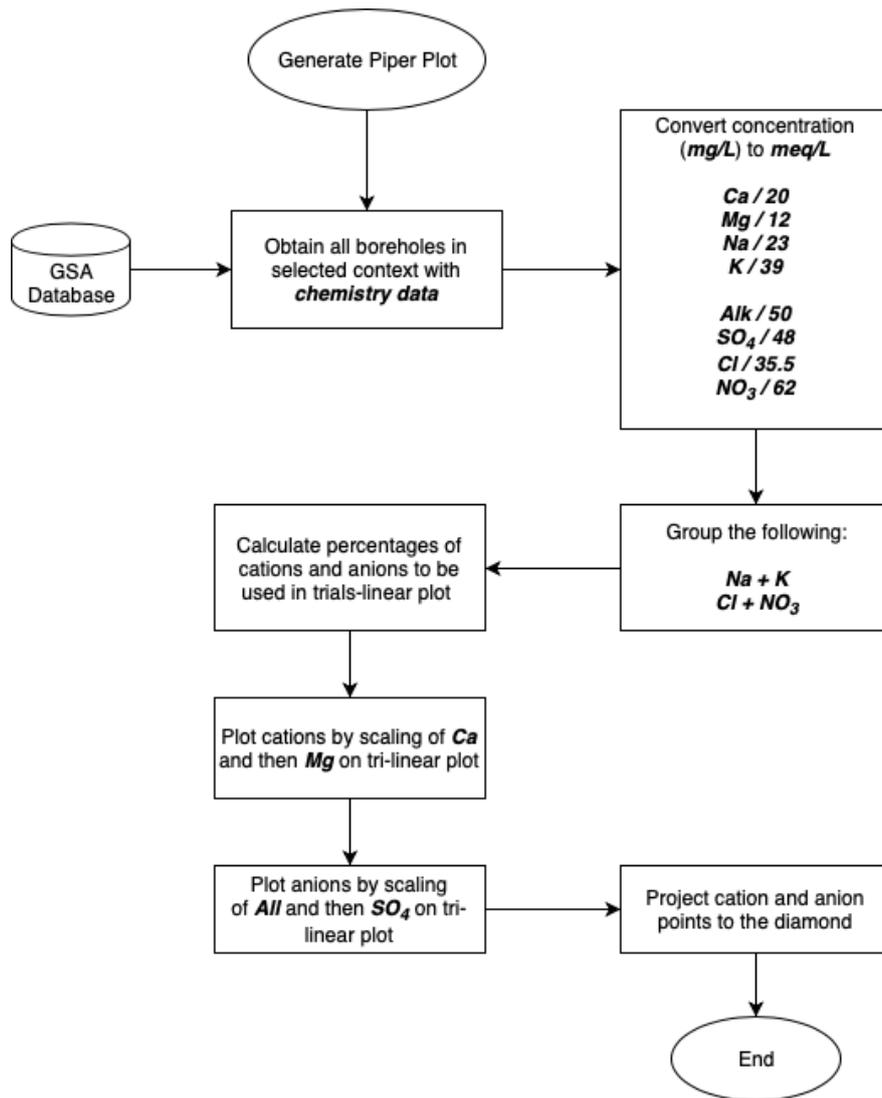


FIGURE 57 – WATER CHARACTER ALGORITHM

### 6.3 DELINEATION

The following delineation algorithms are presented here:

- Delineation classification tree (proposed method)
- Cooper-Jacob delineation (based on Equation 4)

#### 6.3.1 Delineation Classification Tree

The delineation algorithm is represented in a flow chart as shown in Figure 58 and is based on the inferred transmissivity classification tree (Figure 50) discussed earlier in this report. The RMSE calculation is documented in Appendix C and the various interpolation methods are described in Appendix F-Appendix H.

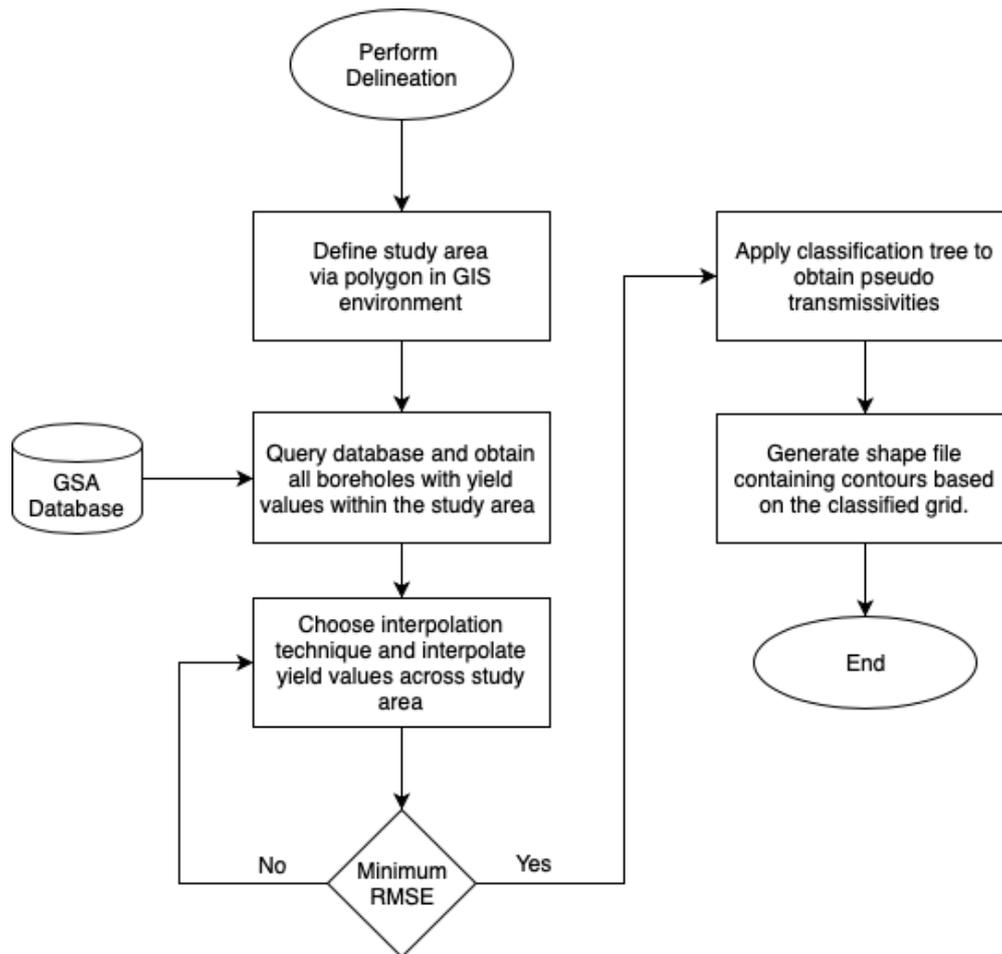


FIGURE 58 – INFERRED TRANSMISSIVITY DELINEATION ALGORITHM

### 6.3.2 Cooper-Jacob Delineation

As an alternative method to calculate the inferred transmissivity, the Cooper-Jacob equation (Equation 4) is used. Interpolated grids are calculated for water level, strike position and yield values. These grids are then used together with a  $r = 0.08 \text{ m}$  and  $t = 365 \text{ days}$ . The inferred transmissivity is then calculated according to Equation 4. The algorithm flow chart is shown in Figure 59. Note that checks are performed for when the water strike is above the water level and when blow yield is zero. The RMSE calculation is documented in Appendix C and the various interpolation methods are described in Appendix F-Appendix H.

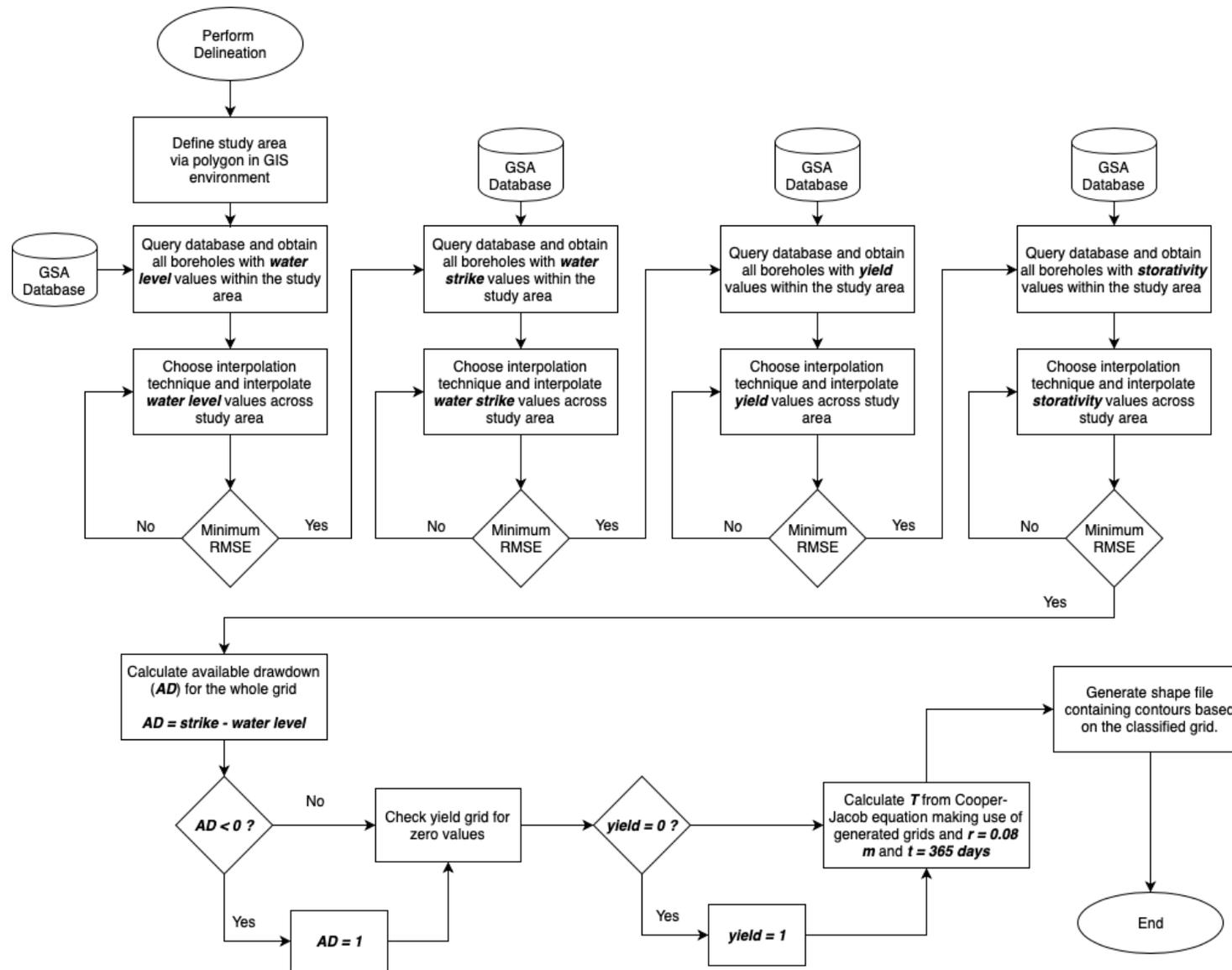


FIGURE 59 – INFERRED TRANSMISSIVITY DELINEATION ALGORITHM BASED ON COOPER-JACOB

## 7 IMPLEMENTATION

This section only focusses on the delineation methodology as the statistical calculations applied remains the same as described by Vegter. Both the delineation based on the classification tree as well as the Cooper-Jacob method are presented here for comparison purposes.

### 7.1 STUDY AREA

The study area considered here is Vegter Region 7 due to the availability of transmissivity data. The simplified geological map of the study area shown in Figure 60 and the groundwater occurrence map is shown in Figure 61.

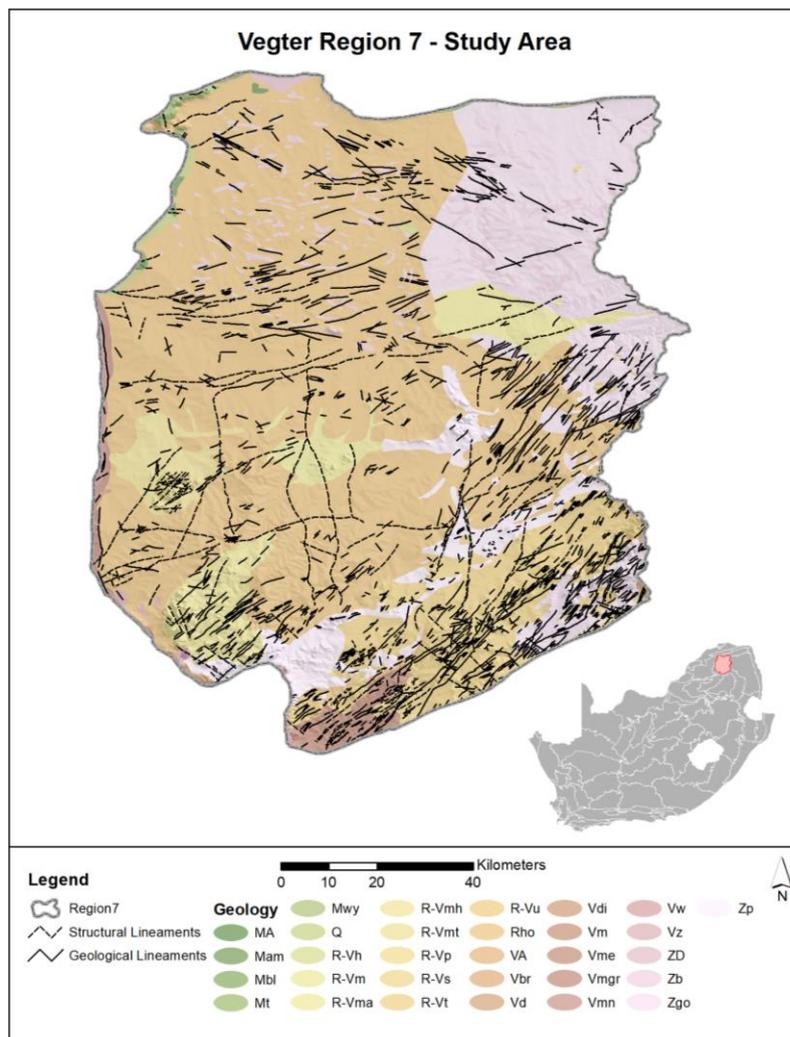


FIGURE 60 – VEGTER REGION 7 GEOLOGICAL MAP

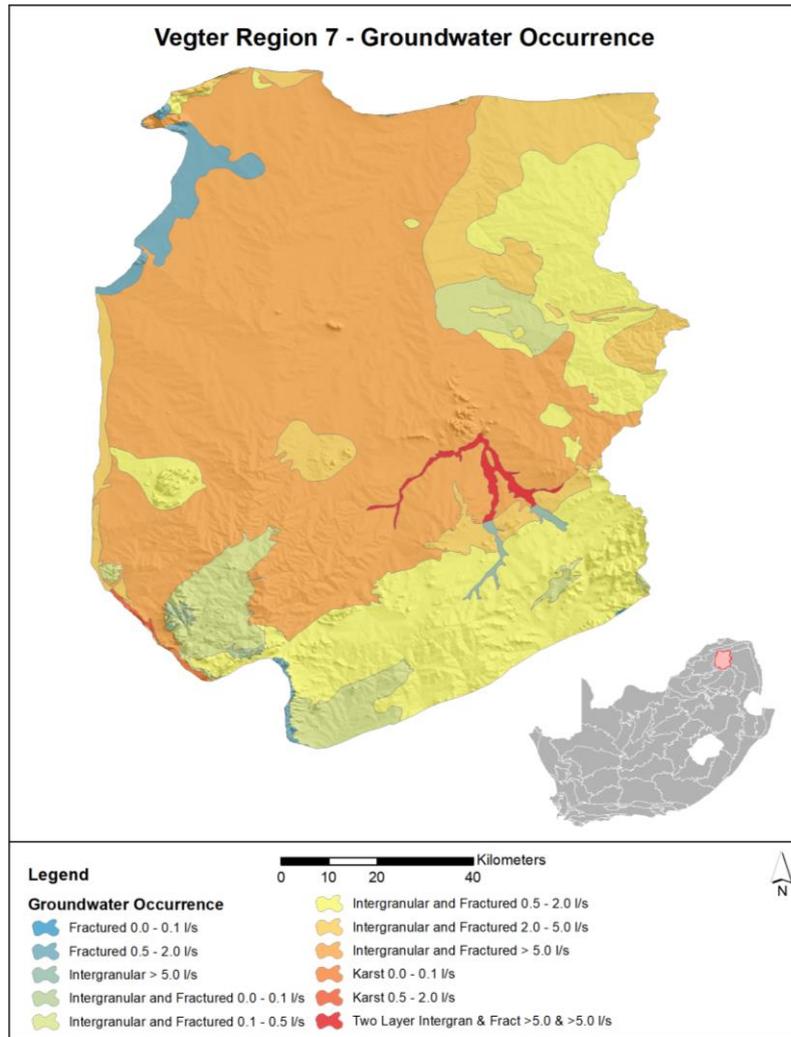


FIGURE 61 – VEGTER REGION 7 GROUNDWATER OCCURRENCE MAP

The transmissivity map of the area based on the GRAII dataset is presented in Figure 62. It is clear from the comparison of the aforementioned maps a correlation exists between the simplified geology and groundwater occurrence which is expected as the simplified geology was used as basis for the geohydrological map which presents the groundwater occurrence.

There is however no distinct correlation between the GRAII transmissivity map and the simplified geology as well as the groundwater occurrence map. It should be noted that geology is a 3D feature and that the geological maps only represent surface geology. Drilled boreholes can intersect multiple lithologies and is therefore not necessarily representative of the surface geology in which they were drilled. The transmissivities for the GRAII is calculated making use of Equation 10 (DWAF, 2006).

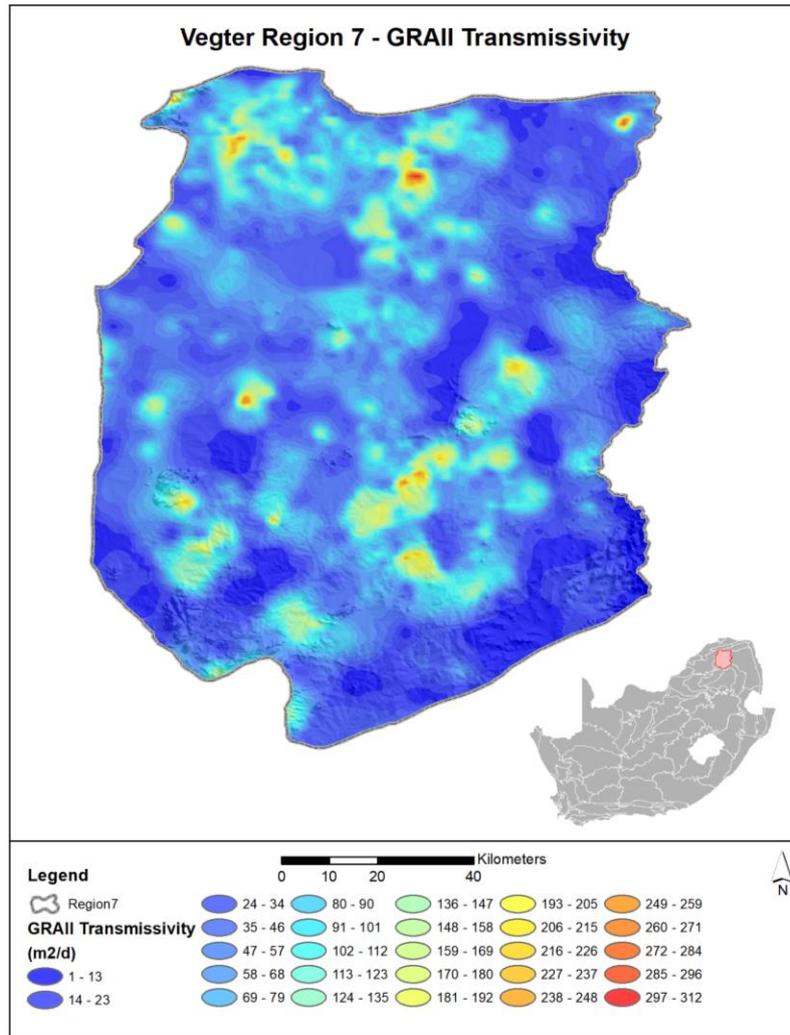


FIGURE 62 – VEGTER REGION 7 GRAII TRANSMISSIVITY MAP

## 7.2 CLASSIFICATION TREE DELINEATION

The borehole distribution within a 25 km buffer zone around Vegter region 7 is shown in Figure 63. The buffer zone is chosen to allow for proper interpolation across the study area boundary. The boreholes with actual transmissivity values compare well with the boreholes that have yield values. This is important for comparison purposes to see how well the inferred transmissivity compare to that of the actual transmissivity.

