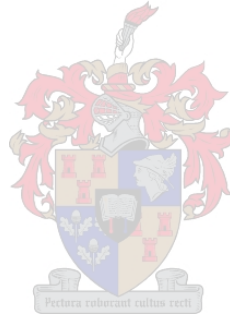# REMOTE SENSING OF SALT-AFFECTED SOILS

by

Zama Eric Mashimbye

Dissertation presented for the degree Doctor of Philosophy in the

Faculty of AgriSciences at Stellenbosch University

Supervisor: Dr Willem Petrus de Clercq

Co-Supervisor: Professor Adriaan van Niekerk

"

""O ctej ''4235

DECLARATION

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes five (5) original papers published in peer reviewed journals. The development and writing of the papers (published and unpublished) were the principal responsibility of myself and for each of the cases where this is not the case a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

Signature…………………………………………… Date…………………………

ABSTRACT

Concrete evidence of dryland salinity was observed in the Berg River catchment in the Western Cape Province of South Africa. Soil salinization is a global land degradation hazard that negatively affects the productivity of soils. Timely and accurate detection of soil salinity is crucial for soil salinity monitoring and mitigation. It would be restrictive in terms of costs to use traditional wet chemistry methods to detect and monitor soil salinity in the entire Berg River catchment. The goal of this study was to investigate less tedious, accurate and cost effective techniques for better monitoring.

Firstly, hyperspectral remote sensing (HRS) techniques that can best predict electrical conductivity (EC) in the soil using individual bands, a unique normalized difference soil salinity index (NDSI), partial least squares regression (PLSR) and bagging PLSR were investigated. Spectral reflectance of dry soil samples was measured using an analytical spectral device FieldSpec spectrometer in a darkroom. Soil salinity predictive models were computed using a training dataset (n = 63). An independent validation dataset (n = 32) was used to validate the models. Also, field-based regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were developed using soil samples (n = 23) collected in the Sandspruit catchment. These soil samples were not ground or sieved and the spectra were measured using the sun as a source of energy to emulate field conditions. Secondly, the value of NIR spectroscopy for the prediction of EC, pH, soluble Ca, Mg, Na, Cl, and SO_4 was evaluated using 49 soil samples. Spectral reflectance of dry soil samples was measured using the Bruker multipurpose analyser spectrometer. "Leave one out" cross validation (LOOCV) was used to calibrate PLSR predictive models for EC, pH, soluble Ca, Mg, Na, Cl, and SO_4. The models were validated using R^2, root mean square error of cross validation (RMSECV), ratio of prediction to deviation (RPD) and the ratio of prediction to interquartile distance (RPIQ). Thirdly, owing to the suitability of land components to map soil properties, the value of digital elevation models (DEMs) to delineate accurate land components was investigated. Land components extracted from the second version of the 30-m advanced spaceborne thermal emission and reflection radiometer global DEM (ASTER GDEM2), the 90-m shuttle radar topography mission DEM (SRTM DEM), two versions of the 5-m Stellenbosch University DEMs (SUDEM L1 and L2) and a 5-m DEM (GEOEYE DEM) derived from GeoEye stereo-images were compared. Land components were

delineated using the slope gradient and aspect derivatives of each DEM. The land components were visually inspected and quantitatively analysed using the slope gradient standard deviation measure and the mean slope gradient local variance ratio for accuracy.

Fourthly, the spatial accuracy of hydrological parameters (streamlines and catchment boundaries) delineated from the 5-m resolution SUDEM (L1 and L2), the 30-m ASTER GDEM2 and the 90-m SRTM was evaluated. Reference catchment boundary and streamlines were generated from the 1.5-m GEOEYE DEM. Catchment boundaries and streamlines were extracted from the DEMs using the Arc Hydro module for ArcGIS. Visual inspection, correctness index, a new Euclidean distance index and figure of merit index were used to validate the results. Finally, the value of terrain attributes to model soil salinity based on the EC of the soil and groundwater was investigated. Soil salinity regression predictive models were developed using CurveExpert software. In addition, stepwise multiple linear regression soil salinity predictive models based on annual evapotranspiration, the aridity index and terrain attributes were developed using Statgraphics software. The models were validated using R^2, standard error and correlation coefficients. The models were also independently validated using groundwater hydro-census data covering the Sandspruit catchment.

This study found that good predictions of soil salinity based on bagging PLSR using first derivative reflectance ($R^2 = 0.85$), PLSR using untransformed reflectance ($R^2 = 0.70$), a unique NDSI ($R^2 = 0.65$) and the untransformed individual band at 2257 nm ($R^2 = 0.60$) predictive models were achieved. Furthermore, it was established that reliable predictions of EC, pH, soluble Ca, Mg, Na, Cl and SO_4 in the field are possible using first derivative reflectance. The R^2 for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 predictive models are 0.85, 0.50, 0.65, 0.84, 0.79, 0.81 and 0.58 respectively. Regarding NIR spectroscopy, validation R^2 for all the PLSR predictive models ranged from 0.62 to 0.87. RPD values were greater than 1.5 for all the models and RMSECV ranged from 0.22 to 0.51. This study affirmed that NIR spectroscopy has the potential to be used as a quick, reliable and less expensive method for evaluating salt-affected soils. As regards hydrological parameters, the study concluded that valuable hydrological parameters can be derived from DEMs. A new Euclidean distance ratio was proved to be a reliable tool to compare raster data sets. Regarding land components, it was concluded that higher resolution DEMs are required for delineating meaningful land components. It seems

probable that land components may improve salinity modelling using hydrological modelling and that they can be integrated with other data sets to map soil salinity more accurately at catchment level. In the case of terrain attributes, the study established that promising soil salinity predictions could be made based on slope, elevation, evapotranspiration and terrain wetness index (TWI). Stepwise multiple linear regressions soil salinity predictive model based on elevation, evapotranspiration and TWI yielded slightly more accurate prediction of soil salinity. Overall, the study showed that it is possible to enhance soil salinity monitoring using HRS, NIR spectroscopy, land components, hydrological parameters and terrain attributes.

UITTREKSEL

Konkrete bewyse van droëland sout is waargeneem in die Bergrivier opvanggebied in die Wes-Kaap van Suid-Afrika. Verbrakking van grond is 'n wêreldwye   probleem wat 'n negatiewe invloed op die produktiwiteit van grond kan hê. Tydige en akkurate herkenning van verandering in grond soutgehalte is 'n noodsaaklike aksie vir voorkoming. Dit sou beperkend wees in terme van koste om konvensionele nat chemiese metodes te gebruik vir die opsporing en monitering daarvan in die hele Bergrivier opvanggebied. Die doel van hierdie studie was om ondersoek in te stel na minder tydsame, akkurate en koste-effektiewe tegnieke vir beter monitering.