Spline interpolation with minimum curvature is applied to the actual transmissivity values as well as the inferred transmissivities produced by the classification tree. Both result sets together with the GRAII transmissivity dataset were classed according to Equation 11 and the results are shown in Figure 64. The boreholes are shown to see how the results relate to borehole distribution. The correlation of the classification tree to that of the actual transmissivity values are presented in Table 20.

## Vegter Region 7 - Yield and Transmissivity Boreholes

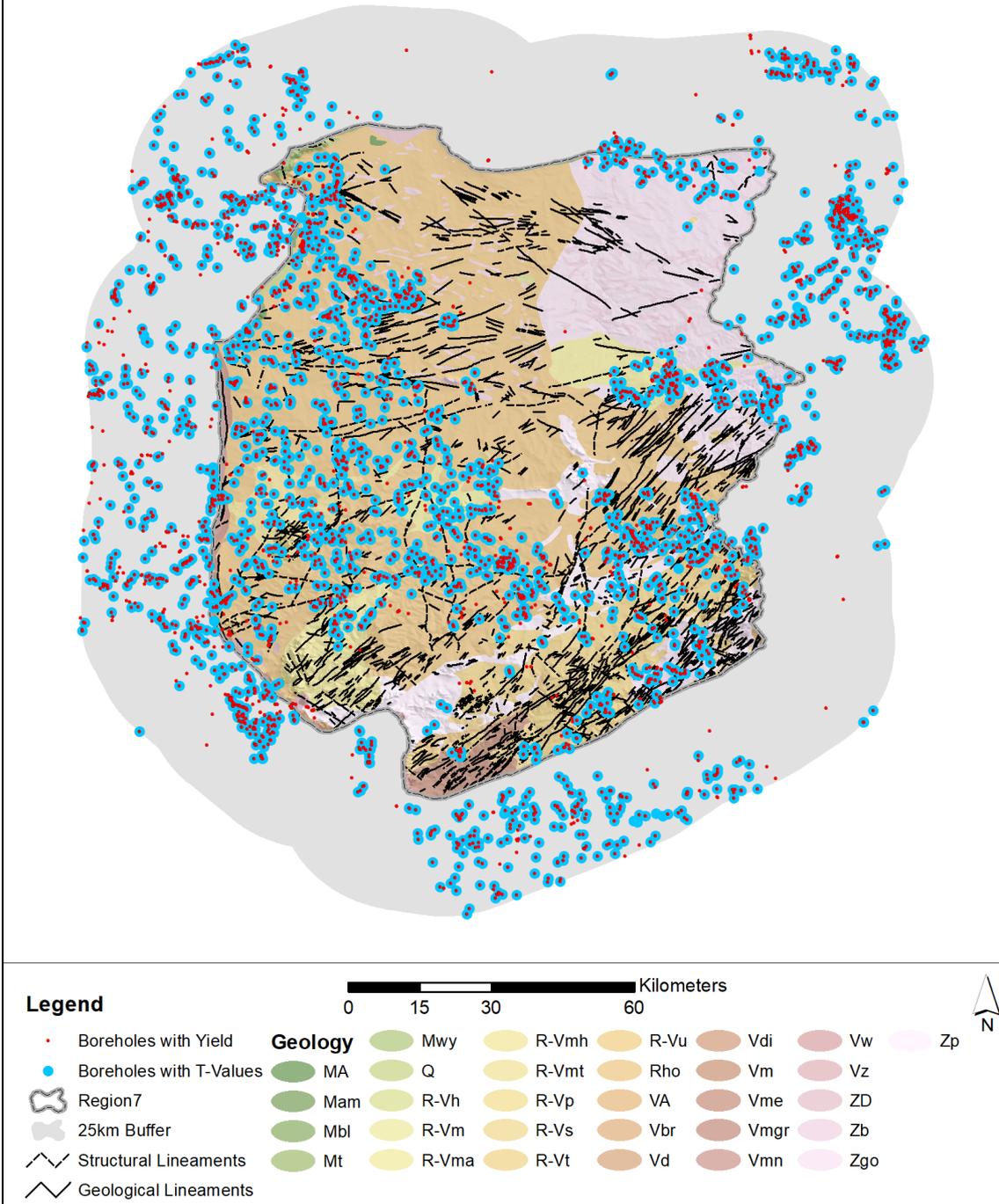


FIGURE 63 – VEGTER REGION 7 BOREHOLES WITH YIELD AND TRANSMISSIVITY VALUES

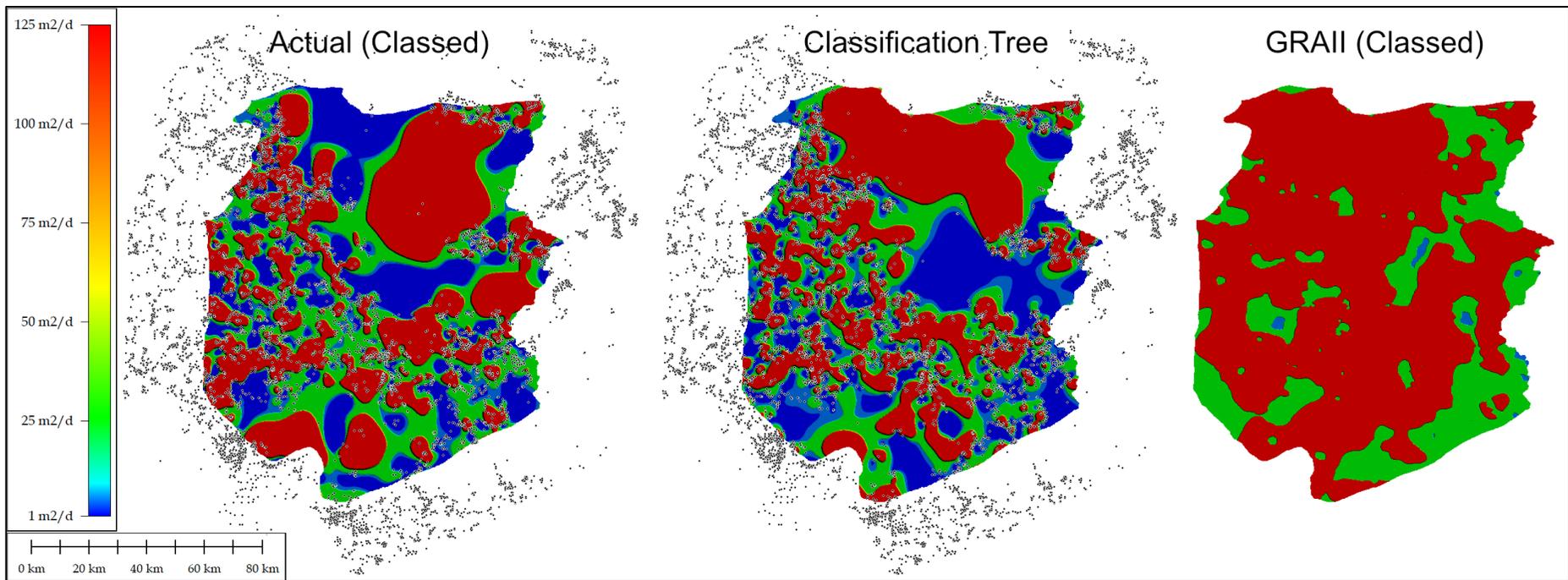


FIGURE 64 – VEGTER REGION 7 CLASSIFICATION TREE RESULTS

TABLE 20 – REGION 7 CLASSED TRANSMISSIVITY CORRELATION WITH CLASSIFICATION TREE

<i>Classed Dataset</i>	<i>Pearson Correlation</i>
Classification Tree	0.32
GRAII (Classed)	0.12

### 7.3 COOPER-JACOB DELINEATION

The borehole distribution within a 25 km buffer zone around Vegter region 7 is shown in Figure 65. The buffer zone is chosen to allow for proper interpolation across the study area boundary. The boreholes with actual transmissivity values (Figure 63) compare well with the boreholes presented in Figure 65. This is important for comparison purposes to see how well the inferred transmissivity compare to that of the actual transmissivity.

Spline interpolation with minimum curvature was applied to the actual transmissivity values as well as the inferred transmissivities produced by the Cooper-Jacob method (Equation 4). Both result sets together with the GRAII transmissivity dataset are classed according to Equation 11 and the results are presented in Figure 66. The correlation of the Cooper-Jacob delineation (Equation 4) to that of the actual transmissivity values are presented in Table 21.

The Cooper-Jacob method exhibits a slight decrease in correlation when compared to the classification tree method which is also the observation in Table 19. Furthermore, the Cooper-Jacob delineation process is computationally more expensive due to the multiple interpolation grids that is required (Figure 59).

## Vegter Region 7 - Yield, Strike and Water Level Boreholes

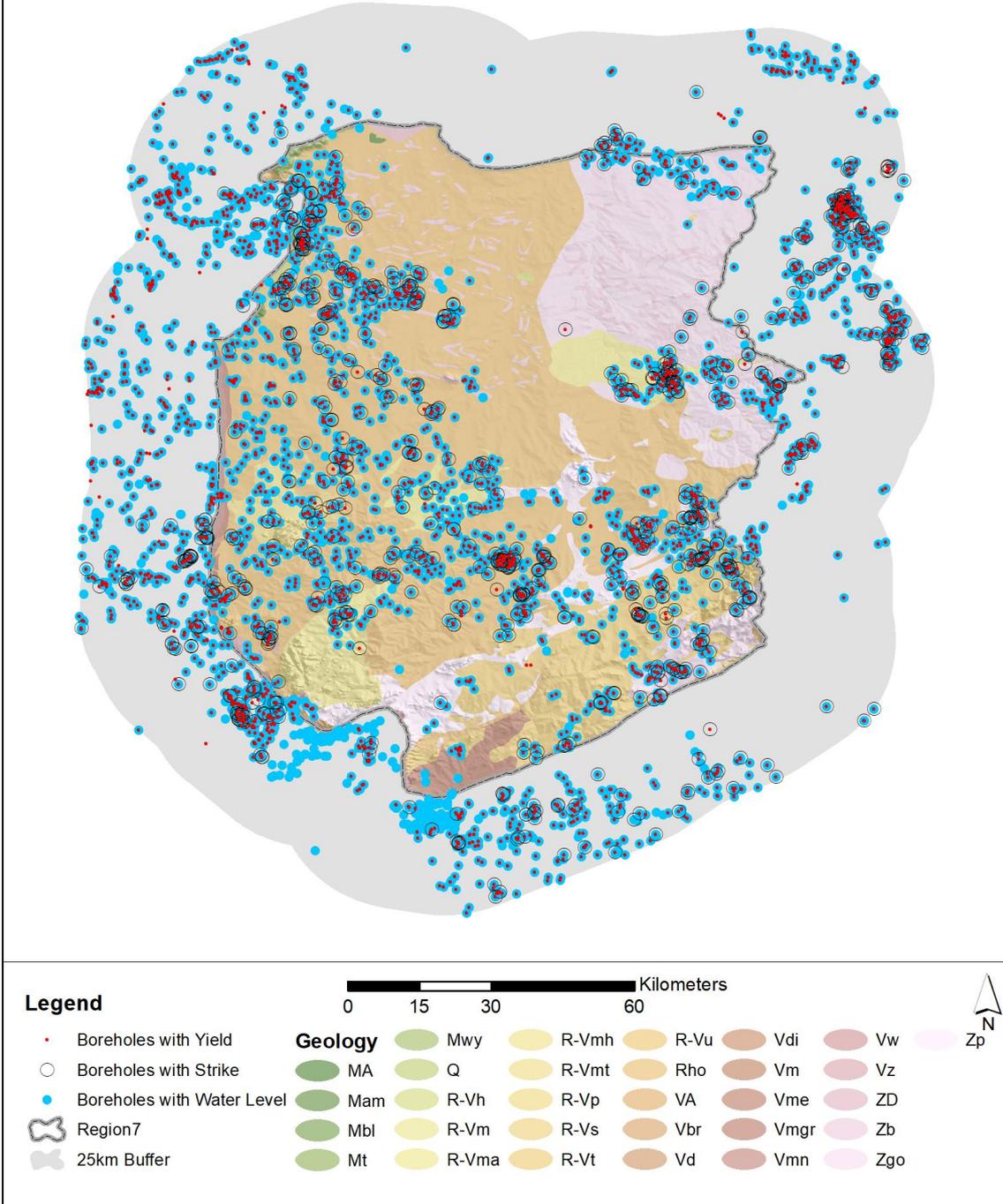


FIGURE 65 – VEGTER REGION 7 BOREHOLES WITH YIELD, STRIKE AND WATER LEVEL VALUES

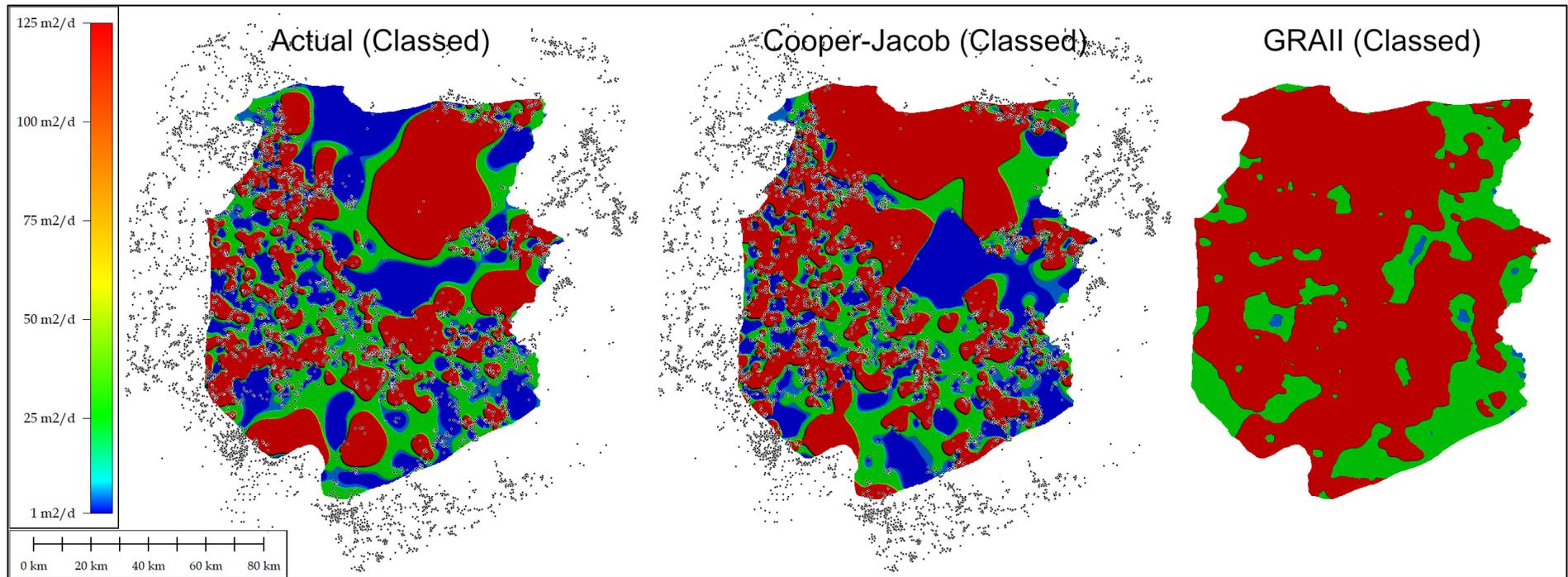


FIGURE 66 – VEGTER REGION 7 COOPER-JACOB RESULTS

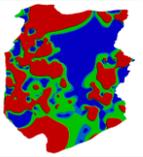
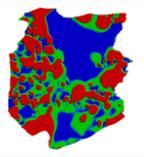
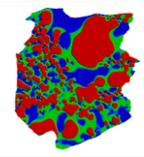
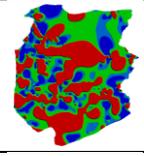
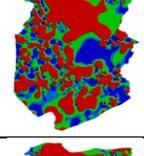
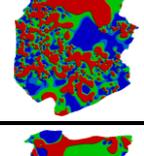
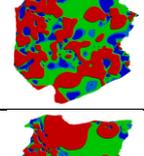
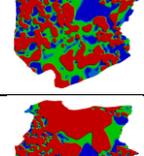
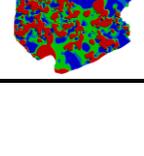
TABLE 21 – REGION 7 CLASSED TRANSMISSIVITY CORRELATION WITH CLASSIFICATION TREE

<i>Classed Dataset</i>	<i>Pearson Correlation</i>
Cooper-Jacob (Classed)	0.28
GRAII (Classed)	0.12

#### 7.4 DATA DENSITY

Since the delineation methods makes use of spatially distributed data, it is accepted that data density will affect the delineation result. The higher the data density the more refined the delineation result will be. The data sets in the preceding sections were reduced twice by 50% to inspect the effect of lowering the data density. The correlation among the various datasets are presented in Table 22 where  $T$  denotes actual transmissivity and  $CT$  denotes the inferred transmissivity obtained from the classification tree.  $CJ$  denotes the inferred transmissivity obtained from the Cooper-Jacob method.

TABLE 22 – EFFECT OF DATA DENSITY ON CORRELATION

				
		25% <sub>T</sub>	50% <sub>T</sub>	100% <sub>T</sub>
	25% <sub>CT</sub>	0.19	0.02	0.13
	50% <sub>CT</sub>	0.13	0.22	0.25
	100% <sub>CT</sub>	0.13	0.05	0.32
	25% <sub>CJ</sub>	0.20	0.12	0.04
	50% <sub>CJ</sub>	0.28	0.15	0.11
	100% <sub>CJ</sub>	0.20	0.11	0.28

It is clear that data distribution will play a major role in the delineation results. The best correlation obtained was 0.32 when using 100% of the available dataset and making use of the classification tree.

The correlation between the inferred transmissivities for the classification tree and the Cooper-Jacob method (Equation 4) is presented in Table 23.

TABLE 23 – REGION 7 CORRELATION BETWEEN CLASSIFICATION TREE AND COOPER-JACOB

<i>Dataset</i>	<i>Pearson Correlation</i>
25%	0.63
50%	0.66
100%	0.72

## 7.5 COMPARATIVE RESULTS

Throughout this report the Vegter analysis of Region 65 (Project K5/2251/1) was presented for illustration purposes only and therefore it is fitting to compare the newly developed delineation methodology results with that of the delineation done for Region 65.

The comparison is shown in Figure 67 and as expected, distinct differences exist between the delineation results as the previous methodology used four parameters to determine a delineation class. The new methodology only relies on borehole yield values to perform delineation based on a classification tree.

What is of interest in these results is that the majority of the boreholes in the northern part of the region have all been classified into a single inferred transmissivity class due to the high yields as opposed to the detailed delineation regions observed from the previous methodology mainly driven by the borehole density in that region (Figure 43). As mentioned in an earlier section, the advantage of the new delineation methodology, is that the results are continuous across the boundaries.

## 7.6 SOFTWARE TOOL

A software tool was developed to perform the delineation making use of the classification tree and the developed database. The user manual for the software tool is available in Appendix I.

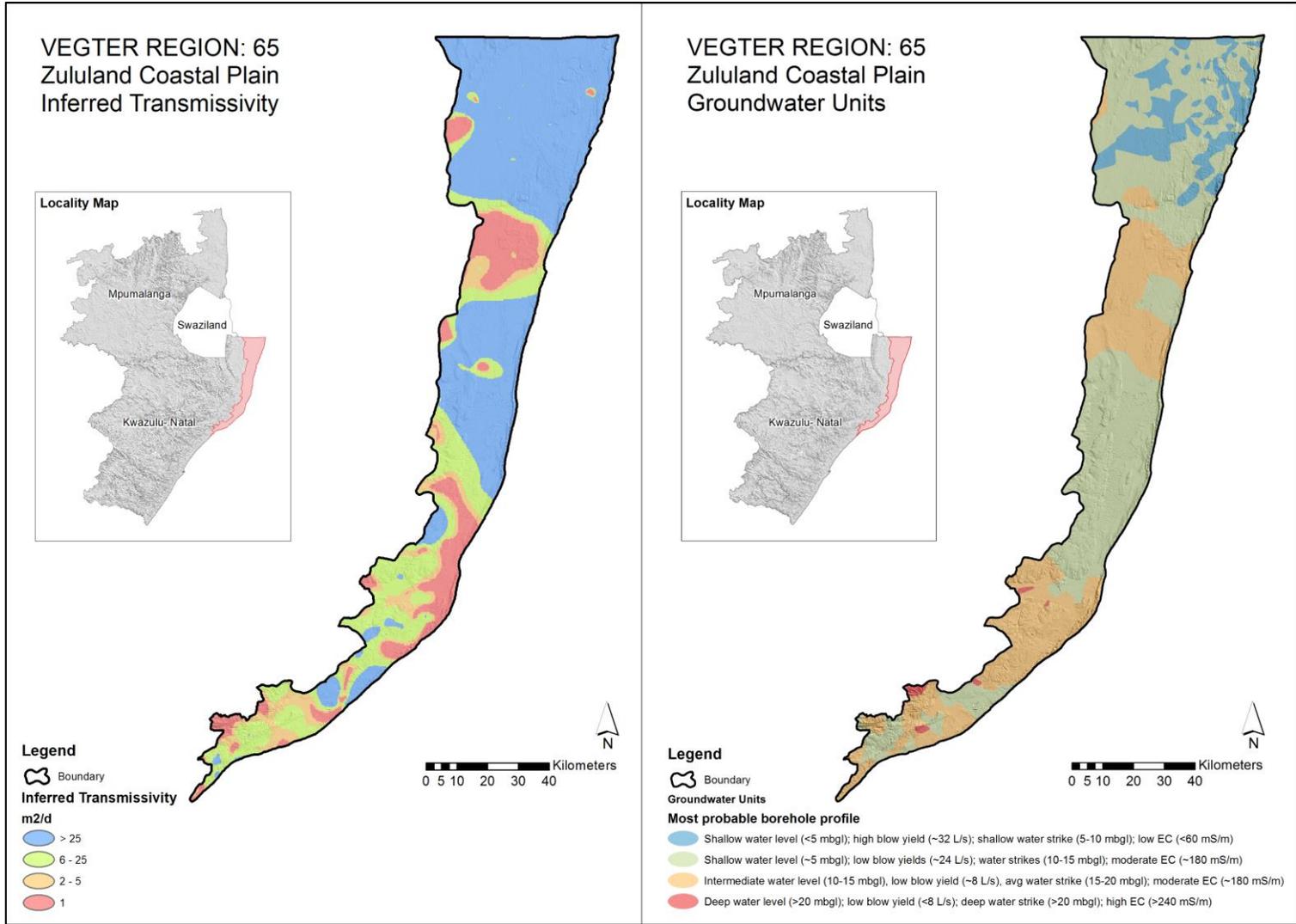


FIGURE 67 – VEGTER REGION 65 DELINEATION COMPARISON

## 8 CONCLUSIONS AND RECCOMENDATIONS

The aim of this project was to automate the geostatistical analysis and sub-delineation of all Vegter regions by means of a software tool.

The formulation of the geostatistical analysis is described in the Vegter methodology and by setting up an appropriate database, these analyses are automated for any Vegter region or sub-region. No changes were made to the exiting methodology, but the concept of contextual analysis is introduced. Traditionally, the Vegter analyses were only performed in the context of the simplified geology. The developed software tool allows the use of the following parameters as the context of analysis in addition to the simplified geology as described by Vegter:

- Quaternary catchment
- Groundwater occurrence
- Aquifer type
- Borehole depth
- Borehole elevation
- Aquifer vulnerability

The developed tool succeeds in automating the geostatistical analysis and allows the user to inspect the analysis in various contexts to obtain a better understanding of the area.

Vegter states that sub-delineation of Vegter regions is required, but does not specify the process. Exiting delineation methodologies, e.g. the delineation class method seems to work in certain study areas for example the Zululand coastal plain, but failed in other areas due to the fact that it becomes increasingly difficult to formulate four unique classes in terms of *water level*, *blow yield*, *water strike depth* and *EC*. In addition, the following also poses a problem over large areas:

- The delineation class method considers the whole range of values for each parameter and transforms them to a delineation class based on the results of the frequency analysis. Therefore, when applied to large areas with a wide range of parameter values, delineation detail is lost in certain areas and will only be represented by a single class.

- Continuity is lost in delineation when considering adjacent areas which exhibit a contrast in parameter values, since the delineation classes for each area is formulated in terms of the frequency analysis of the total range of values.

To allow for continuity over adjacent areas and retain delineation detail across areas with highly contrasting parameter values, a delineation methodology was required that will relate the calculated delineation class to a physical property.

A classification algorithm based on borehole yield values is developed that resulted in an inferred transmissivity of the area. The result classed the inferred transmissivity into four classes:

- 0-1
- 1-5
- 5-25
- >25

The correlation between the actual and inferred transmissivity resulted in roughly 42%. The inferred transmissivity is also calculated making use of the Cooper-Jacob equation under certain assumptions, which resulted in a correlation of 38% compared to the actual transmissivity. The reason for these low correlations is attributed to the fact that the mean value of actual transmissivities obtained from pump tests are an order of magnitude smaller than the inferred transmissivities, although the probability distribution of higher values seem to be quite similar.

The delineation based on the inferred transmissivity based on the classification algorithm is then used as basis for delineation using an appropriate interpolation technique. It is recommended that the RMSE be used as guide in choosing the appropriate interpolation technique. Visual verification of the result is also very important. Spline interpolation with minimum curvature shows good results in the Vegter Region 7 case study.

It is further recommended that users buffer the Vegter region under consideration to ensure proper interpolation across the study area boundary, as interpolation methods perform poorly when extrapolating. This measure increases continuity across adjacent study areas.

The current classification tree is based on observed behaviour of borehole yields and actual transmissivities in the Limpopo province due to the availability of high-density data for each of the aforementioned parameters. It is however recommended to revise the classification algorithm as more data becomes available across the country.

Finally, the newly developed delineation methodology succeeds in producing continuous delineation across boundaries and accommodates areas with highly contrasting parameter values. Data density and the final interpolation method plays a major role in the resultant delineation, although general trends should be preserved.

## **REFERENCES**

- Bohling, G. (2005). Introduction to Geostatistics and Variogram analysis, Kansas Geological Survey.
- Dennis, I. and Dennis, S.R. (2012). Climate change vulnerability index for South African aquifers. Water SA Vol. 38 No. 3. International Conference on Groundwater Special Edition 2012.
- Dressler, M. (2007). Interpolation methods for construction of surfaces, Faculty of Mechatronics and Interdisciplinary Engineering Studies, Technical University of Liberec.
- DWAF (2004). Standard Descriptors for Geosites. Department of Water Affairs and Forestry, Pretoria, South Africa.
- DWAF (2006). Groundwater Resource Assessment II - Task 2C Groundwater Planning Potential. Department of Water Affairs and Forestry, Pretoria, South Africa.
- DWS (2017). Distribution of boreholes stored in the National Groundwater Archive (NGA). Available from [www.dwa.gov.za/Groundwater/Documents.aspx](http://www.dwa.gov.za/Groundwater/Documents.aspx).
- Ikechukwu, M.N., Ebinne, E., Idorenyin, U. and Raphael, N.I. (2017). Accuracy assessment and comparative analysis of IDW, Spline and Kriging in spatial interpolation of landform (topography): An experimental study. Journal of Geographic Information System. ISSN Online: 2151-1969.
- Seymore, A. (1994). Groundwater Resources of the Republic of South Africa. Geographic Information Systems, ARCIMS Mapservices, [www.dwaf.gov.za](http://www.dwaf.gov.za).
- Van Tonder, G.J., Van Sandwyk, L. & Buys, J. (1996). Program TRIPOL, version 1.0, Institute for Groundwater studies, and Water Research Commission.

## APPENDIX A - DRASTIC INDEX CALCULATION FOR SOUTH AFRICA

Groundwater vulnerability was considered in terms of the DRASTIC method of assessment of the intrinsic vulnerability of an aquifer to contamination from the surface (Parsons & Conrad, 1998). The method considers the following factors which control the vulnerability of an aquifer to contamination from surface:

Parameter	Input dataset
Depth to water table (D)	126 263 groundwater levels from the NGDB (for 4 280 of these, the mean groundwater level was calculated from time-series data) were interpolated to a groundwater level grid.
Recharge (R)	Recharge calculated as part of GRAII-3 project.
Aquifer material (A)	1:1 million Geology from CGS
Soils (S)	WR90 soils data set
Topography and slope (T)	DWAF 20 m DTM resampled to 1×1 km
Impact of the vadose (unsaturated) zone (I)	1:1 million Geology from CGS
Hydraulic conductivity (C)	1:1 million Geology from CGS

The overall DRASTIC equation is shown below:

$$Index = D_r D_w + R_r R_w + A_r A_w + S_r S_w + T_r T_w + I_r I_w + C_r C_w \quad (1A)$$

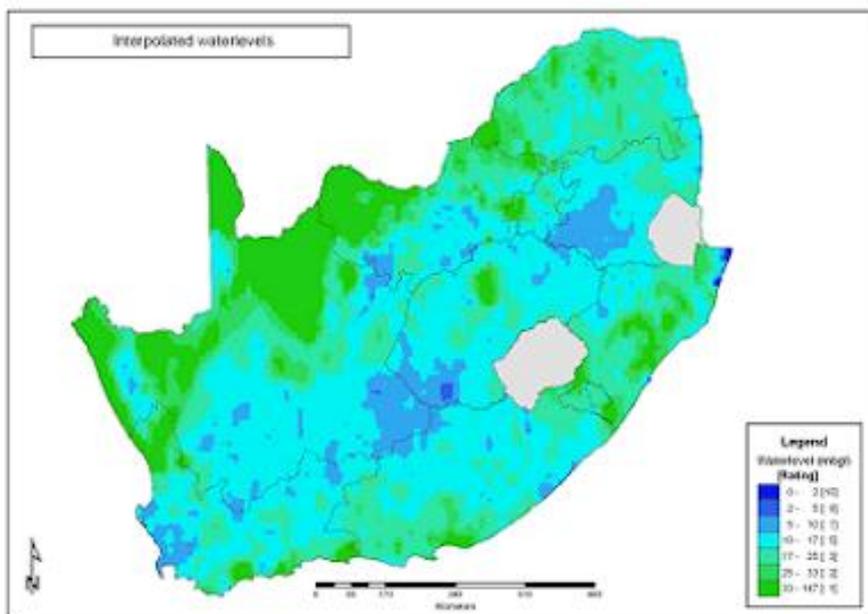
where  $r$  = rating and  $w$  = weighting

The ratings and weights used were based on the categories and values defined by Parsons & Conrad (1998) and are shown below.

## DRASTIC Ratings and Weights

*Table 1A – Depth to groundwater ratings used in the DRASTIC method*

Range (m)	Rating
0-1.5	10
1.5-4.5	9
4.5- 9.0	7
9.0-15.0	5
15.0-22.5	3
22.5-30.0	2
>30	1
Weight	5