Eerstens, is hiperspektrale afstandswaarnemings (HRS) tegnieke wat die beste in staat is elektriese geleidingsvermoë (EG) in die grond te kan voorspel deur gebruik te maak van individuele bande, 'n unieke genormaliseerde grond soutindeks verskil (NDSI), parsiële kleinste kwadratiese regressie (PLSR) en afwyking in PLSR, is ondersoek. Spektrale reflektansie van droë grondmonsters is gemeet deur gebruik te maak van 'n spektrale analitiese toestel: FieldSpec spektrometer in 'n donkerkamer. Voorspellings modelle vir grond soutgehalte is bereken met behulp van 'n toets datastel (n = 63). 'n onafhanklike validasie datastel (n = 32) is gebruik om die modelle te evalueer. Daarbenewens is veld-gebaseerde regressie voorspellings modelle vir EG, pH oplosbare Ca, Mg, Na, Cl and SO_4 ontwikkel deur gebruik te maak van grondmonsters (n = 23) versamel in the Sandpruit opvangsgebied. Hierdie grondmonsters is nie gemaal of gesif nie en die spectra is gemeet deur gebruik te maak van die son as 'n bron van energie om veld toestande na te boots. Tweedens, is die waarde van NIR spektroskopie vir die voorspelling van die EG, pH, oplosbare Ca, Mg, Na, Cl, en SO_4 met behulp van 49 grondmonsters geëvalueer. Spektrale reflektansie van droë grondmonsters is gemeet deur gebruik te maak van die Bruker NIR veeldoelige analiseerder . Kruisvalidering (LOOCV) is gebruik om PLSR voorspellings modelle vir EG, pH, oplosbare Ca, Mg, Na, Cl, en SO_4 te kalibreer. Hierdie modelle is gevalideer: R^2, wortel-gemiddelde-kwadraat fout kruisvalidering (RMSECV), verhouding van voorspellings afwyking (RPD) en die verhouding van die voorspelling se inter-kwartiel afstand (RPIQ). Derdens is land komponente gekarteer vanweë die nut daat van tov grondeienskappe, en die waarde van DEMs is ondersoek om akkurate land komponente af te baken. Land komponente uit die tweede weergawe van die 30 m gevorderde ruimte termiese emissie en refleksie radio

globale DEM (ASTER GDEM2), die 90-m ruimtetuig radar topografie sending DEM (SRTM DEM), twee weergawes van die 5 m Universiteit van Stellenbosch DEMs (SUDEM L1 en L2) en 'n 5 m DEM (GEOEYE DEM) afgelei van GeoEye stereo-beelde, is vergelyk. Land komponente is afgebaken met behulp van helling, gradiënt en aspek afgeleides van elke DEM. Die land komponente is visueel geïnspekteer en kwantitatief ontleed met behulp van die helling gradiënt standaardafwyking te meet en die gemiddelde helling-gradiënt-plaaslike variansie verhouding vir akkuraatheid.

Vierdens, is die ruimtelike akkuraatheid van hidrologiese parameters (stroomlyn en opvanggebied grense) geëvalueer soos afgelei vanaf die 5 m resolusie SUDEM (L1 en L2), die 30 m ASTER GDEM2 en die 90 m SRTM . Die verwysings opvanggebied grens en stroomlyn is gegenereer vanaf die 1,5-m GEOEYE DEM. Opvanggebied grense en stroomlyn uit die DEMs is bepaal deur gebruik te maak van die Arc Hydro module in ArcGIS. Visuele inspeksie, korrektheid indeks, 'n nuwe Euklidiese afstand indeks en die indikasie-van-meriete indeks is gebruik om die resultate te valideer. Laastens is die waarde van die terrein eienskappe om grond southalte te modeleer ondersoek, gebaseer op die EG van die grond en grondwater. Grond soutgehalte regressie voorspellings modelle is ontwikkel met behulp van CurveExpert sagteware. Verder, stapsgewyse meervoudige lineêre regressie grond soutgehalte voorspellings modelle gebaseer op jaarlikse evapotranspirasie, die dorheids indeks en terrein eienskappe is ontwikkel met behulp van Statgraphics sagteware. Die modelle is gevalideer deur gebruik te maak van R^2, standaardfout en korrelasiekoëffisiënte. Die modelle is ook onafhanklik bekragtig deur die gebruik van grondwater hidro-sensus-data wat die Sandspruit opvanggebied insluit.

Hierdie studie het bevind dat 'n goeie voorspelling van grond soutgehalte gebaseer op uitsak PLSR met behulp van eerste orde afgeleide reflektansie ($R^2 = 0,85$), PLSR deur gebruik te maak van ongetransformeerde reflektansie ($R^2 = 0,70$), 'n unieke NDSI ($R^2 = 0,65$) en die ongetransformeerde individuele band op 2257 nm ($R^2 = 0,60$) voorspellings modelle verkry is . Verder is vasgestel dat betroubare voorspellings van die EG, pH, oplosbare Ca, Mg, Na, Cl en SO_4 in die veld moontlik is met behulp van eerste afgeleide reflektansie. Die R^2 van EG, pH, oplosbare Ca, Mg, Na, Cl en SO_4 is 0.85, 0.50, 0.65, 0.84, 0.79, 0.81 en 0.58 onderskeidelik. Ten opsigte van NIR spektroskopie het die validasie van R^2 vir al die PLSR voorspellings modelle

gewissel tussen 0,62-0,87. Die RPD waardes was groter as 1,5 vir al die modelle en RMSECV het gewissel tussen 0,22-0,51. Hierdie studie het bevestig dat NIR spektroskopie die potensiaal het om gebruik te word as 'n vinnige, betroubare en goedkoper metode vir die analise van sout-geaffekteerde gronde. T.o.v. hidrologiese parameters, het die studie tot die gevolgtrekking gekom dat waardevolle hidrologiese parameters afgelei kan word uit DEMs. 'n nuwe Euklidiese afstand verhouding is bevestig as 'n betroubare hulpmiddel om raster datastelle te vergelyk. Ten opsigte van grond komponente, is daar tot die gevolgtrekking gekom dat hoër resolusie DEMs nodig is vir die bepaling van sinvolle land komponente. Dit lyk waarskynlik dat die land komponent soutgehalte modellering hidrologiese modellering verbeter en dat hulle geïntegreer kan word met ander datastelle vir meer akkurate kaarte op opvangsgebied skaal. In die geval van die terrein eienskappe het, die studie vasgestel dat belowende grond soutgehalte voorspellings gemaak kan word gebaseer op helling, elevasie, evapotranspirasie en terrein natheid indeks (TWI). 'n stapsgewyse meervoudige lineêre regressie grond soutgehalte voorspellings model wat gebaseer is op elevasie, evapotranspirasie en TWI het effens meer akkurate voorspellings van die grond soutgehalte gelewer. In geheel gesien, het die studie getoon dat dit moontlik is om grond soutgehalte monitering te verbeter met behulp van HRS, NIR spektroskopie, land komponente, hidrologiese parameters en terrein eienskappe.

ACKNOWLEDGEMENTS

I express my sincere thanks and appreciation to:

Mr. Richard Tswai, Mr. Phila Sibandze (of the ARC-ISCW) and Mr. Leonard Adams (of the Stellenbosch University Soil Science Department) for assistance with data collection and spectral data measurement.

Tarina Vermeulen for assistance with saturated pastes extracts preparation in the laboratory.

Mr. Pieter Norval of Bruker South Africa for technical support during NIR spectroscopic data measurements.

Colleagues Trynos Gumbo, Walter Musakwa and Evance Mwuthunga for support and for taking some time to read some of my work.

Staff of Stellenbosch University Soil Science, and Geography and Environmental Studies departments for support and motivation.

My wife Basani for support and for taking responsibility of our family during my occasional absence from home.

My children Thulani, Kurshell and Mfihlo for their patience and understanding of my occasional absence from home during the course of the study.