*Figure 1A - Interpolated water levels*

*Table 2A -Recharge ratings used in the DRASTIC method*

Range (mm)	Rating
0-50	1
50-100	3
100-175	6
175-250	8
>250	9
Weight	4

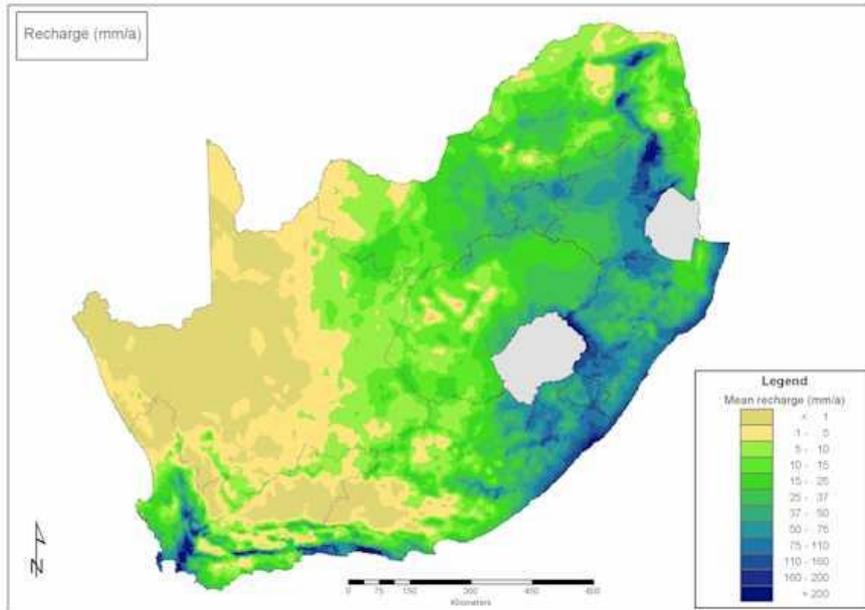


Figure 2A - Recharge output from GRAII-3

Table 3A -Topography (slope) ratings used in the DRASTIC method

Range (%)	Rating
<2	10
2- 6	9
6-12	5
12-18	3
> 18	1
Weight	1

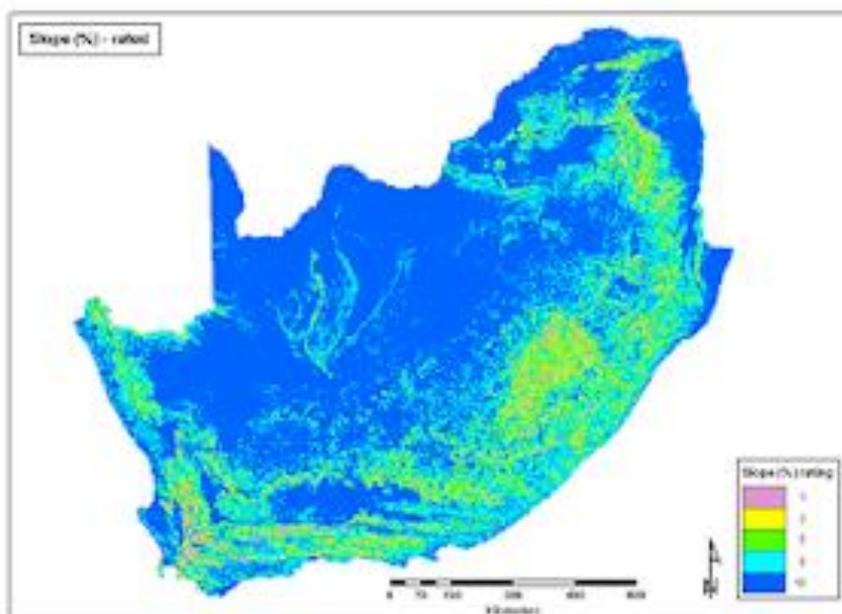


Figure 3A -Slope

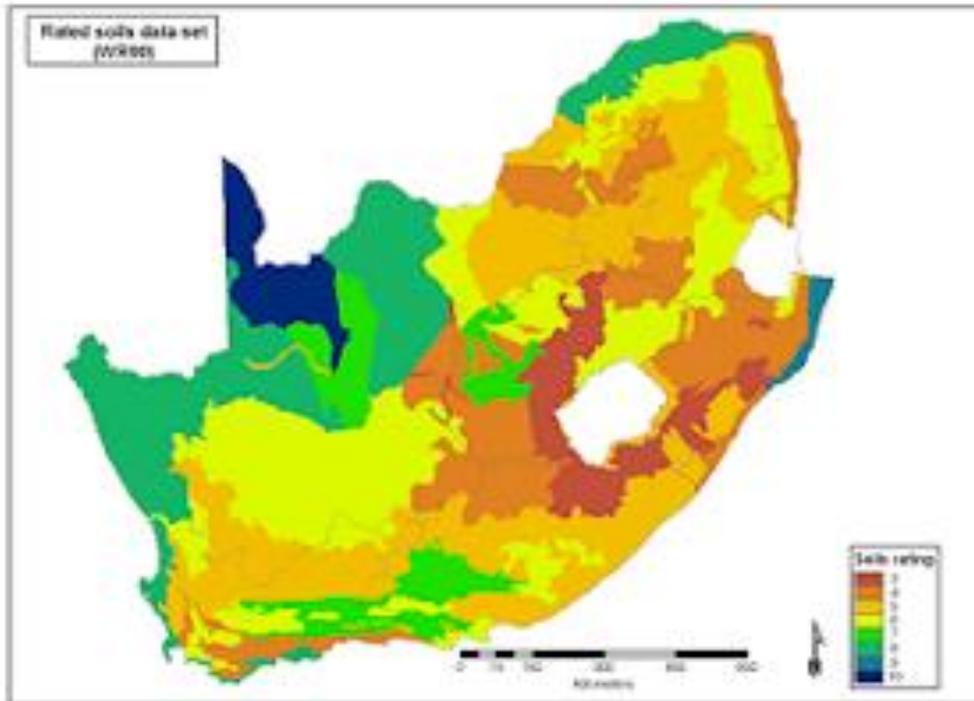


Figure 4A – Soils

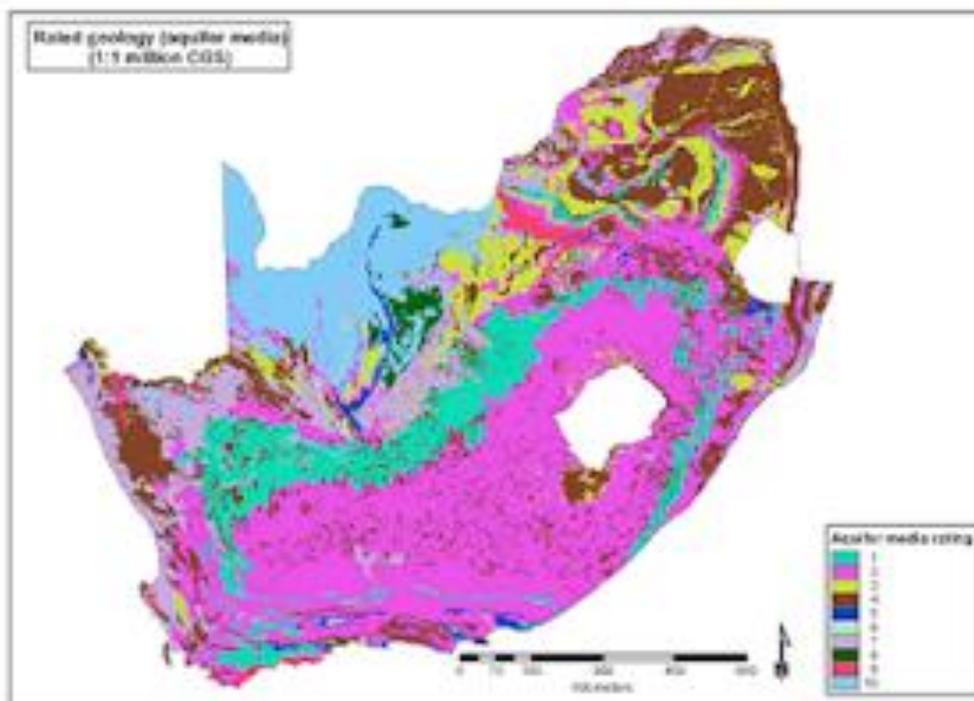


Figure 5A – Geology: Aquifer media

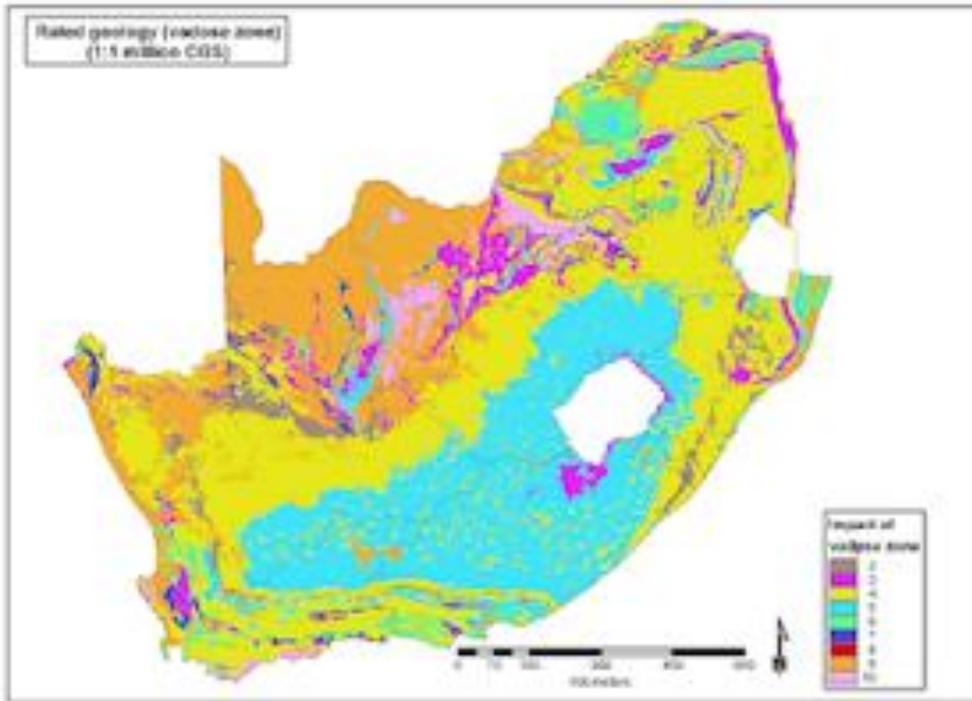


Figure 6A – Geology: Impact of vadose zone

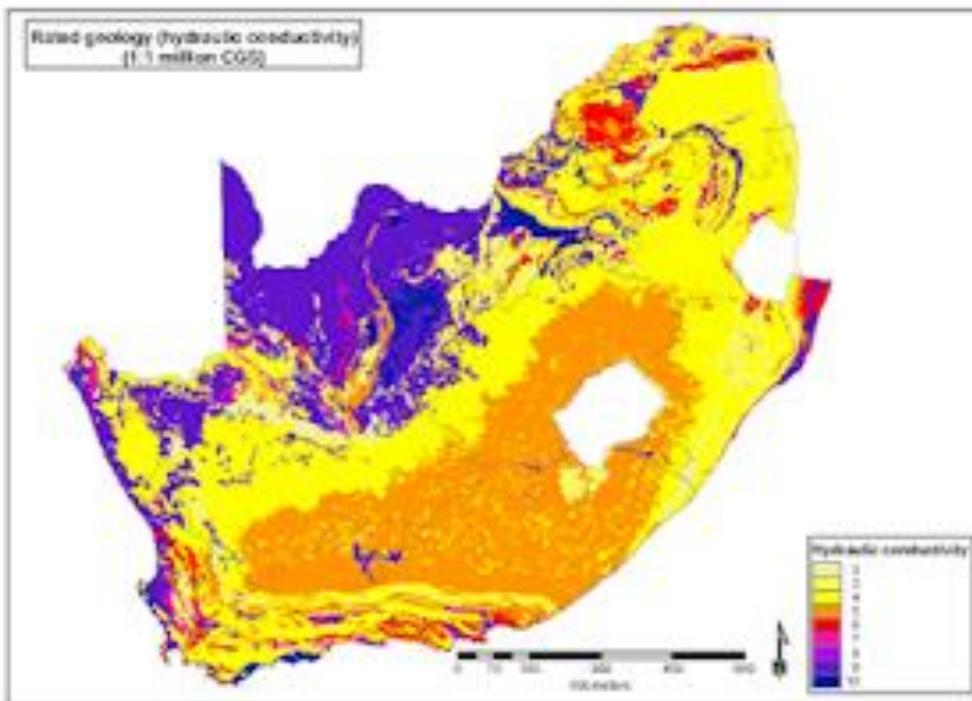


Figure 7A- Geology: Hydraulic conductivity

Table 4A- Soils ratings used in the DRASTIC method

Soil type	Rating
LmSa - 60 SaLm - SaCILm 35	5
LmSa - SaLm 20 SaCILm - 70	5
LmSa - SaLm 20 SaCILm - 75	5
LmSa - SaLm 25 SaCILm - Cl 70	4
LmSa - SaLm 35 SaCILm - SaCl 60	3
LmSa - SaLm 40 SaCILm - SaCl 55	3
LmSa - SaLm 45 SaCILm - 50	4
LmSa - SaLm 50 SaCILm - SaCl 45	4
LmSa - SaLm 60 SaCILm - 30	5
LmSa - SaLm 60 SaCILm - SaCl 30	4
LmSa - SaLm 65 SaCILm - SaCl 25	5
LmSa - SaLm 70 SaCl - Cl 30	5
LmSa - SaLm 70 SaCILm - 30	5
LmSa - SaLm 80 SaCILm - 15	6
LmSa - SaLm 100	6
Sa - 80 SaCILm - SaCl 15	8
Sa - 90 LmSa - SaLm 10	9
Sa - 90 SaCILm - SaCl 10	9
Sa - 100	10
Sa - LmSa 20 SaLm - SaCl 75	4
Sa - LmSa 25 SaLm - SaCILm 70	6
Sa - LmSa 50 SaLm - SaCILm 45	6
Sa - LmSa 50 SaLm - SaCILm 50	6
Sa - LmSa 60 SaLm - 35	7
Sa - LmSa 70 SaLm - SaCILm 25	7
Sa - LmSa 70 SaLm - SaCILm 30	7
Sa - LmSa 80 SaLm - SaCILm 15	7
Sa - LmSa 100	8
Sa - SaLm 100	8
SaCl - 100	5
SaCl - Cl 100	4
SaCILm - 20 SaCl - Cl 75	4
SaCILm - 30 SaCl - 65	4
SaCILm - 30 SaCl - Cl 50	4
SaCILm - 30 SaCl - Cl 65	4
SaCILm - 50 SaCl - 40	5
SaCILm - 60 LmSa - SaLm 35	5
SaCILm - 70 SaCl - Cl 25	5
SaCILm - 80 SaCl - Cl 20	5
SaCILm - 100	5
SaCILm - SaCl 15 LmSa - SaLm 80	5
SaCILm - SaCl 30 Sa - SaLm 60	6
SaCILm - SaCl 65 SaLm - 30	4
SaCILm - SaCl 100	3
SaLm - 15 SaCl - Cl 80	4
SaLm - 20 SaCILm - SaCl 70	4
SaLm - 25 SaCILm - SaCl 70	4
SaLm - 35 SaCILm - SaCl 60	5
SaLm - 60 SaCILm - 40	6
SaLm - 60 SaCILm - SaCl 35	5
SaLm - 60 SaCILm - SaCl 40	5
SaLm - 100	6
SaLm - SaCILm 25 SaCILm - SaCl 70	4
SaLm - SaCILm 30 SaCILm - SaCl 65	4
SaLm - SaCILm 60 SaCl - Cl 30	4
SaLm - SaCILm 100	5
Weight	2

Table 5A – Ratings used in the DRASTIC method based on underlying geology: aquifer media, impact of vadose zone and hydraulic conductivity

LABEL	LITHO_1	LITHO_2	LITHO_3	Aquifer media	Vadose Zone	Hyd_ Cond
Mpf	ALKALI-FELDSPAR SYENITE			2	3	3
Mj	AMPHIBOLITE	CALC-SILICATE ROCK		3	3	3
Ntu	AMPHIBOLITE	GNEISS	SCHIST	3	3	3
Rdo	AMPHIBOLITE	IRON FORMATION		3	3	3
Mha	ANDESITE	PYROCLASTIC		2	3	3
Mro	ANDESITE	DIORITE	PYROXENITE	2	3	3
Rd	ANDESITE	QUARTZ PORPHYRY	QUARTZITE	3	3	3
Rk	ANDESITE	TUFF		2	3	3
R-Vz	ANDESITE	DACITE	TUFF	2	3	3
Val	ANDESITE			3	3	3
Vh	ANDESITE			3	3	3
Vha	ANDESITE	TUFF	CONGLOMERATE	2	3	3
Vo	ANDESITE			3	3	3
Vri	ANDESITE	DACITE		3	3	3
Zme	ANORTHOSITE	SERPENTINITE	PYROXENITE	2	3	3
Cmk	ARENITE	CONGLOMERATE	SHALE	2	6	6
Dw	ARENITE	SHALE		2	5	5
Kma	ARENITE	CONGLOMERATE		3	6	6
Ks	ARENITE	MUDSTONE	SHALE	2	5	5
Ma	ARENITE	RUDITE	CONGLOMERATE	2	6	6
Mam	ARENITE	MUDSTONE		2	5	5
Mbl	ARENITE	RUDITE	CONGLOMERATE	2	6	6
Mc	ARENITE			4	6	6
Mf	ARENITE	BASALT		3	5	5
Mkr	ARENITE	QUARTZ PORPHYRY	BASALT	4	6	6
Mmb	ARENITE	SHALE	CONGLOMERATE	2	5	5
Mnz	ARENITE	SHALE	BASALT	2	5	5
Msm	ARENITE	CONGLOMERATE		3	6	6
Msw	ARENITE	TRACHYTOID		4	6	6
Mv	ARENITE	SHALE		2	5	5
Mwi	ARENITE	CONGLOMERATE		3	6	6
Mwy	ARENITE	CONGLOMERATE		3	6	6
Nf	ARENITE	CONGLOMERATE	SHALE	2	5	5
Nfi	ARENITE	SHALE		2	4	4
Nfl	ARENITE			4	6	6
Nh	ARENITE	DOLOMITE	DIAMICTITE	5	6	6
Nk	ARENITE	SHALE	CONGLOMERATE	2	4	4
Nkb	ARENITE	SHALE		2	4	4
Op	ARENITE	SHALE		2	4	4
Ope	ARENITE			4	6	6
O-S	ARENITE	SHALE		2	4	4
Pc	ARENITE	SHALE		2	4	4
Pko	ARENITE	SHALE		2	4	4
Pm	ARENITE	SHALE	COAL	2	4	4
Pr	ARENITE	SHALE		2	4	4
Pv	ARENITE	SHALE	COAL	2	4	4
Pwa	ARENITE	SHALE		2	4	4
Qb	ARENITE			4	6	6
Qpd	ARENITE	MUDSTONE	LIGNITE	2	5	5
Qs	ARENITE			4	6	6
Rw	ARENITE	SHALE		2	4	4
Sn	ARENITE	SHALE	TILLITE	2	4	4
T-Qn	ARENITE	SAND		5	6	6
TRc	ARENITE	SILTSTONE		3	5	5
TRm	ARENITE	MUDSTONE	SHALE	2	5	5
TRmc	ARENITE	MUDSTONE		2	5	5
TRnt	ARENITE			4	6	6
Vbt	ARENITE			4	6	6
Vhd	ARENITE	SILTSTONE	CONGLOMERATE	3	5	5
Vle	ARENITE	SHALE		2	4	4
Vlm	ARENITE	ARENITE	HORNFELS	4	6	6
Vlu	ARENITE	LIMESTONE		5	6	6
Vma	ARENITE			4	6	6
Vmg	ARENITE			4	6	6
Vrk	ARENITE			4	6	6
Vrw	ARENITE	ARENITE		4	6	6
Vsm	ARENITE	ANDESITE		4	6	6
Vst	ARENITE			4	6	6
Vsu	ARENITE			4	6	6
Zm	ARENITE	CONGLOMERATE	SHALE	2	5	5
Jdr	BASALT			4	3	3
Jl	BASALT			4	3	3

Jm	BASALT			4	3	3
Js	BASALT	TUFF	PYROCLASTIC BRECCIA	4	3	3
Ms	BASALT	QUARTZITE	CONGLOMERATE	4	3	3
Mt	BASALT	TUFF	ARENITE	4	3	3
Vdu	BASALT	ANDESITE		4	3	3
Zns	BASALT	ANDESITE	QUARTZITE	4	3	3
Mgn	CARBONATITE	PYROXENITE		5	9	9
Mno	CARBONATITE	SYENITE		5	9	9
Mtw	CARBONATITE			5	9	9
Mfr	CHARNOCKITE			3	3	3
Mst	CHARNOCKITE			3	3	3
Zk	CHERT	IRON FORMATION	SCHIST	3	10	10
Vc	CLINOPYROXENITE	HARZBURGITE	NORITE	3	4	4
Vvl	CLINOPYROXENITE	HARZBURGITE	NORITE	3	4	4
Vz	CLINOPYROXENITE	HARZBURGITE		3	4	4
Je	CONGLOMERATE	ARENITE		5	7	7
Kmb	CONGLOMERATE	ARENITE		5	7	7
Kmg	CONGLOMERATE	ARENITE		5	7	7
Nnt	CONGLOMERATE	MUDSTONE	LIMESTONE	4	6	6
Rka	CONGLOMERATE	SHALE		3	6	6
R-Vbo	CONGLOMERATE	QUARTZITE	SHALE	4	6	6
Zpr	CONGLOMERATE	LUTACEOUS ARENITE	LAVA	5	7	7
Nnu	DIAMICTITE	DOLOMITE	ARENITE	8	8	8
Vmk	DIAMICTITE	IRON FORMATION	ARENITE	7	7	7
Rro	DIORITE	GABBRO		4	3	3
Jd	DOLERITE			4	4	4
Vdi	DOLERITE			4	4	4
Vas	DOLOMITE	QUARTZITE		10	10	10
Vd	DOLOMITE	LIMESTONE	SHALE	7	9	9
Vgh	DOLOMITE	LIMESTONE	CHERT	8	10	10
Vm	DOLOMITE	CHERT		9	10	10
Vsc	DOLOMITE	SHALE		7	9	9
Vvo	DOLOMITE	IRON FORMATION	LAVA	9	10	10
Mka	DUNITE	PYROXENITE	NORITE	3	4	4
Z-Rm	DUNITE	HARZBURGITE	PYROXENITE	3	4	4
Mri	GABBRO	WEHRLITE	GRANITE	3	4	4
Nma	GABBRO	NORITE		3	4	4
Nmb	GABBRO	NORITE	PYROXENITE	3	4	4
Nti	GABBRO			3	4	4
Rmd	GABBRO	NORITE	PYROXENITE	3	4	4
Ru	GABBRO	GABBRO	GRANITE	3	4	4
Vbi	GABBRO			3	4	4
Vds	GABBRO	NORITE		3	4	4
Vmgr	GABBRO	NORITE		3	4	4
Vmn	GABBRO			3	4	4
Vpy	GABBRO	NORITE		3	4	4
Vrs	GABBRO	DIORITE		3	4	4
Vvi	GABBRO	ANORTHOSITE		3	4	4
MB	GNEISS	GRANITE		4	4	4
Mbt	GNEISS	QUARTZITE	SCHIST	3	4	4
Mga	GNEISS			4	4	4
Mgl	GNEISS			4	4	4
Mgo	GNEISS	QUARTZITE	SCHIST	3	4	4
Mh	GNEISS	METAMORPHIC		4	4	4
Mho	GNEISS			4	4	4
Mli	GNEISS			4	4	4
Mva	GNEISS			4	4	4
Nmp	GNEISS	GRANULITE		4	4	4
Nng	GNEISS			4	4	4
Rbu	GNEISS			4	4	4
Rdr	GNEISS			4	4	4
Rho	GNEISS	MIGMATITE		4	4	4
Zb	GNEISS			4	4	4
ZC	GNEISS			4	4	4
ZD	GNEISS			4	4	4
Zgo	GNEISS	GRANITE		4	4	4
Zhh	GNEISS	MIGMATITE	GRANODIORITE	4	4	4
Zma	GNEISS	QUARTZITE	PELITE	4	4	4
Zs	GNEISS			4	4	4
Mbi	GRANITE			4	4	4
Mbk	GRANITE			4	4	4
Mee	GRANITE	GNEISS		4	4	4
Mje	GRANITE			4	4	4
Mke	GRANITE			4	4	4

Mle	GRANITE			4	4	4
Mnr	GRANITE			4	4	4
Mpa	GRANITE			4	4	4
Msc	GRANITE			4	4	4
Msp	GRANITE	GNEISS		4	4	4
Mup	GRANITE			4	4	4
Mwt	GRANITE			4	4	4
N-Cma	GRANITE			4	4	4
N-Cmak	GRANITE			4	4	4
Nd	GRANITE			4	4	4
Nl	GRANITE			4	4	4
RB	GRANITE			4	4	4
Rga	GRANITE	QUARTZ PORPHYRY		4	4	4
Rha	GRANITE			4	4	4
Rsa	GRANITE			4	4	4
Rsk	GRANITE			4	4	4
R-VA	GRANITE			4	4	4
R-Vb	GRANITE			4	4	4
R-Vh	GRANITE			4	4	4
R-Vl	GRANITE			4	4	4
R-Vm	GRANITE			4	4	4
R-Vma	GRANITE	CHARNOCKITE		4	4	4
R-Vmh	GRANITE			4	4	4
R-Vmo	GRANITE			4	4	4
R-Vms	GRANITE			4	4	4
R-Vmt	GRANITE			4	4	4
R-Vp	GRANITE			4	4	4
R-Vs	GRANITE			4	4	4
R-Vsa	GRANITE			4	4	4
R-Vsh	GRANITE			4	4	4
R-Vt	GRANITE			4	4	4
R-Vu	GRANITE			4	4	4
VA	GRANITE			4	4	4
VB	GRANITE			4	4	4
Vme	GRANITE			4	4	4
Vmpg	GRANITE			4	4	4
Vra	GRANITE			4	4	4
ZA	GRANITE	GNEISS		4	4	4
ZB	GRANITE			4	4	4
Zka	GRANITE			4	4	4
Zne	GRANITE	MIGMATITE		4	4	4
Mau	GRANITOID			4	4	4
Mna	GRANITOID			4	4	4
Mcn	GRANODIORITE	GRANITE	QUARTZ MONZONITE	4	4	4
Mda	GRANODIORITE			4	4	4
Mvi	GRANODIORITE	QUARTZ MONZONITE		4	4	4
VC	GRANODIORITE			4	4	4
Zda	GRANODIORITE			4	4	4
Zh	GRANODIORITE			4	4	4
Jt	GRANOPHYRE			4	4	4
Nbi	GREENSTONE			3	3	3
Zn	GREENSTONE	AMPHIBOLITE	GRANULITE	3	3	3
Mlo	HARZBURGITE	NORITE	GABBRO	3	3	3
Vho	HORNFELS	QUARTZITE	LIMESTONE	5	3	3
Vn	HORNFELS	ARENITE	ARENITE	4	3	3
Vve	HORNFELS			4	3	3
R-Vhe	IGNIMBRITE	TUFF	PYROCLASTIC	4	3	3
Va	IRON FORMATION			5	5	5
Vla	IRON FORMATION			5	5	5
Vp	IRON FORMATION	SHALE		4	4	4
Mgr	KINZIGITE			3	3	3
Mto	KINZIGITE			3	3	3
Md	LAVA	TUFF	METAMORPHIC	3	4	4
Mk	LAVA	TUFF	CONGLOMERATE	3	4	4
Mku	LAVA	PYROCLASTIC	CARBONATITE	3	4	4
Mo	LAVA	GNEISS		3	4	4
Mpl	LAVA	TUFF		3	4	4
Vde	LAVA	TUFF	SCHIST	3	4	4
Zg	LAVA	SCHIST		3	4	4
Zgi	LAVA	SCHIST		3	4	4
Zl	LAVA			3	4	4
Zmu	LAVA			3	4	4
Zo	LAVA	PYROCLASTIC		3	4	4
Zp	LAVA	QUARTZITE	CONGLOMERATE	3	4	4

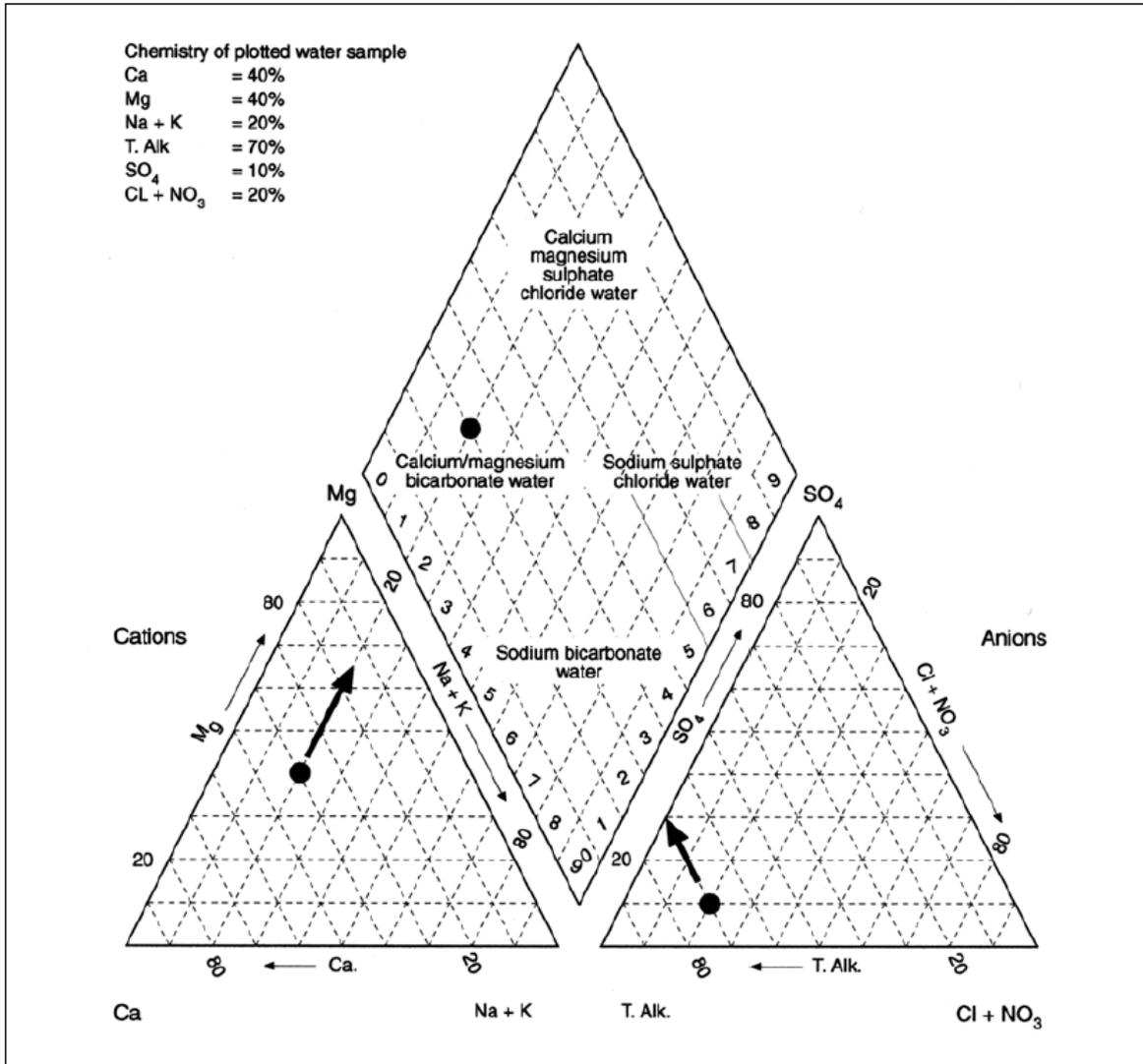
Zr	LAVA	TUFF		3	4	4
Ki	LIMESTONE			10	10	10
Kmz	LIMESTONE	CLAY		9	9	9
Nga	LIMESTONE	PHYLLITE	RUDITE	10	10	10
Nsc	LIMESTONE	SHALE		9	9	9
T-Qa	LIMESTONE	CLAY	CONGLOMERATE	9	9	9
T-Qb	LIMESTONE	ARENITE	CONGLOMERATE	9	10	10
Mm	LUTACEOUS ARENITE	QUARTZITE	CONGLOMERATE	7	7	7
Nbr	LUTACEOUS ARENITE			7	7	7
Nmo	LUTACEOUS ARENITE	SHALE	LIMESTONE	7	7	7
Rt	LUTACEOUS ARENITE	CONGLOMERATE		7	7	7
Zf	LUTACEOUS ARENITE	VOLCANIC ROCKS		7	7	7
Nmz	MARBLE	DOLOMITE	GRANULITE	7	7	7
Zgu	MARBLE	CALC-SILICATE ROCK	GNEISS	7	7	7
Ml	METAMORPHIC			3	4	4
Z	MIGMATITE	GNEISS	ULTRAMAFIC ROCKS	3	4	4
J-K	MUDSTONE	ARENITE	CONGLOMERATE	2	5	5
Pa	MUDSTONE	ARENITE		2	5	5
Pem	MUDSTONE	SHALE	ARENITE	2	4	4
P-TRi	MUDSTONE	ARENITE		2	5	5
TRb	MUDSTONE			2	5	5
TRe	MUDSTONE	ARENITE		2	5	5
TRny	MUDSTONE	ARENITE		2	5	5
TRt	MUDSTONE	ARENITE		2	5	5
Vk	MUDSTONE	IRON FORMATION		2	5	5
MA	NORITE	EPIDIORITE		3	4	4
Vdr	NORITE	PYROXENITE	ANORTHOSITE	3	4	4
Vsl	NORITE	PYROXENITE	ANORTHOSITE	3	4	4
Nhl	OLIVINE GABBRO	GABBRO		3	4	4
Ntr	PERIDOTITE	PYROXENITE	GABBRO	3	3	3
Mz	PHYLLITE	QUARTZITE	GREENSTONE	2	3	3
Nm	PHYLLITE	LUTACEOUS ARENITE	CONGLOMERATE	2	3	3
Np	PHYLLITE	LUTACEOUS ARENITE	LIMESTONE	2	3	3
Npo	PHYLLITE	LUTACEOUS ARENITE	LIMESTONE	2	3	3
My	PYROCLASTIC	CARBONATITE		4	9	9
Jp	PYROCLASTIC BRECCIA	TUFF		4	9	9
Ksu	PYROCLASTIC BRECCIA	TUFF	TRACHYTOID	4	9	9
Mkk	PYROXENITE	SERPENTINITE		3	4	4
V-Mp	PYROXENITE	DUNITE	CARBONATITE	3	4	4
Rmp	QUARTZ MONZONITE			3	4	4
Rkr	QUARTZ PORPHYRY	RHYOLITE	TRACHYTOID	4	5	5
Rm	QUARTZ PORPHYRY			4	5	5
R-Va	QUARTZ PORPHYRY	TUFF		4	5	5
R-Vr	QUARTZ PORPHYRY			4	5	5
Vgl	QUARTZ PORPHYRY	CONGLOMERATE	ARENITE	4	5	5
Mbr	QUARTZITE	SHALE	ARENITE	6	6	6
Mge	QUARTZITE	CALC-SILICATE ROCK		7	7	7
Mkh	QUARTZITE	SCHIST		6	7	7
Mko	QUARTZITE	CONGLOMERATE		6	7	7
Mr	QUARTZITE			7	7	7
Mu	QUARTZITE	SCHIST		6	7	7
Nka	QUARTZITE	PHYLLITE	SCHIST	6	7	7
Nku	QUARTZITE			7	7	7
Ns	QUARTZITE	ARENITE	DOLOMITE	7	7	7
Rg	QUARTZITE	SHALE		5	6	6
Rjo	QUARTZITE	CONGLOMERATE	SHALE	5	7	7
Rmz	QUARTZITE	SHALE	HORNFELS	5	6	6
Vbr	QUARTZITE	CONGLOMERATE	SHALE	5	7	7
Vda	QUARTZITE			7	7	7
Vdw	QUARTZITE	SILTSTONE	CONGLOMERATE	6	7	7
Vg	QUARTZITE	SHALE	LAVA	5	6	6
VI	QUARTZITE	ARENITE		7	7	7
Zmo	QUARTZITE	GNEISS		7	7	7
Zw	QUARTZITE	CONGLOMERATE	SCHIST	6	7	7
Jb	RHYOLITE	SYENITE	BASALT	4	4	4
Jj	RHYOLITE			4	4	4
Vdm	RHYOLITE	PYROCLASTIC		4	4	4
Vkw	RHYOLITE			4	4	4
Vro	RHYOLITE			4	4	4
Vse	RHYOLITE			4	4	4
Vsh	RHYOLITE			4	4	4
Mss	RUDITE	CONGLOMERATE	ARENITE	3	3	3
Qm	SAND			10	9	9
I-Qk	SAND	LIMESTONE		10	9	9

Mb	SCHIST	CONGLOMERATE	QUARTZITE	2	3	3
Me	SCHIST	GNEISS	QUARTZITE	2	3	3
Mg	SCHIST	QUARTZITE	LAVA	2	3	3
Msr	SCHIST	GNEISS	KINZIGITE	2	3	3
Nbe	SCHIST	LUTACEOUS ARENITE	LIMESTONE	3	3	3
Ng	SCHIST	ANDESITE	BASALT	2	3	3
Ngj	SCHIST	LIMESTONE	DOLOMITE	4	3	3
Nho	SCHIST	GNEISS	ARENITE	2	3	3
Nkl	SCHIST			3	3	3
Nmf	SCHIST			3	3	3
No	SCHIST	PHYLLITE	DOLOMITE	4	3	3
Npr	SCHIST	PHYLLITE		3	3	3
Vdg	SCHIST	QUARTZITE	AMPHIBOLITE	4	3	3
Nv	SEDIMENTARY			7	9	9
Q	SEDIMENTARY	SAND	CALCRETE	7	9	9
Nml	SERPENTINITE	GABBRO		2	3	3
Nsi	SERPENTINITE			2	3	3
Db	SHALE			1	4	4
Dbi	SHALE	SILTSTONE	ARENITE	1	4	4
Dc	SHALE	ARENITE		2	4	4
Dl	SHALE	ARENITE	DIAMICTITE	2	4	4
Dt	SHALE	SILTSTONE	ARENITE	2	4	4
Nkn	SHALE	SILTSTONE	ARENITE	2	4	4
Nt	SHALE	LUTACEOUS ARENITE	QUARTZITE	2	4	4
Pe	SHALE			1	4	4
Pf	SHALE			1	4	4
Pk	SHALE			1	4	4
Pp	SHALE			1	4	4
Ppr	SHALE			1	4	4
Ppw	SHALE			1	4	4
Ps	SHALE	ARENITE		2	4	4
Pt	SHALE			1	4	4
P-TR	SHALE	ARENITE	MUDSTONE	2	4	4
P-TRsk	SHALE	MUDSTONE	ARENITE	1	4	4
Pvo	SHALE			1	4	4
Pw	SHALE			1	4	4
Rh	SHALE	QUARTZITE		2	4	4
Rj	SHALE	QUARTZITE	LAVA	2	4	4
Vga	SHALE	QUARTZITE	CONGLOMERATE	2	4	4
Vlo	SHALE	ARENITE	CONGLOMERATE	2	4	4
Vmt	SHALE			1	4	4
Vrt	SHALE	ARENITE		2	4	4
Vry	SHALE	ARENITE		2	4	4
Vs	SHALE			1	4	4
Vsi	SHALE			1	4	4
Vt	SHALE	ARENITE		2	4	4
Vw	SHALE	ARENITE	CONGLOMERATE	2	4	4
Tg	SILCRETE			2	3	3
Vvs	SILICICLASTIC			2	3	3
Kz	SILTSTONE	ARENITE	CONGLOMERATE	2	5	5
Tu	SILTSTONE	LIMESTONE	LIMESTONE	5	5	5
Vv	SILTSTONE	SHALE	ARENITE	2	4	4
Vbl	SLATE	ANDESITE	QUARTZITE	2	3	3
Mpi	SYENITE	ALKALI-FELDSPAR SYENITE	ANDESITE	4	4	4
Mps	SYENITE			4	4	4
Msi	SYENITE	ALKALI-FELDSPAR SYENITE	CARBONATITE	4	4	4
Nke	SYENITE	GRANITE		3	4	4
Nr	SYENITE	GRANITE		3	4	4
Rbo	SYENITE			4	4	4
VD	SYENITE			4	4	4
Vmd	SYENITE			4	4	4
Vsa	SYENITE	GRANITE		3	4	4
C-Pd	TILLITE	ARENITE	MUDSTONE	2	2	2
Nko	TONALITE			3	4	4
Rc	TONALITE			3	4	4
Ra	TUFF			4	5	5
R-Vha	TUFF	ANDESITE	CHERT	4	5	5
Vvg	TUFF	PYROCLASTIC BRECCIA		7	7	7
Z-R	ULTRAMAFIC ROCKS			3	3	3
Mw	VOLCANIC ROCKS			4	4	4
R-Vso	VOLCANIC ROCKS			4	4	4
Vb	VOLCANIC ROCKS	ARENITE		4	4	4
Weight				3	5	3

APPENDIX B - PLOTTING PROCEDURE FOR PIPER DIAGRAM

Piper Diagram Plotting Procedure

Plotting procedures for the Piper diagram: Convert mg/l to meq/l by division  
 (Ca/20, Mg/12, Na/23, K/39, T.Alk./50, SO<sub>4</sub>/48, Cl/35.5, NO<sub>3</sub>/62. Add Na +K and Cl + NO<sub>3</sub>.  
 Calculate percentage of cations and anions. Plot cations by scaling off Ca, then Mg.  
 Plot anions by scaling off T.Alk. then SO<sub>4</sub>. Project cation and anion points to triangle



## APPENDIX C - ROOT MEAN SQUARE ERROR

RMSE provides a measure of the error size, but is sensitive to outliers as it places a lot of weight on large errors and is calculated as follows:

$$RSME = \left[ \frac{1}{n} \sum_{i=1}^n (p_i - o_i)^2 \right]^{1/2} \quad (1C)$$

where

$n$  = Number of observations

$i$  = Index

$p$  = Predicted value

$o$  = Observed value

## APPENDIX D - MEAN AND STANDARD DEVIATION CALCULATIONS

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (1D)$$

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}} \quad (2D)$$

where

- $N$  = Total number of boreholes
- $x$  = Parameter value of interest
- $i$  = Index of parameter list
- $\bar{x}$  = Mean value of  $x$
- $s$  = Standard deviation

## APPENDIX E - PEARSON CORRELATION

The Pearson's correlation coefficient determines the degree to which two variables are linearly related. The results range from -1 (negative linear relationship) to 1 (positive linear relationship), and 0 indicates no relationship present between the predicted data and the observed data. The Pearson correlation coefficient is calculated as follows:

$$PCC = \frac{\sum(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum(O_i - \bar{O})^2 \sum(P_i - \bar{P})^2}} \quad (1E)$$

where

- $PCC$  = *Pearson Correlation Coefficient*
- $i$  = *Total number of observation points*
- $P_i$  = *predicted value*
- $O_i$  = *the observed values for the  $n$  observations*
- $\bar{O}$  = *the mean of the observed values*

## APPENDIX F - INVERSE DISTANCE WEIGHTING

The following is taken directly from Ikechukwu *et al.*, 2017

This method assumes that the value at an unknown location can be approximated as a weighted average of values at points within a certain cut-off distance, or from a given number of the closest points (typically 10 to 30). Weights are usually inversely proportional to a power of distance which, at unsampled locations, leads to an estimator as contained in equation below:

$$F(s) = \sum_{i=1}^n w_i z(s_i) = \frac{\sum_{i=1}^m z(s_i)}{\sum_{j=1}^m \frac{1}{|s - s_j|^p}} \quad (1F)$$

where  $p$  is a parameter (typically = 2). IDW is a method that is easy to use and readily available; it frequently does not produce the local shape implied by data and produces local extrema at the data points. Some modifications have given rise to a class of multivariate blended IDW surfaces and volumes.

The assumption for IDW is that measured points closer to the unknown point are more like it than those that are further away in their values. The weight is given as:

$$\lambda = \frac{\frac{1}{d_i^p}}{\sum_{i=1}^n \frac{1}{d_i^p}} \quad (2F)$$

Where  $d_i$  is the distance between  $x_0$  and  $x_i$ ,  $p$  is a power parameter, and  $n$  is the number of measured points used for the estimation. The main factor affecting the accuracy of IDW is the value of the power parameter. Weights diminish as the distance increases, especially when the value of the power parameter increases, so nearby samples have a heavier weight and have more influence on the estimation, and the resultant spatial interpolation is local.

The choice of power parameter and neighborhood size is arbitrary. The most popular choice of  $p$  is 2 and the resulting method is often called inverse square distance or inverse distance squared (IDS). IDW is referred to as “moving average” when  $p$  is zero, “linear interpolation” when  $p$  is 1 and “weighted moving average” when  $p$  is not equal to 1.

## APPENDIX G - SPLINE INTERPOLATION

The following is taken directly from Ikechukwu *et al.*, 2017

### *Regularized Spline with Tension*

For regularized spline with tension and smoothing, the prediction is given by:

$$z(s_0) = a_1 + \sum_{i=1}^n w_i R(v_i) \quad (1G)$$

where  $a_1$  is a constant and  $R(v_i)$  is the radial basis function given by:

$$R(v_i) = -[E_1(v_i) + \ln(v_i) + C_e] \quad (2G)$$

and

$$v_i = \left[ \frac{\phi h_0}{2} \right]^2 \quad (3G)$$

where  $E_1(v_i)$  is the exponential integral function,  $C_e = 0.577215$  is the Euler constant,  $\phi$  is the generalized tension parameter and  $h_0$  is the separation between the new and interpolation point. The coefficients  $a_1$  and  $w_i$  are obtained by solving the system,

$$\sum_{i=1}^n w_i = 0 \quad (4G)$$

$$a_1 + \sum_{i=1}^n w_i \left[ R(v_i) + \delta_{ij} \frac{\varpi_0}{\varpi_i} \right] = z(s_j); j = 1, \dots, n \quad (5G)$$

where  $\varpi_0/\varpi_i$  are positive weighting factors for a smoothing parameter at each location. The tension parameter  $\phi$  determines the distance over which the given points influence the resulting surface, while the smoothing parameter

controls the vertical deviation of the surface from the sample locations. The use of an appropriate combination of tension and smoothing produces a surface that correctly fits the empirical knowledge about the expected variation.