My uncle Dr. KZ Nkuna and aunt Carlifornia Nkuna for guidance, support and motivation.

My niece Khombani Ngulele for support and for agreeing to stay with my family when I was away on a one year study leave at Stellenbosch University.

My parents, brothers and sisters for support and encouragement.

ABBREVIATIONS AND ACRONYMS

AGIS   Agricultural Geo-Referenced Information Systems

ALCoM   Automated land component mapper

ARC-ISCW   Agricultural Research Council-Institute for Soil, Climate and Water

ASD   Analytical spectral devices

ASTER GDEM   Aster global digital elevation model

BRC   Berg River catchment

BRDF   Bidirectional reflectance distribution effects

CDF   Cumulative distribution function

CDNG   Chief Directorate National Geo-spatial Information

CGA   Centre for Geographical Analysis

CGIAR-CSI   Consultative Group on International Agricultural Research Consortium for
                      Spatial Information

CONUS   Conterminous United States

C_r Correctness index

DEM   Digital elevation model

DTM   Digital terrain model

DWAF   Department of Water Affairs and Forestry

EC   Electrical conductivity

ED   Euclidean distance

ESP   Estimation of scale parameter

FDR   First derivative reflectance

FMI   Figure of merit index

FT   Fourier-transform

GIS   Geographical information systems

HRS   Hyperspectral remote sensing

LiDAR   light detection and ranging

LSU   Linear spectral unmixing

LTD   Land type database

LV   Local variance

ME	mean error
MAE	mean absolute error
METI	Ministry of Economy, Trade and Industry
MF	Matched filtering
MIR	Mid infrared
MPA	Multipurpose analyser
MRS	Multiresolution image segmentation
MSC	Multiplicative scattering correction
MSGLV	Mean slope gradient local variance
MTMF	Mixture tuned matched filtering
NDSI	Normalized difference salinity index
NDVI	Normalized difference vegetation index
NIR	Near infrared
OBIA	Object-based image analysis
PLSR	Partial least squares regression
RADAR	Radio detection and ranging
RMSE	Root mean squares error
RMSEP	Root mean squares error of prediction
RMSECV	Root mean squares error of cross validation
RPC	Rational polynomial coefficients
RPD	Ratio of prediction to deviation
RPIQ	Ratio of prediction to inter-quartile distance
SD	Standard deviation
SDE	Standard deviation of the error distribution
SDR	Second derivative reflectance
SEP	Standard error of prediction
SGSD	Slope gradient standard deviation
SNV	Vector normalization
SRTM	Shuttle Radar Topographic Mission
SUDEM	Stellenbosch University digital elevation model
SWIR	Shortwave infrared

UV	Ultra violet
VIS	Visible
VNIR	Visible-near infrared reflectance
WCDEM	Western Cape digital elevation model
WRC	Water Research Commission

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION: THE SOIL SALINTY QUESTION

INTRODUCTION

The link between what can be observed from space and the detailed work normally carried out in field surveys for agricultural, environmental and hydrological purposes, are mostly difficult to define. The work that will be presented in this dissertation is focussed on finding and defining the most functional link between the two modes of observation with elaboration on the extrapolation of ground based measurements through using remotely sensed information as co-factors. Our diminishing ability to use ground based measurements, and the further restrictions due to the cost of sampling and chemical analysis, makes relating our results to remotely sensed information very attractive.

As will be demonstrated in this dissertation, a huge number of articles already exist trying to bridge the gap between point based information and remotely sensed pixel based information. To be able to make a comparison between the two sources, one needs to find the neutral ground and this is mostly hazed by the questions of scale, sometimes called support, the quality of the data and the ability of analytical software. The work therefore presented in this dissertation will fall within the latter, and it will be indicated that for two specific uses, namely the adequate mapping of salinity in a landscape and the derivation of hydrologically correct information, how the information generated could vary and how to get the most out of available remotely sensed information.

It is very important to acknowledge that this type of study demands people that are skilled in field observation techniques (Soil Science), geographical information systems (GIS), remote sensing and geostatistical techniques. This study, though done across three departments at

Stellenbosch University, could develop into a standard program as the demand for this type of information is on the increase.

The amount of reasons remotely sensed information is important, is increasing daily while our ability to utilize all the resources is getting less and less as we discover more reasons why and how remotely sensed information is affected. Therefore, the knowledge and skills of defining remotely sensed information for specific uses and generate credibility for the utilisation of these products, is a very important step we need to take forward to also make sure that actions taken by government and environmentalists were not based on wrong information. As a simple example, a catchment boundary is normally determined using digital elevation models, derived from satellite information. If the boundary is delineated oncorrectly, all other hydrological calculations from then on are calculated wrongly.

With the above in mind, the chapters that follow indicate the filling of specific gaps in our quest for generating better information towards the hydrological modelling of the Sandspruit and the role salinity plays in this region.

1.1 BACKGROUND

In the desire to provide food for the continually growing global population, more land is made available for crop production. As a result, deep-rooted plants are cleared and irrigation is introduced in arid and semi-arid regions of the world. In the study area, the removal of natural vegetation made room for winter crop production, and the landscape is left bare during summer. The elimination of deep-rooted native plants and the use of incorrect irrigation methods can result in soil salinization (Maianu, 1984; Allison *et al*., 1990; Cox and McFarlane; 1995; Greiner; 1998; Pannell and Ewing, 2005; Kingswell and John, 2007; Hughes *et al*., 2008). According to Farifteh (2007) soil salinization is the accumulation of soluble salts at the surface or near-surface of the soil horizon. Two main groups of soil salinization occur, namely primary soil salinization and secondary soil salinization. Primary soil salinization occurs as a consequence of natural processes and secondary soil salinization is that which is caused by the activities of human intervention (for example incorrect irrigation methods and removal of deep-rooted plants). Soil salinity degrades agricultural land causing the decline in the productivity of plants and thus leading to the loss of agricultural yields (Patel *et al*., 2009). Salt-affected soils

occur across all continents and it is estimated that about 1 billion hectares of land around the world is affected by salts (Szabolcs, 1994; Metternicht and Zinck, 2003).

In spite of the awareness of the damaging effects that excessive salts in the soil have on agricultural production and the environment, it is recorded that the problem is growing rather than decreasing (Szabolcs, 1994; Metternicht and Zinck, 2003). For instance, Mirlas (2012) studied the soil salinity problem in cultivated areas in the Jezre'el Valley in Israel using the MODFLOW groundwater model and GIS techniques. They established that the area affected by soil salinity has been increasing. Gao *et al*. (2011) monitored the temporal and spatial dynamics of soil salinization changes in the upper stream of the Tarim River in China based on remote sensing and global positioning system. Their study found that the total area of salinized land increased. Martínez-Sánchez *et al*. (2011) evaluated salinization problems in Murcia Region using two chemical degradation indicators, salinization state and salinization rate. They concluded that salinization increased in certain areas. Nell (2009) reported that about 23% of South African soils are slightly saline, 5.1% saline, 1.4% moderately saline, 0.4% strongly saline, 3.8% saline-sodic (non-alkaline), 6.3% saline-sodic (alkaline), and 0.4% can be considered as sodic. And, Fey and de Clercq (2004) found convincing proof of dryland salinity in the Berg River catchment (BRC) in the Western Cape Province of South Africa. Timely identification of areas affected by salinization is essential so that reclamation strategies can be implemented to regain degraded land.