### *Thin Plate Spline*

Thin plate splines (TPS), previously called Laplacian smoothing splines, for modeling climatic data. A basic solution to the bi-harmonic equation, has the form

$$Z(r) = r^2 \log(r) \quad (6G)$$

where  $r$  is the distance between sample points and unsampled locations. The relation below approximates the surface with minimum bend

$$f(s) = a_1 + a_2x + a_3y + \sum_{i=1}^n w_i z(|s_i - s_0|) \quad (7G)$$

where the terms  $a_1, a_2x, a_3y$  model the linear portion of the surface defining a flat plain that best fits all control points using least squares, the last term models the bending forces due to  $m$  sampled points,  $w_i$  are control points coefficients and  $|s_i - s_0|$  is the separation of sampled point  $s_i$  and location  $s_0$ . The unknowns  $a_1, a_2x, a_3y$  and  $w_i$  are evaluated using the relation:

$$L^{-1} = (w|a_1 a_2 a_3)^T \quad (8G)$$

where

$$L = \begin{bmatrix} K & P \\ P^T & 0 \end{bmatrix} \quad (9G)$$

and  $V$  is a vector of point heights.  $K$  is a matrix of the distance between sampled points and  $P$  is a matrix of the sampled points coordinates.  $L^{-1}$  is obtained by calculating the inverse of  $L$ . Once the unknowns are evaluated, one can compute  $f(s)$  to determine the heights of unknown points. TPS computes a smoothing factor by limiting the Generalized Cross Validation function, GVC, making for a comparatively sturdy model as limiting the GVC improves the accuracy of estimations and is less reliant on the accuracy of the model itself. TPS gives a determination of spatial accuracy.

### *Inverse Multi-Quadratic Spline*

The relation below gives the inverse multi-quadratic spline function

$$f_i(s) = \frac{1}{1 + \|s - s_i\|^2} \quad (10G)$$

where  $\|s - s_i\|$  is the Euclidean distance between control points  $s_i$  and the unknown point  $s$ . The surface is modelled by the function

$$z(s) = \sum_{j=1}^n a_j f_j(s) \quad (11G)$$

where the weights  $a_i$  are selected to ensure exact estimations at each data point such that

$$z_i = z(s_i) = \sum_{j=1}^n a_j f_j(s_i), i = 1, \dots, n \quad (12G)$$

and is computed by the relation

$$z = Fa \quad (13G)$$

where  $z$  is replaced by a vector of sampled data values,  $F$  is a square function matrix given by,

$$L = \begin{bmatrix} K & P \\ P^T & 0 \end{bmatrix} \quad (14G)$$

The estimation function generated with these weights is smooth and exact at sampled data points.

Splines have been widely seen as highly suitable for estimation of densely sampled heights and climatic variables. Among its disadvantages, the inability to integrate larger amounts of auxiliary maps in modeling the deterministic part of change as well as the arbitrary selection of the smoothing and tension parameters have been widely criticized. Predictions obtained from splines therefore are largely dependent on decisions like the order of polynomial used and the number of break points taken by the user. Splines may also be modeled not to be exact to avoid the generation of excessively high or low values common with some exact splines. Unlike the IDW methods, the values predicted by RBFs are not constrained to the range of measured values, i.e. predicted values can be above the maximum or below the minimum measured value.

## APPENDIX H - KRIGING INTERPOLATION

This section is adapted from Van Tonder *et al.*, 1996.

### Notations

*Regionalized variables* are variables distributed in space (and/or time). Mathematically, one can state that a regionalized variable is simply a function that describes the value of a characteristic quantity  $z$  at point  $\mathbf{x} = (x,y)$  in space. We denote This quantity  $z$  is denoted as  $z(\mathbf{x})$ .

A *random variable* is, by definition, a variable that can attain different numerical values, subject to a certain probability distribution. The random function associated with a random variable,  $z(\mathbf{x})$ , is conventionally denoted by  $Z(\mathbf{x})$ .

The basic estimation problem can now be defined as obtaining some estimate for the function  $Z(\mathbf{x})$  at a site  $\mathbf{x}_0$ , where no observations on  $Z(\mathbf{x})$  are available.

The *estimator* for a function  $Z(\mathbf{x})$  at a site  $\mathbf{x}_0$  will be denoted as  $Z^*(\mathbf{x}_0)$ .

The *observations* available for a given set of  $n$  regionalized variables or data points are denoted as  $Z(\mathbf{x}_i)$ ,  $i=1,\dots,n$ .

### Semi-Variogram Computation

Because of the nature of Kriging, a semi-variogram, computed from the regionalized variables or data points, is needed to estimate the manner in which the mean values of the phenomena behave over the region, often referred to as the *drift* or *trend* of the regionalized variable. A mathematical function is then fitted to the semi-variogram values, to obtain certain parameters that are needed for interpolation by Kriging and Bayesian Kriging.

A semi-variogram describes the connection between two points at distance  $h$  from each other. It can be estimated by the function  $\gamma(h)$  in the following equation:

$$\gamma(h) = \frac{1}{2n(h)} \sum [z(\mathbf{x}_i) - z(\mathbf{x}_j)]^2 \text{ for all } i,j \leq n$$

such that

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} = h ,$$

where  $d_{ij}$  is the distance between two points  $x_i$  and  $x_j$ ,  $\gamma(h)$  is the semi-variogram value of lag  $h$ ,  $n$  the number of observations and  $n(h)$  is the number of pairs  $(x_i, x_j)$ , such that  $d_{ij} = h$ .

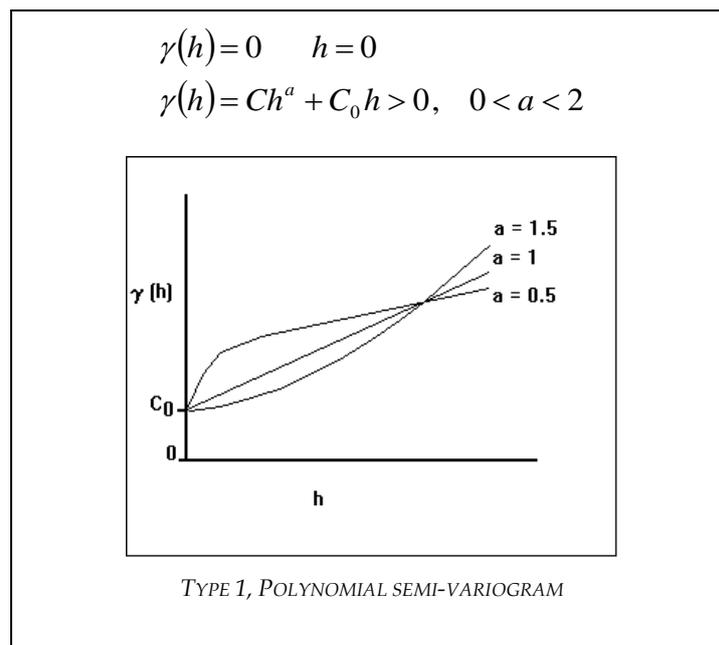
In practice, the approximation for the semi-variogram  $\gamma(h)$  is computed for fixed values of  $h$ , given by some *basic lag distance*,  $a$ , say, thus for  $\gamma(a)$ ,  $\gamma(2a)$ ,  $\gamma(3a)$ , and so forth, where  $\gamma(a)$  is computed for all pairs  $(x_i, x_j)$  such that  $a - 1 < d_{ij} \leq a$ . Further,  $h$  is approximated by the mean distance between all pairs  $(x_i, x_j)$  used in the computation of  $\gamma(a)$ . Thus  $\gamma(d) \equiv \gamma(a)$  for

$$d = \frac{1}{n(h)} \sum^{n(h)} d_{ij}$$

Experience indicates that  $\gamma(h)$  generally tends to increase with  $h$ , until it reaches a maximum value, called the *sill* at some lag  $a$ . This distance is customarily referred to as the *range* of  $\gamma(h)$ . The semi-variogram is then approximated with either one of the six model equations, given in the next section.

### Fitting

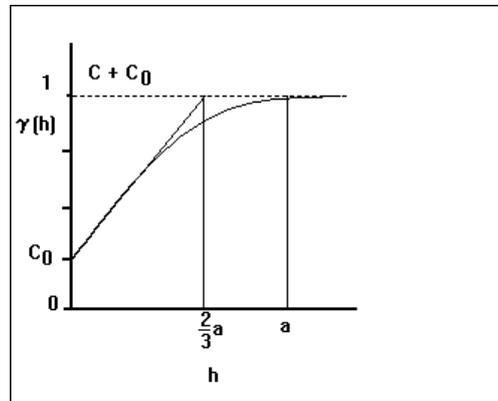
Interpolation by Kriging requires a theoretical semi-variogram and functional terms of an assumed trend in a given neighbourhood of the input data. The neighbourhood is determined by the number of nearest points that are used for interpolation. The most commonly used theoretical semi-variograms are shown in the following figures. The quantity,  $C_0$ , which corresponds to  $\gamma(0)$ , is usually referred to as the *nugget effect*.



$$\gamma(h) = 0 \quad h = 0$$

$$\gamma(h) = C \left[ 1.5 \left( \frac{h}{a} \right) - 0.5 \left( \frac{h}{a} \right)^3 \right] + C_0 \quad 0 < h < a$$

$$\gamma(h) = C + C_0 \quad h > a$$

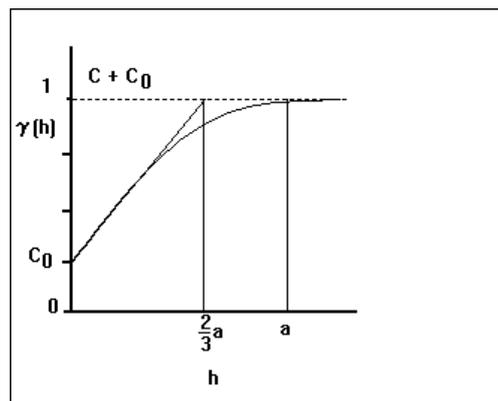


TYPE 2, SPHERICAL SEMI-VARIOGRAM

$$\gamma(h) = 0 \quad h = 0$$

$$\gamma(h) = C \left[ 1.5 \left( \frac{h}{a} \right) - 0.5 \left( \frac{h}{a} \right)^3 \right] + C_0 \quad 0 < h < a$$

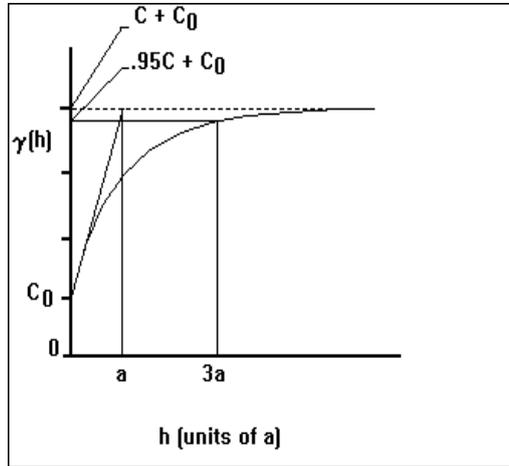
$$\gamma(h) = C + C_0 \quad h > a$$



TYPE 2, SPHERICAL SEMI-VARIOGRAM

$$\gamma(h) = 0 \quad h = 0$$

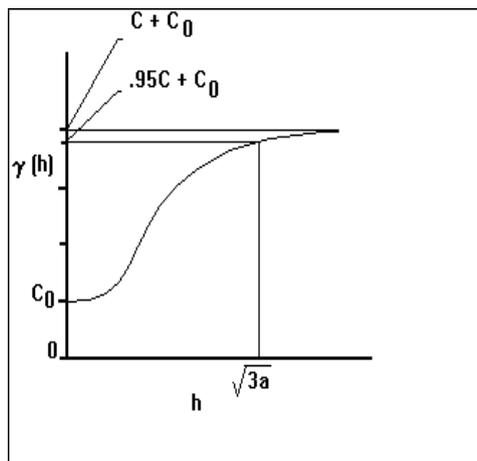
$$\gamma(h) = C \left[ 1 - e^{-\left(\frac{h}{a}\right)} \right] + C_0 \quad h > 0$$



TYPE 3, EXPONENTIAL SEMI-VARIOGRAM

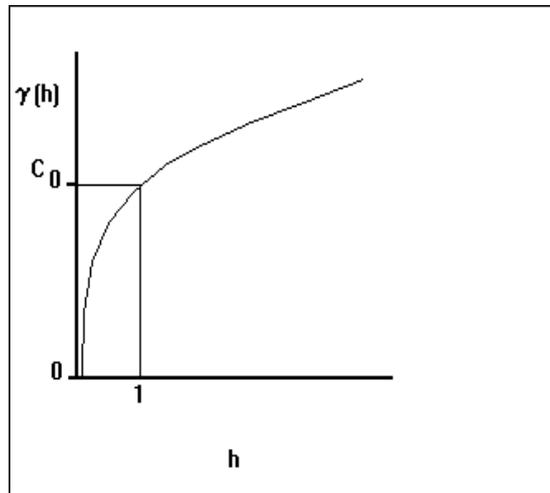
$$\gamma(h) = 0 \quad h = 0$$

$$\gamma(h) = C \left[ 1 - e^{-\left(\frac{h}{a}\right)^2} \right] + C_0 \quad h > 0$$



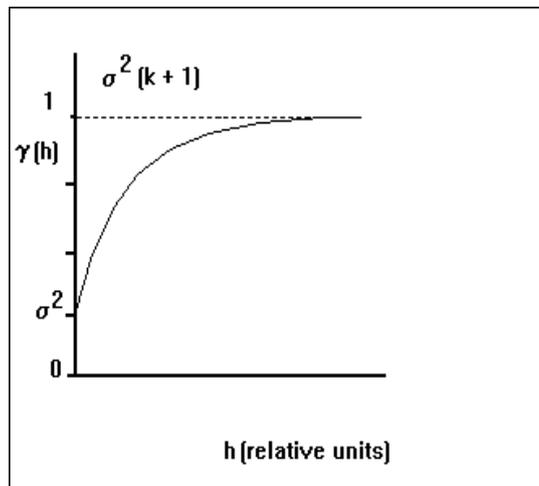
TYPE 4, GAUSSIAN SEMI-VARIOGRAM

$$\begin{aligned} \gamma(h) &= 0 & h &= 0 \\ \gamma(h) &= a \ln(h) + C_0 & h &> 0 \end{aligned}$$



TYPE 5, DE WIJSIAN SEMI-VARIOGRAM

$$\begin{aligned} \gamma(h) &= \sigma^2 & h &= 0 \\ \gamma(h) &= \sigma^2 [1 + k(1 - \rho^h)] & h &> 0 \end{aligned}$$



TYPE 6, KRHO OF DE WAAL SEMI-VARIOGRAM

### Formulation

The most appropriate way to describe the spatial variability of environmental variables, is to present them with random functions. This approach has the advantage that it allows one to describe an environmental variable in statistical terms, through the Theory of Regional Variables. The best-known estimation method, based on this approach, is Ordinary Kriging or *Kriging*, as it is conventionally known.

Interpolation done with Kriging, where the mean value of  $Z(\mathbf{x})$  is unknown, is given by:

$$Z^*(\mathbf{x}_o) = \sum_{i=1}^n w_i Z_i$$

where the weight function  $w_i$  is calculated by solving the system of linear equations

$$\sum_{j=1}^n w_j \gamma_{ij} = \gamma_{io}, \text{ with } \sum_{i=1}^n w_i = 1 \text{ where } \gamma_{ij} = \gamma(d_{ij})$$

The semi-variogram function,  $\gamma(h)$ , as a function of  $h$ , must be known for all values of  $h$ . This condition requires the approximations of the semi-variogram with any of the models in the previous section.

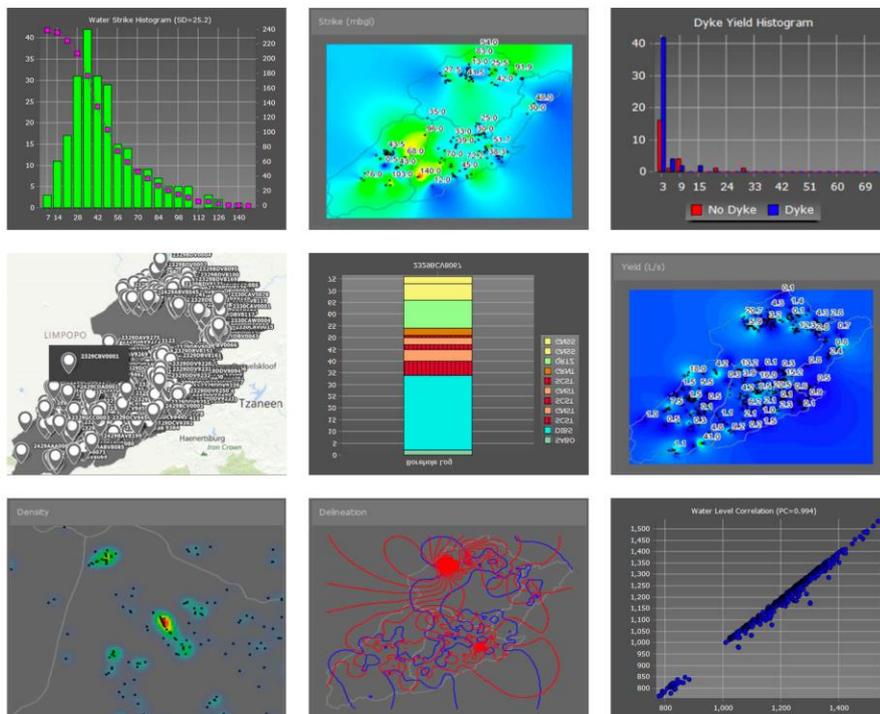
Since Kriging is a linear procedure, difficulties are experienced if the variable to be estimated contains a non-linear trend, or drift as it is called in geostatistical literature. To solve this, *Universal Kriging* was developed, but it is numerically unstable and often singular.

# GSA USER GUIDE

## *Geohydrological Statistical Analysis*

Version 1.0

Get latest update from [www.waterscience.co.za/GSA/](http://www.waterscience.co.za/GSA/)



WRC Project K5/2745

Geo-statistical analysis and sub-delineation of all Vegter Regions

# TABLE OF CONTENTS

---

<i>Computer requirements</i> .....	3
<i>Software Installation</i> .....	3
<i>Graphical User Interface</i> .....	4
<i>Menu Ribbon</i> .....	5
<i>File Menu</i> .....	5
<i>Spatial Tab</i> .....	6
Connect Toolbar .....	6
Data Toolbar .....	6
Map Toolbar.....	7
Area Toolbar .....	9
Spatial Result Window .....	10
<i>Analysis Tab</i> .....	11
Date Toolbar .....	11
Context Toolbar .....	11
Filter Toolbar.....	12
Analysis Result Window .....	12
<i>Delineation Tab</i> .....	15
Level, Strike, Yield and EC Toolbar .....	15
Density Toolbar .....	16
Contour Toolbar .....	16
Delineation Result Window.....	17
<i>Object Tree</i> .....	20
<i>Object Inspector</i> .....	22
<i>SStatus Bar</i> .....	23
<i>Trouble Shooting</i> .....	24

# COMPUTER REQUIREMENTS

Component	Requirement
Operating System	Windows 10 64bit (will also run on some older 32bit versions of Windows)
Disk Space	500Mb base installation plus additional space required for databases
Memory (RAM)	16Gb (will run on some machines with lower memory, but at the cost of reduced performance)

## SOFTWARE INSTALLATION

The software deployment is handled through the use of a self-extracting archive that will prompt the user where it should be deployed. No administrator rights are required for the installation.

### *Installation steps:*

1. Download latest installation file from [www.waterscience.co.za/GSA/](http://www.waterscience.co.za/GSA/)
2. Double click on **GSASetup.exe** and select the desired installation folder. The default folder of the installer is **C:\GSA**

Note: The installer does not create a shortcut to the application. To create a shortcut on your desktop, browse to the installation folder via Windows Explorer and right click on **GSA.exe** and select **Send to > Desktop (create shortcut)**.