Traditionally, soil salinization is analysed and monitored by wet chemistry methods. Soil samples are collected at targeted areas and analysed in the laboratory for salinity. In recent times, the amount of land that is associated with the agricultural resources is vast. Consequently, the use of wet chemistry methods of analysing and monitoring soil salinity will be restrictive in terms of costs and labour involved. It is desirable to investigate reliable, less tedious and cost effective techniques to monitor soil salinization. Most recently, remote sensing, GIS and spectroscopy are among cheaper and less labour intensive methods that have been used for quantitative soil salinity analysis and monitoring.

Despite the success of broadband sensors (for example SPOT and Landsat) in identifying severely saline from non-saline land, broadband sensors have not been useful in identifying saline soils at early stages of development. Also, the use of near-infrared (NIR) spectroscopy as a

less expensive and less labour intensive technique for soil salinity studies is limited. Additionally, the use of terrain attributes for studying soil salinity has not yet been adequately investigated.

This study investigated hyperspectral remote sensing (HRS), NIR spectroscopy, terrain attributes, hydrological parameters and land components for improved soil salinity monitoring. HRS is the acquisition of spectral data using many contiguous bands. HRS provides near-laboratory spectra and has the potential to overcome the shortcomings of broadband sensors for studying soil properties. NIR spectroscopy is an easy to use and less expensive technique that has the potential to replace traditional wet chemistry methods of soil analysis. Owing to the value of hydrological parameters (for example streamlines and catchment boundaries) in modeling salinity at catchment scales using hydrological models, the suitability of digital elevation models (DEMs) to derive accurate hydrological parameters deserves evaluation. Accurate hydrological parameters would be very useful to accurately model salinity at catchment scales. Regarding land components, they are landform elements with a constant value of elevation, or of two or more readily interpretable morphometric variables, bordered by lines of discontinuities (Minar and Evans, 2008). Land component borders commonly coincide with environmental land properties such as soil, climate and biology (Speight, 1997; MacMillan *et al*., 2004; Van Niekerk, 2010). Accordingly, land components have great potential to map soil properties and thus can be useful for mapping soil salinity. Finally, topography is an important soil forming property and determines areas of groundwater discharge in the landscape. Thus, the value of terrain attributes for modeling soil salinity in the landscape is vital.

1.2 PROBLEM STATEMENT

The Berg River is a crucial water source for the Western Cape Province in South Africa. Declining water quality in the BRC triggered concerns of the emergence of dryland salinity. According to Hughes *et al*. (2008) river salinity is a typical indicator of land salinity. A study conducted in a small catchment which is representative of the semi-arid conditions in the BRC by Fey and de Clercq (2004) confirmed the existence of dryland salinity in the catchment. The clearing of deep-rooted native plants for wheat production is the source of dryland salinity in the BRC. Dryland salinity in the BRC will have severe implications for irrigation agriculture, wheat

production, the supply of clean water, ecology of the river itself and for irrigation agriculture (de Clercq *et al*., 2010).

It will be prohibitive in terms of costs and labour to use traditional ground-based methods to monitor salinity in the BRC due to its vast extent. Remote sensing, spectroscopy and DEM-delineated terrain parameters are attractive less labour intensive and cheaper alternatives for quantitatively analysing, mapping and modelling soil salinity. Even though broadband sensors have been successfully used to isolate severely salinized from non-salinized soils, salinized soils in the BRC are predominantly small patches that may be challenging to map using broadband sensors due to their inadequate spatial and spectral resolution. It is critical to accurately identify, quantify and map salinized soils timely so that mitigation strategies can be implemented when the situation is still manageable. Thus, it is valuable to investigate the utility of hyperspectral remote sensing, NIR spectroscopy and DEM-based methods to quantitatively analyse and map soil salinity for improved monitoring.

1.3 RESEARCH AIM AND OBJECTIVES

The aim of this study is to evaluate the value of HRS, NIR spectroscopy, DEM-based hydrological parameters and land components, and terrain attributes as accurate, less tedious and cost effective methods for enhanced monitoring of soil salinity. Soil salinity analysis, modeling and mapping techniques developed in this study will be tested in the Sandspruit catchment and ultimately used in the entire BRC. The Sandspruit catchment is representative of the BRC in terms of soils, climate, geology and terrain.

The specific objectives of this study are to:

1. Conduct a review of literature to reveal the value of remote sensing, spectroscopy, hydrological parameters, land components and terrain attributes for monitoring soil salinity,
2. Investigate the value of hyperspectral remote sensing for enhanced detection of soil salinity by remote sensing,
3. Evaluate the potential of NIR spectroscopy for quantitative analysis of soil salinity,
4. Assess the potential of DEMs to derive useful hydrological parameters for use in soil salinity modeling,

5. Investigate the accuracy of DEM-extracted land components for use in mapping soil properties (including soil salinity), and

6. Evaluate the value of terrain attributes to map the risk of soil salinity.

1.4 DATA REQUIREMENTS

The data sets used to realize the aim of this study are; soil samples, soil spectral data, groundwater hydro-census data, DEMs and orthorectified digital aerial images. Each of the data sets is briefly described in the subsequent subsections. The description of the data sets may be partly repeated in Chapters 4, 5, 6 and 7 for the purpose of clarity of the specific data set used because this PhD study was presented in a manuscript format.

1.4.1 Soil samples

Soil samples were sourced from databases held at the Agricultural Research Council Institute for Soil, Climate and Water (ARC-ISCW). The ARC-ISCW soil samples included Land Type Database (LTD) and *ad hoc* soil sample data. The LTD arose from the 1:250 000 scale soil mapping programme, carried out over a period of 30 years (1972-2002) by the ARC-ISCW (Land Type Survey Staff, 1972-2003). The information was systematically transferred to a GIS, along with the composition of each of the more than 7000 unique land type mapping units, as well as a supporting database containing the soil profile information. The ARC-ISCW *ad hoc* soil samples selected were collected on a monthly or bi-monthly basis over a period of 14 years from fixed sites south east of Johannesburg in the Gauteng Province in South Africa. Other soil samples (n = 23) were collected from within the Sandspruit River catchment and the nearby Langgewens Experimental Farm north of Cape Town in South Africa. The samples were dried, ground and sieved with a 2 mm sieve to remove large particles and plant remains. The soil samples collected from the Sandspruit catchment and the nearby Langgewens Experimental Farm were not ground or sieved. These samples were used for developing field-based soil salinity predictive models. The samples were analysed for electrical conductivity (EC), organic carbon, texture, pH, and soluble Ca, Mg, Na, Cl, and SO_4. EC was measured by a 1:5 saturated extract.

1.4.2 Soil spectral data

Soil spectral data were collected using two different spectrometer instruments, namely the analytical spectral device (ASD) FieldSpec spectrometer (http://www.asdi.com) and the Fourier-Transform (FT) Bruker multi-purpose analyser (MPA) laboratory spectrometer (http://www.bruker.com). The ASD Fieldspec instrument covers the visible to shortwave infrared wavelength range (350 to 2500 nm). The ASD spectrometer has a sampling interval of 1.4 nm for the region 350 to 1000 nm and 2 nm for the region 1000 to 2500 nm with a spectral resolution of 3 and 10 nm, respectively. A halogen lamp (Lowel Light Pro, JCV 14.5V-50WC) was used as a source of light with the ASD Feildspec spectrometer. Also, a white reference was used to calibrate the ASD Fieldspec spectrometer. The white reference is a calibrated white spectralon with a near-100% diffuse (Lambertian) reference reflectance panel made from a sintered poly-tetra-flourethylene based material. The spectra measured by the ASD spectrometer were used to investigate quantitative hyperspectral models for predicting soil salinity in dry soils by hyperspectral remote sensing.