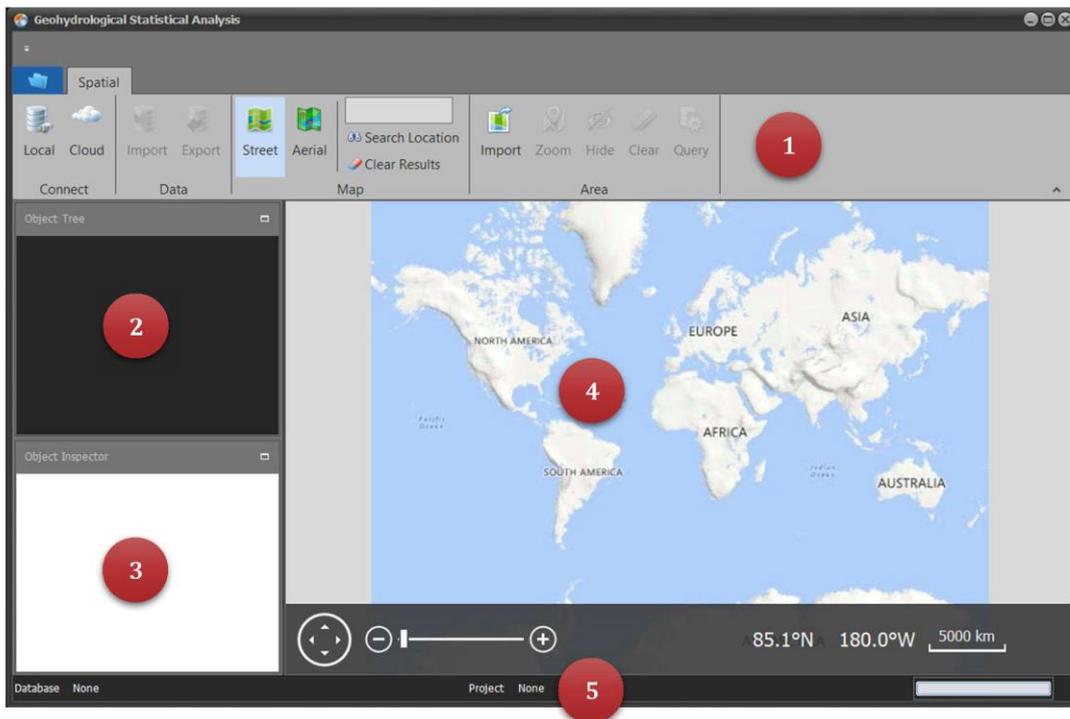
### *Uninstallation steps:*

1. Browse to the folder where the installation took place and simply delete the folder. No system files are deployed during installation and no registry entries are made during the use of the program, so by deleting the folder all elements of the GSA software are removed.

# GRAPHICAL USER INTERFACE

The main graphical user interface (GUI) is shown in the image below and consists of the following components:

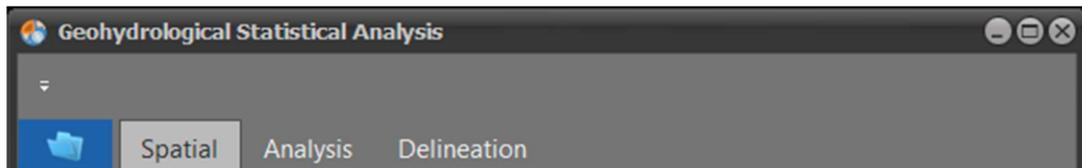
1. A ribbon which consists of multiple tabs, each of which consist of toolbars that provides access to functionality within the software.
2. An object tree which organize all loaded borehole data in a tree structure.
3. An object inspector which allows access to properties of the selected object.
4. A result window that display the various results. The visible results depend on which ribbon tab is active at the time.
5. A status bar that shows which database is connected, which project file is open and also houses the progress bar that is active during operations.



## MENU RIBBON

---

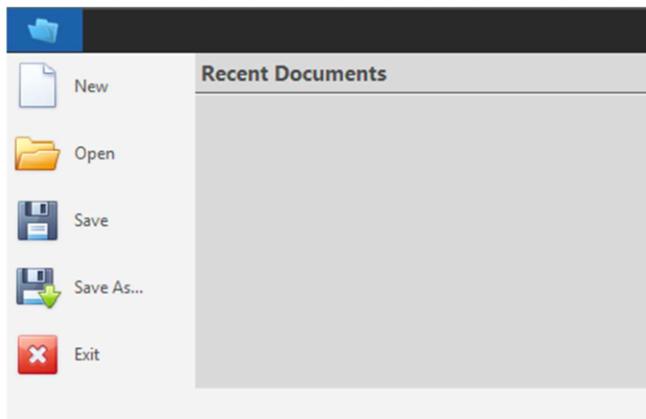
The menu ribbon consists of four tabs namely **File**, **Spatial**, **Analysis** and **Delineation** as shown in the image below. Note the File tab does not show a caption, only an icon. On start-up, only the **File** and **Spatial** tabs are visible and the others will only become visible once a successful connection is established to a database.



## FILE MENU

---

The file menu is shown by clicking on the tab containing the blue file icon as shown in the image as below.



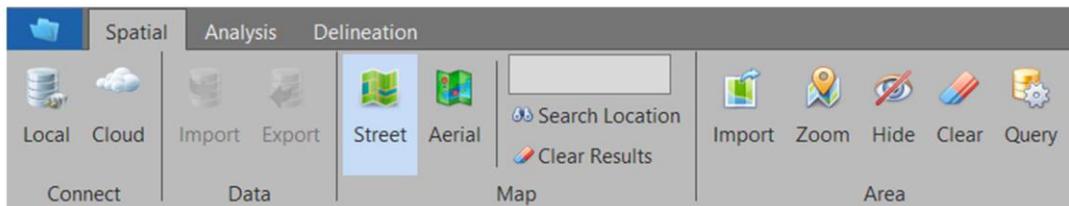
The file menu is used to manage file operations on the project file. The project file is used to save the borehole selections and associated settings, so that a user can continue to work on a specific project without having to query the source database again. A list of recent project files is stored under the Recent Documents section to allow a user to quickly open an of the recent documents.

# SPATIAL TAB

---

The *Spatial* tab house all the toolbars related to spatial functionality and is visible at start-up to allow the user to connect to a database and select the area of interest.

- **Connect** to database source
- **Data** import and export from database
- **Map** functionality used to setup view of online maps
- **Area** selection and query functionality



## Connect Toolbar

---

Offer the user two options on connecting to a source database:

-  **Local** The local option allows the user to select a database that exists on file
-  **Cloud** The cloud option is future functionality that will allow the software to connect to an online database that is sourced via a mobile application

## Data Toolbar

---

Facilitate the import and export of data to and from the connected database:

-  **Import** The import option allows the user to select CSV files, specify the mapping to the database structure and import the data to the selected database.
-  **Export** The export functionality allows the user to export all selected data to CSV format. A CSV file is created for each table in the database with the name of the project file appended.

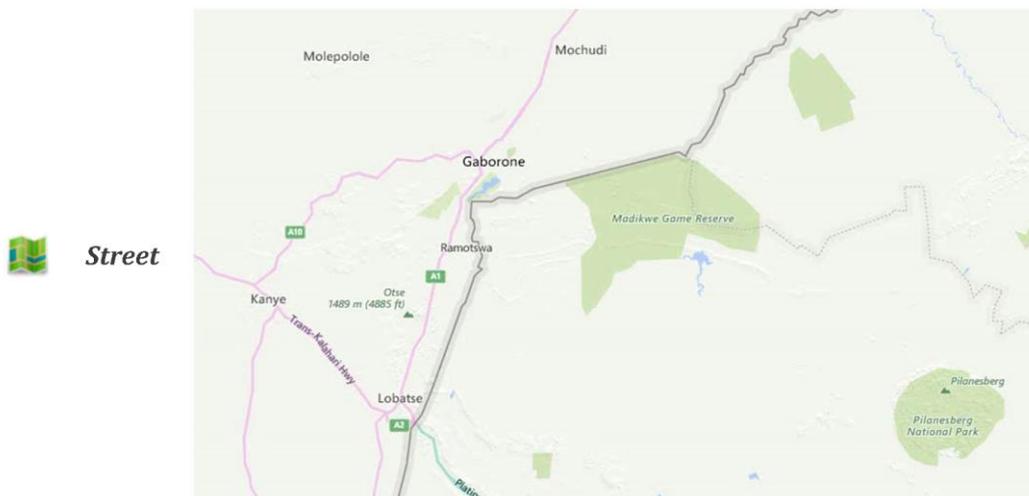
## Map Toolbar

---

Set visible layers of the online maps and provide search functionality:

-  **Street** The street view of the online Bing maps
-  **Aerial** The aerial photo view of the online Bing maps.
-  **Search** The search functionality allows the user to enter a search string e.g. a town name in the text box and search for that feature.
-  **Clear** Once search features are located, they are marked with a map pin. The clear button clears the map from all located search results and removes the map pins.

Note that both the Street and Aerial buttons is check buttons to enable the relevant view. When both buttons are checked simultaneously a hybrid view of the area is obtained as shown in the images below:



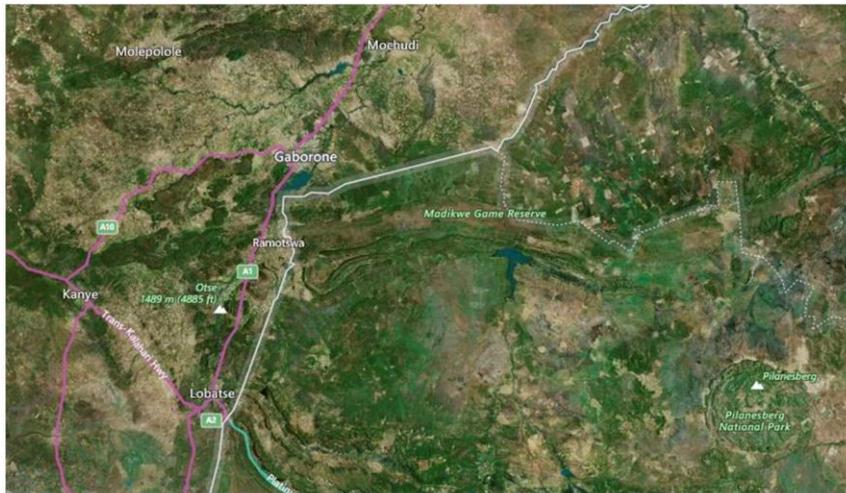


*Aerial*



+

*Hybrid*

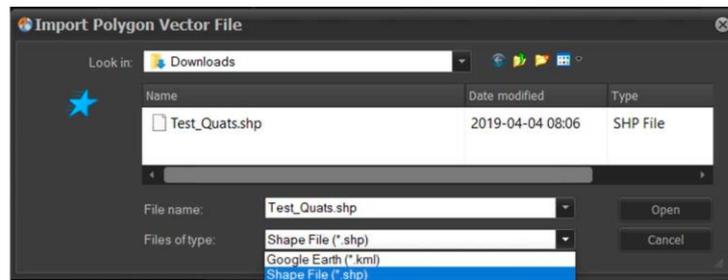


## Area Toolbar

---

Functionality to select area of interest which will be queried for borehole information from the selected database and to be used in the analysis:

 **Import** Import vector data (SHP or KML) that describe the study area.



An example of imported polygons is shown in the map below:



-  **Zoom** Zoom to the extent of the imported polygons.
-  **Hide** Hide the imported polygons.
-  **Show** Show the imported polygons.
-  **Clear** Remove the imported polygons.
-  **Query** Query the database to obtain all boreholes that lie within the selected study area. If no selection is made, the button is disabled.

## Spatial Result Window

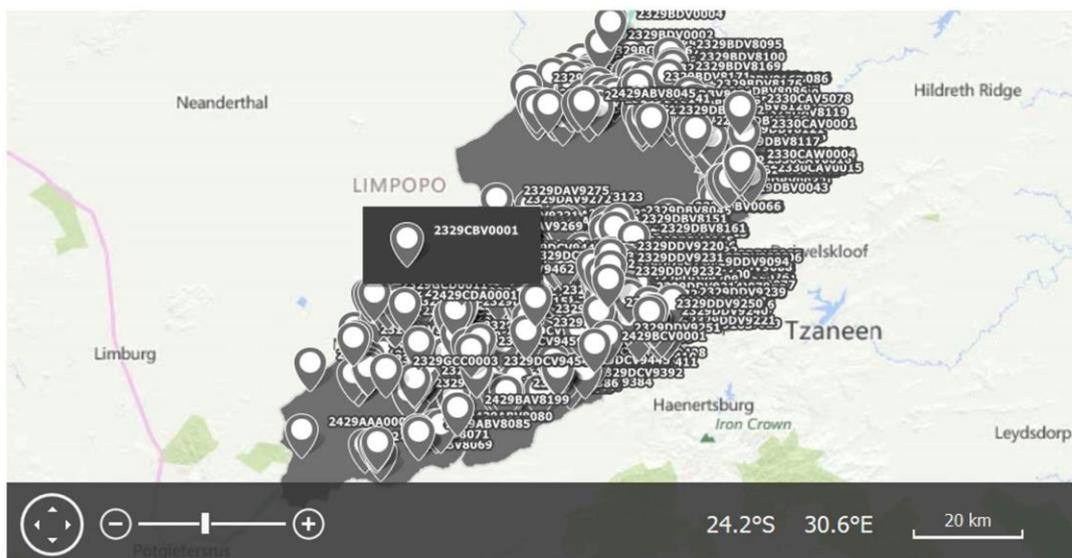
The spatial results are presented within a Bing map that requires an active internet connection to operate. In the absence of an internet connection the software will still operate, only the Bing map will not be able to update and show selected views.

The map view has a toolbar at the bottom which includes a navigation section on the left with pan and zoom functionality and a coordinate readout and scale bar on the right as shown in the image below.

Zoom and pan functionality is also available through the mouse wheel and click-and-drag operations respectively.

The spatial results comprise of the selected study area obtained from the **Import** function and all boreholes within the selected study area obtained from the **Query** operation as shown in the image below.

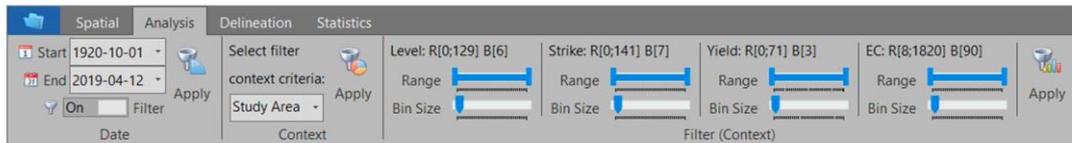
Borehole selection in the **Object Tree** is also highlighted as a selection in the map window and vice-versa.



# ANALYSIS TAB

The **Analysis** tab house all the toolbars related to the analysis functionality and is only visible once an area has been selected and the relevant boreholes queried from the database:

- **Date** selection to enable temporal filter
- **Context** selection to specify tree format
- **Filter** criteria based on data range for specified context



## Date Toolbar

Offer the user the option to specify a start and end date that will be used as temporal filter across all parameters (*Level, Strike, Yield and EC*):

 Start	1920-10-01	<b>Start</b>	Select start date of filter.
 End	2019-04-12	<b>End</b>	Select end date of filter.
<input type="checkbox"/>	On	<b>Filter</b>	Enable the filter.
		<b>Apply</b>	Apply the selected filter settings to the dataset.

## Context Toolbar

Offer the user the option to specify a start and end date that will be used as temporal filter across all parameters (*Level, Strike, Yield and EC*):

 Study Area	<b>Criteria</b>	Select the criteria according to which the tree is build.
	<b>Apply</b>	Apply the selected filter settings to the dataset.

## Filter Toolbar

---

Allow the user to filter the dataset based on minimum and maximum values for each of the parameters considered (*Level*, *Strike*, *Yield* and *EC*). Since all the sub-components are the same for the parameters considered, only the *Level* parameter section is presented here for illustration purposes.

Level: R[0;129] B[6]

### *Info*

The numerical values for both the selected range and bin size are presented in the given label. The range is presented in square brackets preceded by R, **R[x1;x2]** and the bin size is also presented in square brackets but preceded by B, **B[x]**.



### *Range*

Select the range of values to be considered for the specific dataset in question. The range bar values, range from zero to the maximum value detected in the selected dataset.



### *Bin Size*

Select the bin size used in the histogram plot.



### *Apply*

Apply the selected filter settings to the dataset.

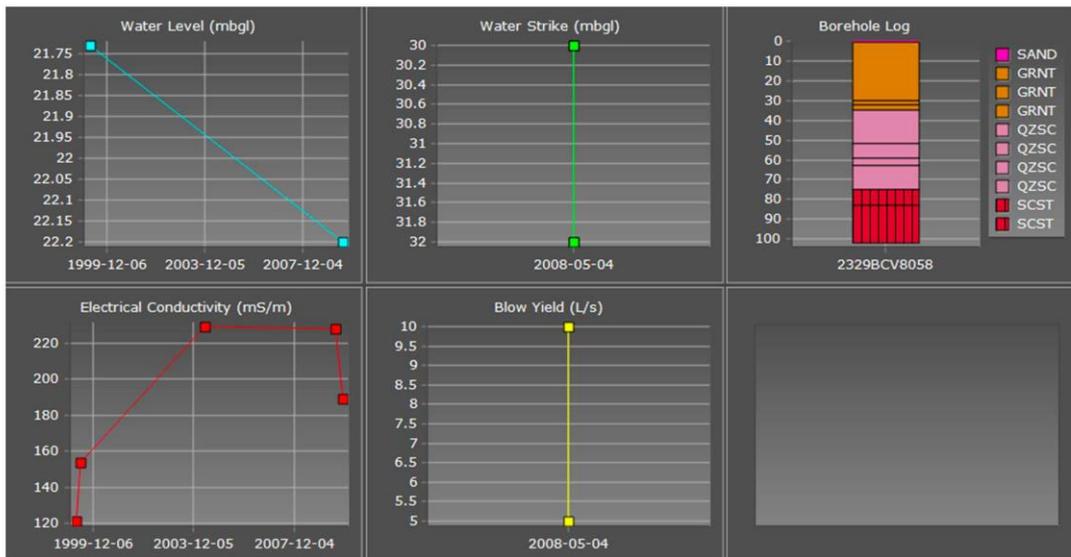
## Analysis Result Window

---

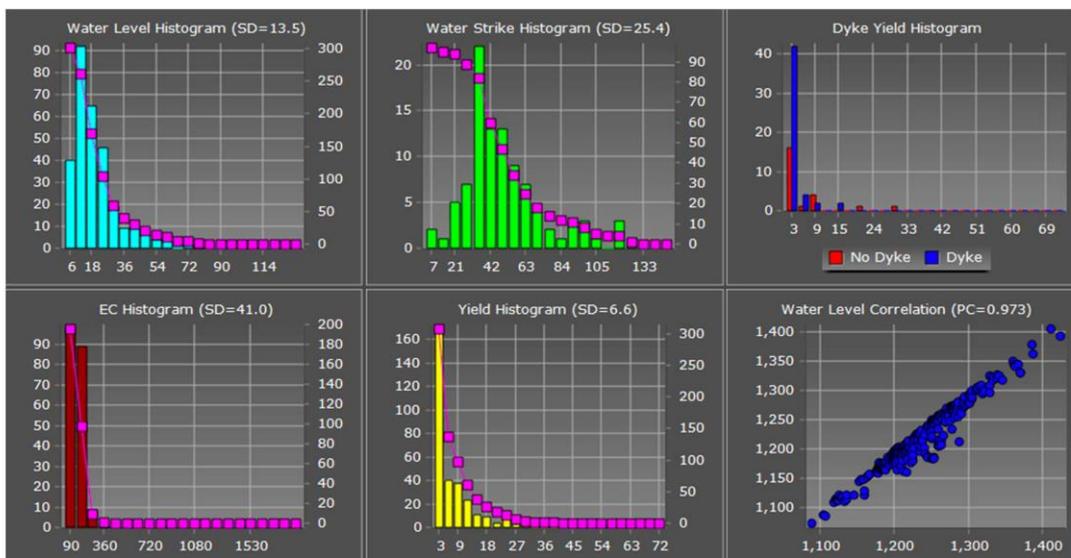
The analysis results are presented as a series of charts comprising a chart dash board.