The Bruker MPA spectrometer covers the wavelength range of 3595 to 12 489 cm^{-1} (equivalent to 800 to 2800 nm). According to Bruker Optics (2011), the MPA uses state-of-the-art optics for exceptional sensitivity and steadiness. The heart of the instrument is Bruker Optics' patented, permanently aligned RockSolid interferometer, which is equipped with gold-coated optics for maximum efficiency and sensitivity. The permanent alignment provides reliable high quality results, less downtime and great stability (Bruker Optics, 2011). The spectra measured by the Bruker MPA were used to investigate the utility of NIR spectroscopy to predict EC, pH, soluble Ca, Mg, Na, Cl and SO_4.

1.4.3 Groundwater hydro-census data

Hydro-census data was obtained from the National Groundwater Information System of the Department of Water Affairs and Forestry (DWAF) (http://www.dwa.gov.za), South Africa. The data records groundwater information and the geographic location. The data records amongst others EC range of the groundwater and precipitation. The EC ranges of the groundwater data were classified into six categories; namely 0 - 70, 70 -150, 150 - 300, 300 – 500, 500 - 1000 and greater than 1000 mS m^{-1}. The data covered the whole of the Berg River catchment and the

points were spaced at approximately 1 000 m from each other. The hydro-census data was used to investigate the relationship of the groundwater EC with terrain attributes.

1.4.4 Digital elevation models

Six DEMs were used in this study. These include the Shuttle Radar Topography Mission (SRTM) DEM, the second version of the 30-m Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model (ASTER GDEM2), very high resolution DEMs generated from GeoEye stereo images (GEOEYE DEM), two versions of the Stellenbosch University DEM's (SUDEM L1 and L2) and the 20-m Western Cape digital elevation model (WCDEM).

The 90-m SRTM DEM, completed in 2000, is the first high resolution DEM to be developed at near-global scale (Farr and Kobrick, 2001; Li and Wong 2010). The SRTM DEM has a vertical accuracy of less than 16 m (Rodriguez *et al*., 2005; Van Niekerk, 2008). According to the Consultative Group on International Agricultural Research Consortium for Spatial Information (CGIAR-CSI), the SRTM DEM has been processed to fill data voids, and can be used by a wide group of potential users (CGIAR, 2011).

ASTER GDEM was developed jointly by the Ministry of Economy, Trade and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA). The full 1.5-million-scene ASTER archive was used to create the DEM. The second version of the ASTER GDEM (GDEM2) was released in October 2011 (ASTER GDEM Validation Team, 2011) with the inclusion of 26 000 additional scenes to improve coverage. A smaller correlation kernel was also used to yield higher spatial resolution and enhanced water masking. ASTER GDEM2 was validated by comparing it to the absolute geodetic references over the conterminous United States (CONUS), the national elevation grids over the US and Japan, the SRTM 1 arc-second DEM over the US and 20 sites around the globe, as well as global space-borne laser altimeter data. The vertical and horizontal accuracy of the GDEM2 is less than 17 m and 71 m respectively (ASTER GDEM Validation Team, 2011; Mukherjee *et al*., 2013). The number of voids and artefacts noted in GDEM1 were substantially reduced in GDEM2 and were almost eliminated in some areas (ASTER GDEM Validation Team, 2011).

The GEOEYE DEMs were created from the GeoEye stereo images acquired in July 2011. Elevations were extracted at 1.5, 2 and 5 m horizontal interval using the rational polynomial coefficients (RPC) model in the LPS module of Erdas Imagine software. The resulting DEMs were validated using five trig beacon height points covering the Sandspruit catchment. They were found to have a vertical accuracy of about 0.70 m. Although the GEOEYE DEM is a surface model, most of the study area is used for cultivation of grains and very few tall objects (e.g. trees and buildings) are present. At the time of acquisition, the crops were still at seedling height and had very little impact on the extracted elevations. The 1.5-m GEOEYE DEM was used to delineate a reference catchment boundary and reference streamlines that were used to validate hydrological parameters delineated from the other DEMs. While the 2-m GEOEYE DEM was used to extract hydrological parameters for the purpose of comparison with previous studies undertaken using a 2 m light detection and ranging (LiDAR) DEM, the 5-m GEOEYE DEM was used to generate land components for comparison with the other DEMs evaluated in this study.

The SUDEM was developed by the Centre for Geographical Analysis (CGA) at the Stellenbosch University, South Africa. Large scale (1: 10 000) contours were used to interpolate two DEM products (Van Niekerk, 2011). The first product (Level 1) only used contours while the second product (Level 2) combined contours and the SRTM DEM (at flat areas). The SUDEMs were used to derive hydrological parameters and land components. The SUDEM L2 was also used to investigate the relationship of soil salinity with terrain attributes.

The 20-m WCDEM was developed by the Stellenbosch University CGA using contours digitized from 1:50 000 national topographic map series (van Niekerk, 2001). The vertical accuracy of the WCDEM was determined by computing mean absolute error (MAE) and root mean squares error (RMSE) using highly accurate elevation data obtained from Chief Directorate National GeoSpatial Information (CDNGI), South Africa (van Niekerk, 2008). It was found to have a MAE and RMSE of 7 and 10 m respectively, and is more accurate than the 90-m SRTM DEM.

1.4.5 Orthorectified digital aerial photographs

The orthorectified digital aerial photographs covering the Sandspruit catchment were obtained from the CDNGI. The resolution of the digital aerial images was 0.5 m. The orthorectified digital

aerial photographs were used as a backdrop for visual inspection of DEM delineated hydrological parameters and for delineating test terrain morphological discontinuities for validating land components.

1.5 THE STUDY AREA

The Sandspruit (quaternary catchment G10J) catchment is study area of interest. Soil salinization quantitative analysis, mapping and modelling techniques developed in this study will be applied in the Sandspuit. Ultimately, the techniques will be utilized in later studies to map dryland salinity risk in the entire BRC and other suitable sites. The Sandspruit River catchment is a tributary to the Berg River. The catchment is 152 km^2 in size and is situated in the vicinity of Riebeek West, north of Cape Town in the Western Cape Province of South Africa (Figure 1.1).
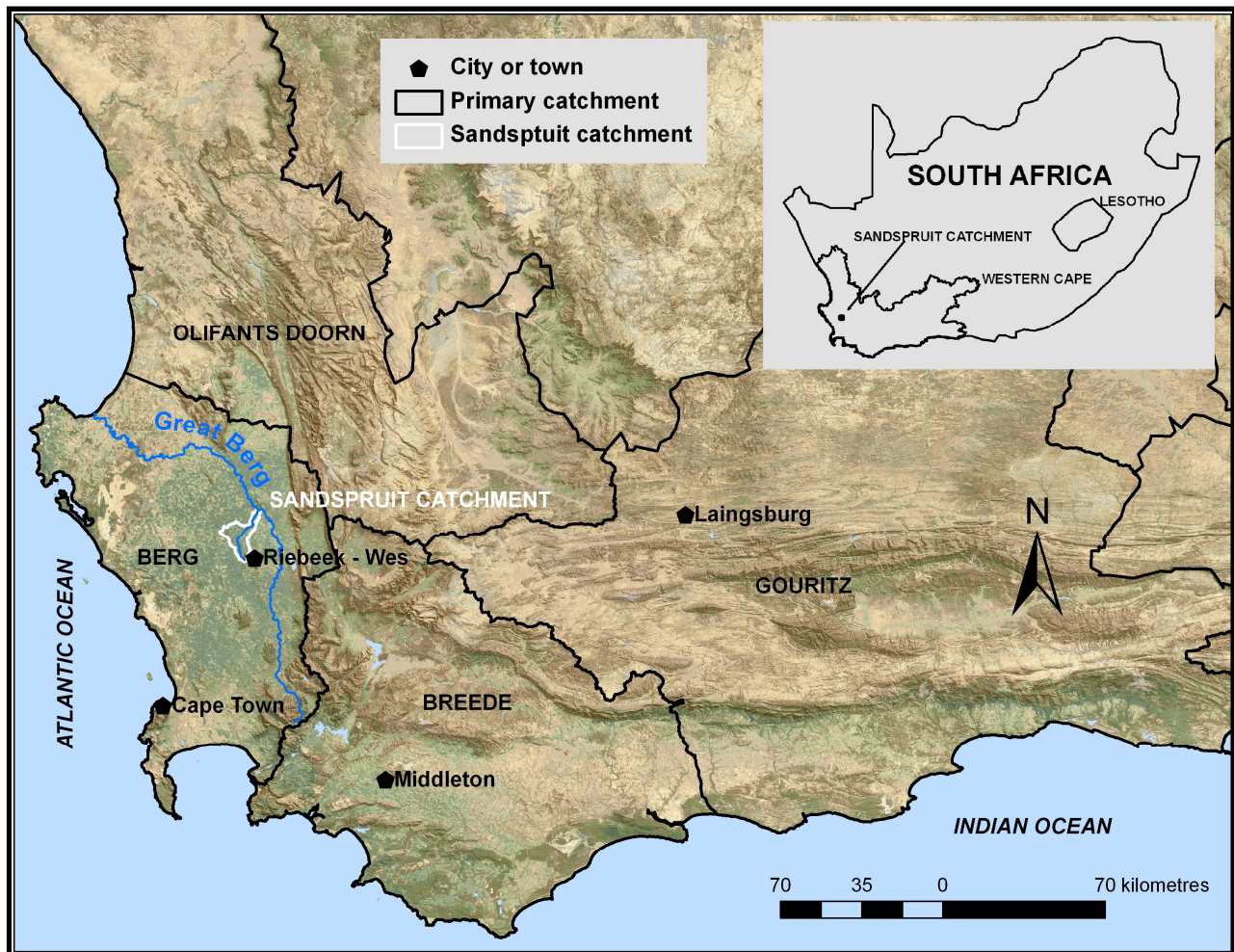
Figure 1.1 Location of the Sandspruit catchment

The geology in the Sandspruit catchment shows minimal variation and is dominated by the Table Mountain Group sandstone in the high elevation areas and Malmesbury shale in the mid- to low elevation parts. Most of the catchment is used for dryland cultivation, in particular winter wheat. Canola and pasturage also occur. Natural vegetation occupies only a small proportion of the catchment. The catchment was mainly chosen because it has a Department of Water Affairs and Forestry (DWAF) monitoring weir with a long-term record of salt and water discharge into the Berg River. Additionally, the Sandspruit catchment is the major contributor to the salinity of water in the Berg River.

The Sandspruit catchment has a semi-arid (Mediterranean) climate and is located in a winter rainfall region with a mean annual rainfall of about 400 mm (Flügel, 1995). Rainfall is generally in the form of frontal rain approaching from the north-west, extending normally over a few days with significant periods of clear weather in between. Annual rainfall in the Sandspruit catchment is between 300 and 500 mm, with slightly higher rainfall in the upper southern reach in the vicinity of Kasteelberg (Riebeeck West).

The catchment generally has an undulating topography with gentle to moderate slopes. About 61% of the catchment has slope gradients between 0 and 4 degrees, 27% has slope gradients between 4 and 7 degrees, and 12% has slope gradients greater than 7 degrees. According to Flügel (1995), the valleys have a moulded shape and a shallow groundwater table occurs in the lower parts during the winter rainfall season. Salt crystallizes in patches during the hot summer between November and March. The Sandspruit catchment is representative of the BRC in respect of climate, geology, soils and geomorphology.

The study area description and the map depicting the Sandspruit catchment will not be repeated in Chapters 5, 6 and 7 where it is needed, but will be referred to. This was done to minimize repetition in the dissertation.

1.6 RESEARCH METHODOLOGY AND AGENDA

The use of conventional wet chemistry methods to monitor soil salinity is not viable for large study areas due to high costs. Again, it will be challenging to use broadband sensors to map dryland salinity in the BRC because it predominantly manifests as small patches that may be partially covered with vegetation. It would be valuable to investigate less tedious, reliable and

cost effective methods for better monitoring. This study provides a unique opportunity to evaluate the utility of HRS, NIR spectroscopy, terrain attributes, DEM-delineated hydrological parameters and land components for quantitative analysis, mapping and modelling of soil salinity for better monitoring. This research is organized into eight (8) chapters. The overview of this research is summarized in the research design in Figure 2.2.

The formulation of the research problem, research aims and objectives, research methodology, description of the study area and explanation of data sets used is given in this chapter (Chapter 1).

RESEARCH PLAN (CHAPTER 1)

- Formulate research problem
- Aims and objectives
- Research methodology and outline

LITERATURE REVIEW (CHAPTER 2)

- Soil salinity hazard
- Remote sensing of salt-affected soils
- Spectroscopy of salt-affected soils
- Land components and soil mapping
- Hydrological parameters and salinity modeling
- Topography and soil properties

SYNTHESIS: IMPROVED SOIL SALINITY MONITORING (CHAPTER 8)

- Evaluate the success of the research
- New developments
- Directions for further research

HYPERSPECTRAL REMOTE SENSING OF SALINE SOILS (CHAPTER 3)

- Individual bands
- Normalized difference soil salinity index
- Partial least squares regression (PLSR)
- Bagging PLSR
- Field-based predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4

TERRAIN ATTRIBUTES AND SOIL SALINITY (CHAPTER 7)

- Investigate relationship of soil salinity with terrain attributes
- Develop terrain attributes-based soil salinity predictive models
- Map salinity risk in the Sandspruit catchment

NEAR-INFRARED SPECTROSCOPY OF SALINE SOILS (CHAPTER 4)

- Predicting EC, pH, soluble Ca, Mg, Na, Cl and SO_4 using NIR spectroscopy

LAND COMPONENT EVALUATION (CHAPTER 6)

- Delineate land components from DEMs
- Evaluate land components visually and quantitatively

EVALUATION OF DEM-DELINEATED HYDROLOGICAL PARAMETERS (CHAPTER 5)

- Delineate hydrological parameters (streamlines and catchment boundaries) from DEMs
- Compare the accuracy of delineated hydrological parameters

Figure 1.2 Research design

Chapter 2 gives an account of the theoretical background of the soil salinity risk. Quantitative analysis using spectroscopy, detection and mapping by remote sensing and the value of DEM-based terrain attributes to study soil salinity is also outlined in this chapter.