The selected node in the **Object Tree** determines which results are displayed. If the tree node is a borehole node then the water level, water strike, blow yield, EC and borehole log information is displayed where available as shown in the image below.

Note that axis labels are not presented to optimize the use of the desktop space due to the amount of information displayed on the dash board. In each instance the units of the graph are displayed in brackets after the title where applicable.



If the selected node in the **Object Tree** is a context node then the associated histograms for the group of boreholes within the selected context is displayed as shown in the image below. The **SD** in the graph titles refer to Standard Deviation and **PC** refers Pearson Correlation. All histograms are expressed in counts on the primary vertical axis. The purple line is read from the secondary axis and indicates the total number of boreholes passing through the particular parameter value. The water level correlation is expressed in *mamsl* with elevation on the x-axis and level on the y-axis.



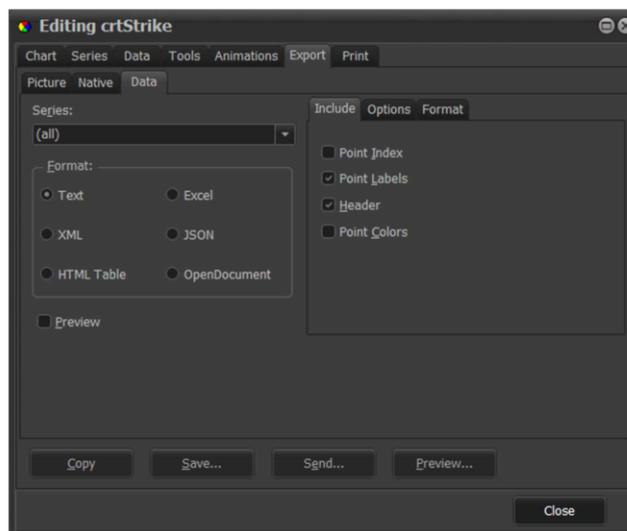
## Chart Functionality

- Pan**
1. Click and hold down the **right mouse button**.
  2. Move the mouse to pan the chart
  3. Let go of the mouse button to exit pan mode.

- Zoom**
1. Click and hold down the **left mouse button**.
  2. Move the mouse from the selected point in (1) and move downward to the right to draw the selection rectangle.
  3. Let go of the mouse button and the chart will zoom to the selected rectangle.

- Full Extent**
1. Click and hold down the **left mouse button**.
  2. Move the mouse from the selected point in (1) and move upward to the left to draw the selection rectangle.
  3. Let go of the mouse button and the chart will zoom out to the full extent.

**Export** To export the chart data, double click on the chart of interest and a chart editor dialog will open as shown below:

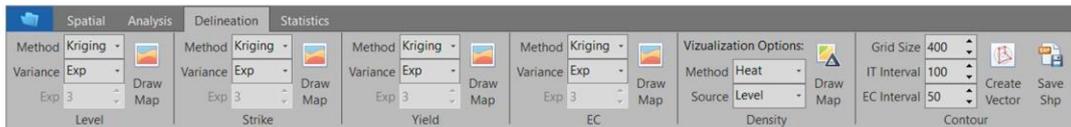


Various other settings are also available through the chart editor, but is only valid while no new selection is made in the **Object Tree**.

# DELINEATION TAB

The **Delineation** tab house all the toolbars related to the delineation functionality and is only visible once an area has been selected and the relevant boreholes queried from the database:

- **Level, Strike, Yield, EC** selection to enable temporal filter
- **Density** selection to specify tree format
- **Contour** criteria based on data range for specified context



## Level, Strike, Yield and EC Toolbar

Since the toolbars for the four parameters considered are identical, only one set will be presented for the purpose of illustration:

- |   |                 |  |
|---|-----------------|--|
|  | <b>Method</b>   | Select the interpolation method from the dropdown list to use. The options are <i>Kriging</i> , <i>Spline</i> and <i>Inverse Distance Weighted (IDW)</i> .   |
|  | <b>Variance</b> | When <i>Kriging</i> is selected as interpolation method, a semi-variance model should be selected from the dropdown list. The options are <i>Power Law</i> , <i>Exponential</i> , <i>Gaussian</i> , <i>Spherical</i> , <i>Circular</i> and <i>Linear</i> . |
|  | <b>Exponent</b> | When the IDW interpolation method is selected an exponent value should be specified (generally a value of 3 yields good results).  |
|  | <b>Draw Map</b> | Interpolate the data based on the selected interpolation settings.   |

## Density Toolbar

---

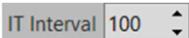
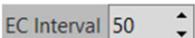
Since the toolbars for the four parameters considered are identical, only one set will be presented for the purpose of illustration:

	<b>Method</b>	Select the density method from the dropdown list to use. The options are <i>Heat Map</i> and <i>Concentration Map</i> .
	<b>Source</b>	Select between <i>Level</i> , <i>Strike</i> , <i>Yield</i> and <i>EC</i> .
	<b>Draw Map</b>	Render map based on the selected density method and source.

## Contour Toolbar

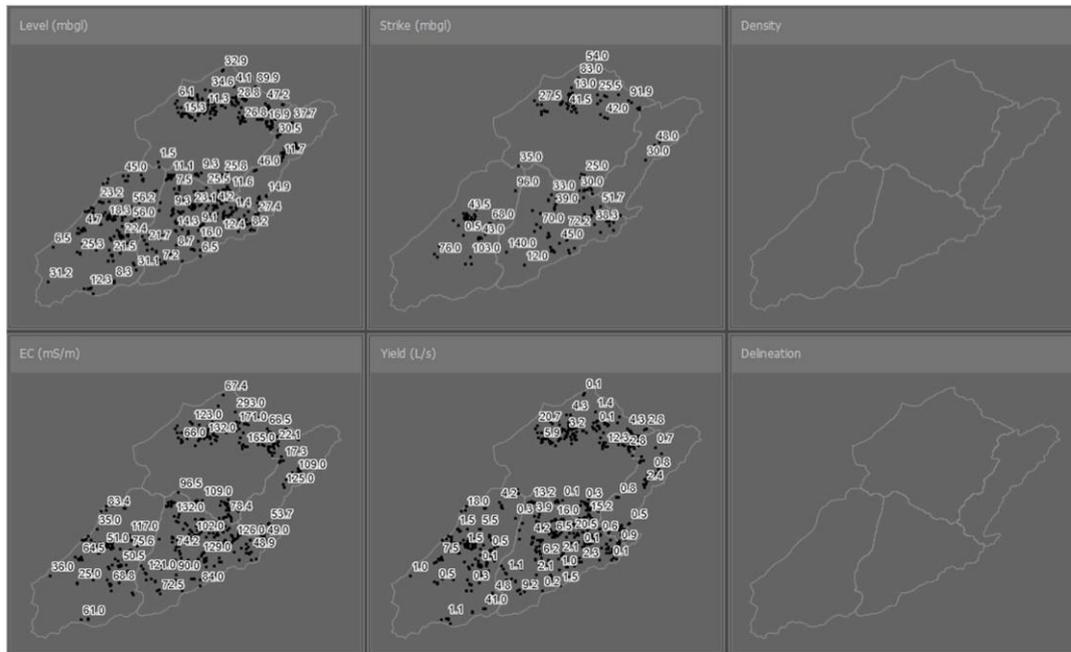
---

Since the toolbars for the four parameters considered are identical, only one set will be presented for the purpose of illustration:

	<b>Grid Size</b>	Select the grid size to be used for all interpolation operations. Please note that the smaller the grid size is chosen the higher resolution maps will be created, but the processing times will be longer.
	<b>IT Interval</b>	Select the <i>Inferred Transmissivity</i> contour interval to be used to produce the delineation shape file.
	<b>EC Interval</b>	Select the <i>Electrical Conductivity</i> contour interval to be used to produce the EC shape file.
	<b>Create Vector</b>	Create the delineation of the study area based on specified methodology ( <i>Inferred Transmissivity</i> ) and also create EC contour map as water quality indicator.
	<b>Save Shp</b>	Once the delineation has been completed and the user can save all delineations in shape file format to be used in a GIS system.

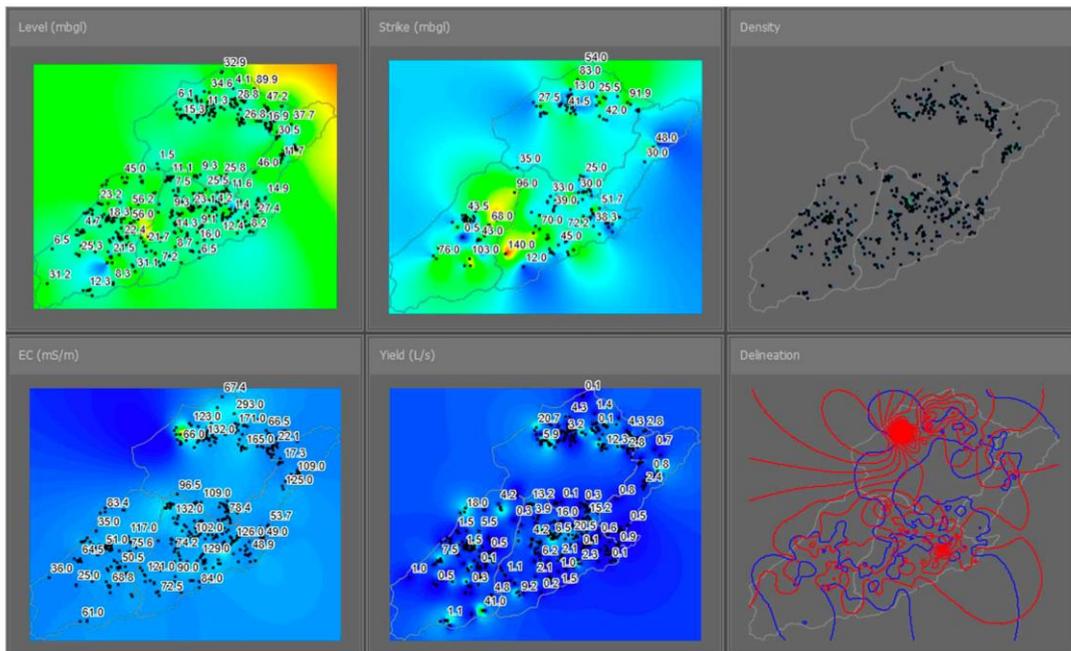
## Delineation Result Window

Once a context node is selected from the **Object Tree**, the borehole distribution across the four parameters (*Level*, *Strike*, *Yield* and *EC*) is presented in a map view as shown in the image below:



Once the user has selected on the interpolation settings for each of the parameters, the interpolation maps is generated as shown in the image below.

The delineation map shown in the bottom right in the image below is only generated once the **Create Vector** button is pressed. The blue contours represent the delineated area and the red contours represent the EC contours.



The following zoom functionality is available:

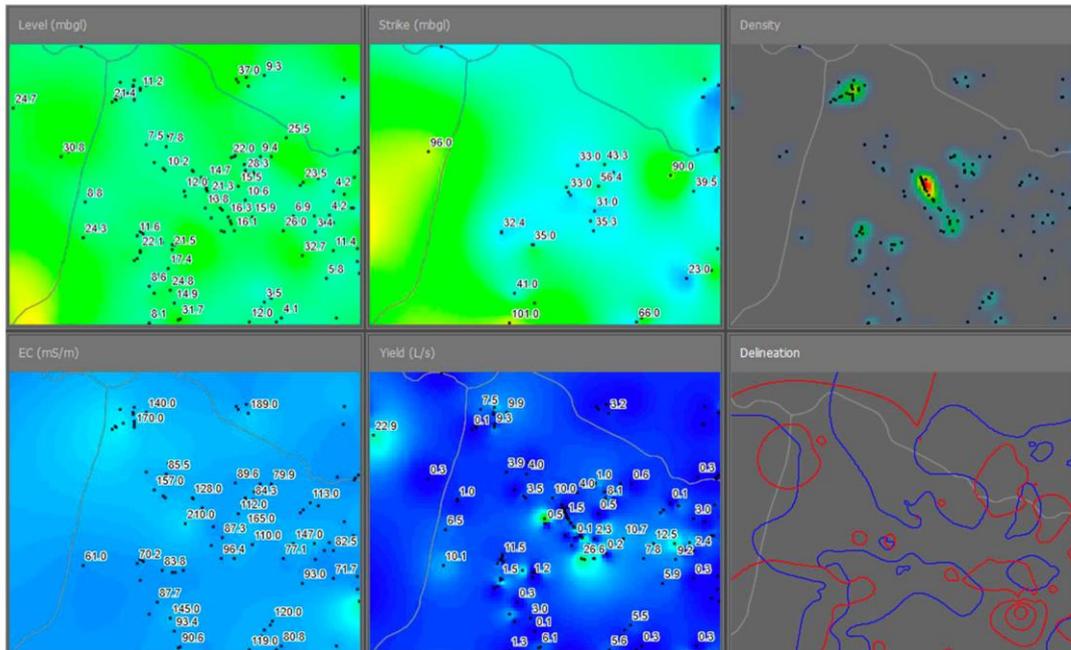
**Zoom**

1. Click and hold down the **left mouse button**.
2. Move the mouse from the selected point in (1) and move downward to the right to draw the selection rectangle.
3. Let go of the mouse button and the chart will zoom to the selected rectangle.

**Full Extent**

1. Click and hold down the **left mouse button**.
2. Move the mouse from the selected point in (1) and move upward to the left to draw the selection rectangle.
3. Let go of the mouse button and the chart will zoom out to the full extent.

All the maps are linked with each other with respect to zoom extent and if one map is zoomed, the others follow suit as shown in the image below:



# OBJECT TREE

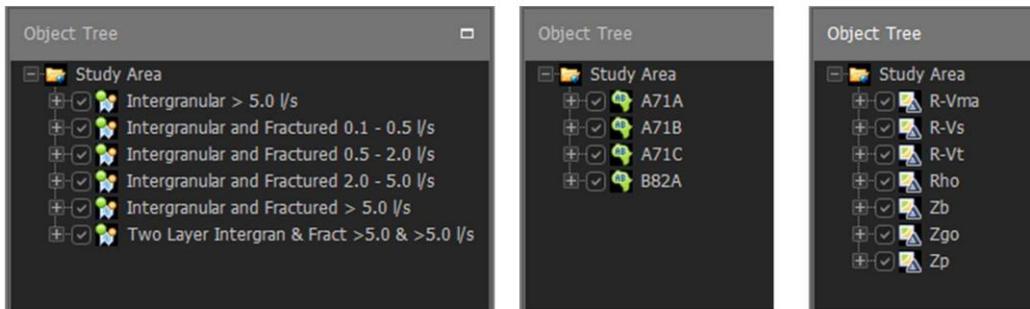
---

The object tree is a representation of all the boreholes that lie within the study area. The context determines the grouping of boreholes into different classes so that the user may analyse the dataset in these different contexts.

Currently the software makes provision to analyse the dataset in terms of the following contexts:

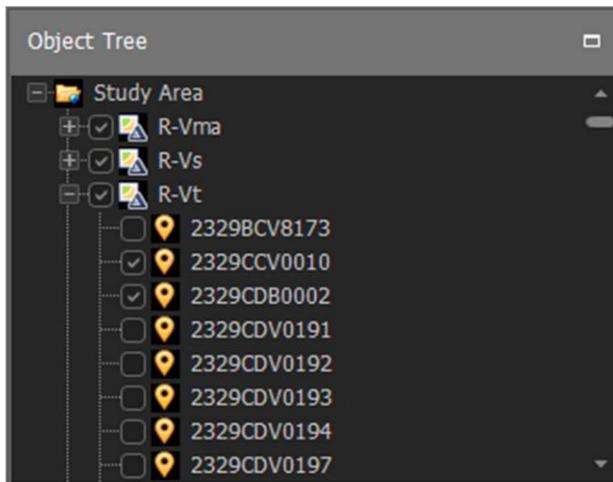
- *Study Area (default)*
- *Quaternary*
- *Geology*
- *Yield (Groundwater Occurrence)*
- *Aquifer Type*
- *Borehole Depth*
- *Borehole Elevation*
- *Aquifer Vulnerability (DRASTIC)*

An example of a study area built in terms of the *Yield (Groundwater Occurrence)*, *Quaternary Catchments* and *Geological Units* is shown in the images below:



The check box in front of the context nodes determine if the specific context is to be used in an overall statistical calculation.

The child nodes of a context tree node are all the boreholes belonging in that particular context as shown in the image below:



Each borehole has a checkbox associated with it and all boreholes that do not own any data with respect of the parameters analysed, will automatically be deselected.

In addition to the automatic system selection of boreholes, the user can also deselect unwanted boreholes by simply removing the check mark in front of the borehole node. Only boreholes with a checked status is displayed in the **Spatial Results Window**.

Only borehole nodes with which is checked for selection is used in the statistical analysis and delineation process. This allows the user to eliminate outliers on an individual basis rather than modifying the range filters, that could possibly eliminate wanted boreholes.

By selecting a node in the tree, the user gain access to the associated **Results Window** and the underlying information of the node as discussed in the **Object Inspector** section.

# OBJECT INSPECTOR

The object inspector presents the node properties to the user in a grid form as shown in the images below. As soon as a user click on a node in the **Object Tree**, the information for the selected node is displayed in the **Object Inspector**.

The borehole node provide site specific information with parameter averages as opposed the context node that provides parameter statistics related to the selected context.

## Borehole Node

Object Inspector	
▲ Site Info	
Name	2329CDV0044
Latitude	-23.897222
Longitude	29.453861
Elevation (mamsl)	0.0
Depth (mbgl)	70.1
Collar (m)	0.26
Quaternary	A71A
Geology	Rho
Occurence	Intergranular and Fract...
SRTM90 (mamsl)	1275.7
DRASTIC	92.0
Aquifer Indicator	4
▲ Dyke/Fault Indicators	
Dyke Lithology	2.00
Fault Proximity	-1.00
▲ Calculated Parameters	
Water Level Trend	0.00
Average Water Level	3.06
Average Strike	33.05
Average Yield (l/s)	0.25
Average EC (mS/m)	63.90

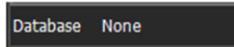
## Context Node

Object Inspector	
▲ Context Info	
Name	Rho
▲ Water Level Stats	
Count	0
Minimum	0.8
Avgerage	17.4
Maximum	75.3
Std.Dev.	13.5
▲ Water Strike Stats	
Count	0
Minimum	0.5
Avgerage	47.3
Maximum	126.0
Std.Dev.	25.4
▲ Yield Stats	
Count	0
Minimum	0.0
Avgerage	5.1
Maximum	37.8
Std.Dev.	6.6
▲ EC Stats	
Count	0
Minimum	23.0
Avgerage	98.7
Maximum	283.0
Std.Dev.	41.0
▲ Depth Stats	
Count	0
Minimum	0.0
Avgerage	71.7
Maximum	202.6
Std.Dev.	35.1

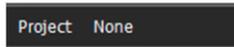
# STATUS BAR

---

The status bar at the bottom of the GUI of the software can be divided into three sections:



**Database** The database section shows the path and name of the current database in use. If the name is *None* it means no database is currently open.



**Project** The project section shows the path and name of the current project in use. If the name is *None*, it implies no project file is open or the current project has not been saved yet.



**Progress Bar** The progress bar is used to indicate the progress of lengthy operations e.g. interpolation. The visual feedback to the user gives an indication of the time required to complete the current operation.

# TROUBLE SHOOTING

---

Should you experience any problem with the software please send an email to the developer explaining the problem:



Rainier Dennis



[rainier.dennis@nwu.ac.za](mailto:rainier.dennis@nwu.ac.za)