Chapters 3, 4, 5, 6 and 7 present the scientific papers responding to the objectives of this study. Chapter 3 investigated the potential of hyperspectral individual bands, a normalized difference salinity index (NDSI), PLSR and bagging PLSR to predict soil salinity. Spectra of dried, ground and sieved soil samples were measured using an ASD Fieldspec spectrometer in a darkroom. The spectral range of the ASD Fieldspec spectrometer is 350 to 2 500 nm. A halogen lamp was used as a source of light. The spectrometer was calibrated with a white reference before taking measurements of each sample. The spectral signatures were taken off-nadir at a height of about 15 cm above the target to minimize the effects of bidirectional reflectance. Soil salinity predictive models were computed based on untransformed spectra, first derivative reflectance (FDR), PLSR and bagging PLSR using calibration samples (n = 63). The predictive models were validated using an independent sample (n = 32) set which was not used in the development of models. Furthermore, field-based regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were developed using soil samples (n = 23) that were not ground or sieved to simulate field conditions. These samples were collected from the Sandspruit catchment and the nearby Langgewens Experimental farm. Spectral signatures of these samples were measured on a clear sky day using the sun as a source of light. A total of 118 soil samples spread throughout South Africa were used for this analysis.

Chapter 4 presents the potential of NIR spectroscopy to predict EC, pH, soluble, Ca, Mg, Na, Cl and SO_4. Spectra of ground, dried and sieved soil samples were measured using a FT Bruker MPA spectrometer. A total of 49 samples were used for this analysis. Due to a small number of samples, a full leave one out cross validation (LOOCV) was used to calibrate the PLSR predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4. Not many studies that used NIR spectroscopy to predict EC and soluble cations and anions were found. To the best of the knowledge of the researcher, it is the first time that spectroscopy and chemometric modelling was used to predict Cl and SO_4. The PLSR predictive models were validated using the R^2, root mean square error of cross validation (RMSECV), ratio of prediction to deviation (RPD) and the

ratio of performance to inter-quartile distance (RPIQ). The use of RPIQ for evaluating pH, EC and soluble cations and anions was not found in the literature.

Chapter 5 presents the potential of DEMs to delineate accurate hydrological parameters (streamlines and catchment boundaries) for enhanced soil salinity modelling by hydrological models. Four DEMs, namely two 5-m SUDEM (L1 and L2), the 30-m ASTER GDEM2 and the 90-m SRTM DEM were used. A 1.5-m GEOEYE DEM which was developed from the GeoEye stereo-images was used to generate a reference catchment boundary and reference streamlines. The Arc Hydro module for ArcGIS software was used to extract streamlines and catchment boundaries at the native resolutions of the DEMs. Outlet (pour) points were selected at the same position for catchment boundary delineation. Streamlines were also delineated from a 2-m GEOEYE DEM to enable comparison with previous studies conducted with high-resolution light detection and ranging (LiDAR) DEMs. The catchment boundaries and streamlines were converted to raster data sets with a cell size of 5 m for comparison purposes. The DEM delineated catchment hydrological parameters were validated using visual inspection, a correctness index (C_r), the figure of merit index (FMI) and a new Euclidean distance (ED) index.

Chapter 6 evaluates the suitability of DEMs to derive useful land components for mapping soil properties. Land components derived from the 30-m ASTER GDEM2, 90-m SRTM DEM, 5-m SUDEM L1 and L2, and a 5-m GEOEYE DEM using the multi-resolution segmentation (MRS) algorithm of eCognition 8.6 software. The SRTM DEM and the ASTER GDEM2 were up-sampled to 5 m resolution for easy comparison. Land components were delineated using the slope gradient percentage and aspect derivatives of each DEM. A suitable scale factor was determined by experimentation and visual inspection using hill-shaded DEMs as backdrops. The experimentation with a suitable scale factor was carried out on the DEM with the highest detail (i.e. GEOEYE DEM). The scale factors for the other DEMs were adjusted so that their segmentations yield a similar number of objects to allow comparison. The resulting land components were visually inspected and quantitatively analysed using the slope gradient standard deviation (SGSD) measure and a novel mean slope gradient local variance (MSGLV) ratio.

Chapter 7 investigates the value of terrain attributes to predict soil salinity. Terrain attributes-based soil salinity regression predictive models based on the EC of top soil and groundwater are

developed using CurveExpert software (www.curveexpert.net). Stepwise multiple linear regression soil salinity predictive models were also developed based on terrain attributes, the annual evapotranspiration and the aridity index. The models are validated using R^2, correlation coefficient and standard error. An independent groundwater hydro-census data covering the Sandspruit catchment is also used to validate the models. Although the models are less accurate, it is promising that further studies using better data sets will improve the results. Sound potential soil salinity maps of the Sandspruit were produced based on elevation, slope, evapotranspiration and terrain wetness index (TWI).

Chapter 8 presents the overall findings, the implication for future soil salinity quantitative analysis, mapping and modelling for accurate monitoring, and the recommendations for further research. Overall this study demonstrated that HRS, NIR spectroscopy, land components, hydrological parameters and terrain attributes present positive developments for improving soil salinity monitoring.

CHAPTER 2

REMOTE SENSING AND MODELLING OF SALT-AFFECTED SOILS:

A REVIEW

2.1 INTRODUCTION

Salinization of soils is a global problem that harmfully affects the productivity of soils consequently threatening the sustainability of agricultural production. This problem is common in soils of arid and semi-arid regions (Allison *et al*., 1969; Fitzpatric, 1980; Rowel, 1994) and is practically non-existent in humid regions except when the soils have been subjected to the effects of seawater in river deltas and low lying areas near the sea (Allison *et al*., 1969). Soil salinization can develop as a consequence of natural processes (primary soil salinization) and also due to human activities (secondary soil salinization). While irrigation is a main source of human-induced soil salinization (Lenney *et al*., 1996; Katerji *et al*., 1998; Utset and Borrotto, 2001; Slavisch *et al*. 2002; Pannell and Ewing, 2005), the removal of deep rooted plants also contribute to soil salinization in areas where the water table is closer to the surface also causes salinization. Irrigation-induced salinization is a result of the extreme use of low quality water (where the soil lacks sufficient drainage) for irrigation (Rietz and Haynes, 2003; Acosta *et al*., 2011; Iwai *et al*., 2012). Dryland salinization is largely caused by rising water tables due to more rain water entering the subsoil and causing the water table to rise. Deep-rooted native plants capture most of the rainfall water, thus minimizing rates of groundwater recharge (Talsma, 1981; Pannell and Ewing, 2005; Robertson *et al*., 2009). All continents of the globe are affected by salts (Figure 2.1). According to Sumner (2000), 20% of cultivated land around the globe is affected by salts. In total, it is estimated that 1 billion hectares of land in the world is affected by salts.
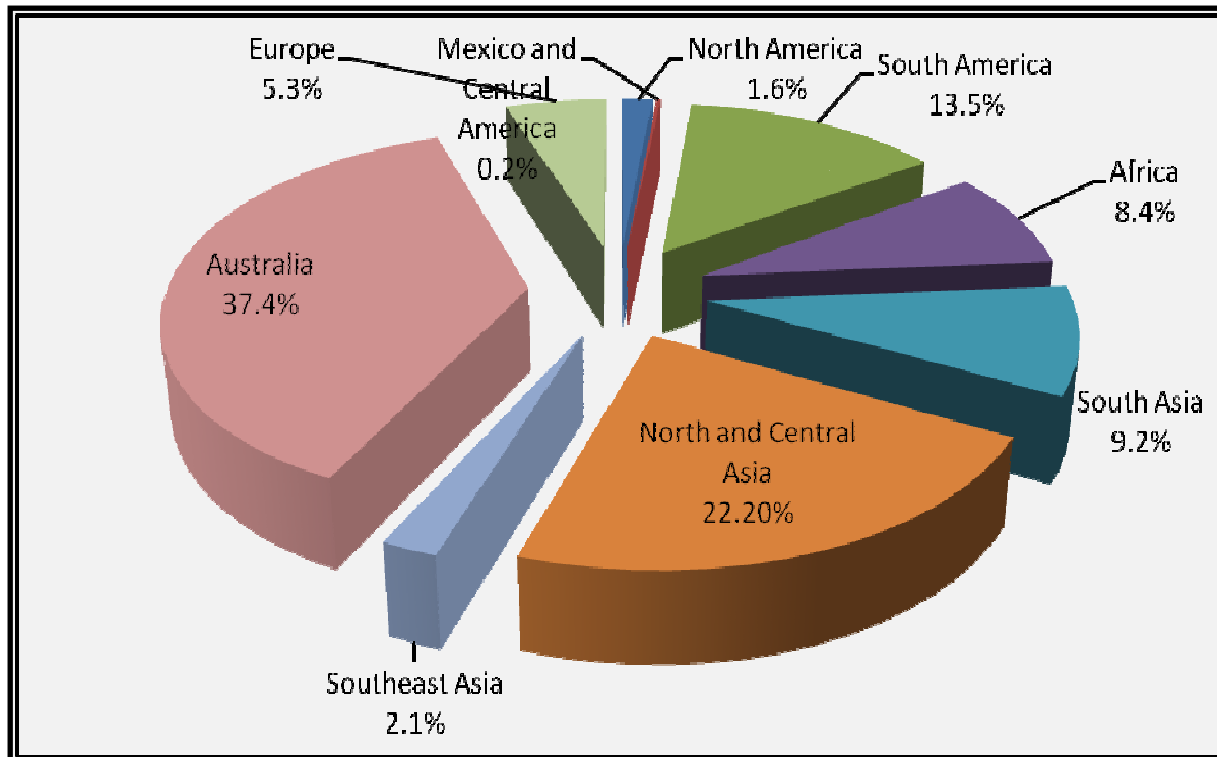
Figure 2.1 The extent of salt-affected soils in the world

Adopted from: Szabolcs. 1994: 6

Conventionally, soil salinization is analysed and thus monitored using ground-based methods. Soil samples are collected at chosen ground points and analysed for EC in the laboratory. Due to the large global acreage that is associated with agricultural resources to date, the use of field survey methods for soil salinity analysis and monitoring would be prohibitive in terms of labour and costs. Recently, GIS, remote sensing and spectroscopy can provide accurate, less labour intensive and cheaper alternatives for soil salinity analysis, mapping and modelling, thus resulting in improved monitoring.

This chapter presents a theoretical account of the methods that are commonly used in the analysis and monitoring of salt-affected soils. Firstly, remote sensing of salt-affected soils is presented. Secondly, laboratory spectroscopy and chemometric modeling of soils is considered followed by the value of DEMs to delineate accurate hydrological parameters and digital terrain mapping for soil property mapping. Lastly, the relationship of soil salinity with terrain attributes is given.

2.2 REMOTE SENSING OF SALT-AFFECTED SOILS

The use of remote sensing to study salinized soils is striking because images can cover large areas and the data is collected in a non-subjective manner. Generally, the identification and monitoring of salt-affected soils by remote sensing is conducted using aerial photographs, video images, infrared thermography, visible and infrared multispectral and microwave images (Metternicht and Zinck, 2003; Abbas *et al*., In press). Basically, two approaches can be used to detect salt-affected soils by remote sensing; the first approach entails the detection of salt-affected soils directly by analysing the spectral reflectance of soils and the second approach infers salt-affected soils by analysing the spectral reflectance of vegetation or crops growing on affected soils. Largely, broadband sensors have been used for studying salt-affected soils. Lately, the use of imaging spectroscopy for monitoring salt-affected soils has been increasing. Although light detection and ranging (LiDAR) and radio detection and ranging (RADAR) can be used for monitoring salinized soils, their use is not widespread. The use of LiDAR and RADAR for studying saline soils is not considered in this review because their use in studying salt-affected soils is limited.

2.2.1 Broadband remote sensing of salt-affected soils

In the main, broadband sensors have been used for detecting salinized soils and yielded satisfactory results (Csillag *et al*., 1993; Metternicht, 2001; Metternicht, 2003; Khan *et al*., 2005; Mashimbye, 2005; Douaoui *et al*., 2006; Fernandez-Buces *et al*., 2006; Gutierrez and Johnson, 2010; Yu *et al*., 2010). For example, Jian-li *et al*. (2011) used a Landsat Enhanced Thematic Mapper Plus (ETM+) image using a decision tree approach to determine the key variables to be used for classification and extraction of salinized soil from other cover and soil types using principal component analysis (PCA). Their study revealed that the PC3 was the best band to identify areas of severely salinized soil while the blue spectral band from the enhanced thematic mapper plus sensor (TM1) was the most appropriate to recognize salinized soil by identifying salt-tolerant vegetation. Bouaziz *et al*. (2011) reported moderate correlations between EC and spectral indices using a linear spectral unmixing (LSU) technique to improve the prediction of salt-affected soils using MODIS data. In addition, they established that the use of the LSU enhanced the correlations. Ding *et al*. (2011) classified land cover into different levels of soil salinity using the Landsat ETM+ image by means of PCA and decision tree approach. They

obtained an accuracy of 95%. Melendez-Pastor (2010) used image based spectra of saline and non-saline training areas, and the spectrum of the halite mineral as a surrogate to the spectra of saline soils to map saline soils using multispectral ASTER images by matched filtering and mixture tuned matched filtering techniques. They reported that the image based approach was the most accurate approach for saline soil mapping and monitoring.

Notwithstanding the success of multispectral sensors in mapping severely salt-affected soils, they have limitations in mapping slight to moderately affected soils (Farifteh *et al*., 2006; Weng *et al*., 2010). This is attributed to their low spectral resolution and the use of traditional classification techniques (Dehaan and Taylor, 2003; Tamas and Lenart, 2006). Additionally, it is not easy to map small areas including areas with a good cover of salt tolerant plants (Dutkiewicz, 2006).

2.2.2 Hyperspectral remote sensing of salt-affected soils

Hyperspectral remote sensing has the potential to overcome the spectral limitations of the broadband data as it provides near-laboratory quality spectra for each pixel. Each picture element contains a distinctive spectrum which can be used for the recognition of earth's surface materials. The spectrum allows the discrimination of slight differences between materials, permitting the investigation of phenomena that impressively extend the capability of traditional remote sensing (Chang, 2003; Lillesand *et al*., 2004; Campbell, 2007). This is possible because of the contiguous nature of the spectral profile of a hyperspectral signature.

Currently, researchers have been investigating the value of hyperspectral remote sensing to enhance the detection of salt-affected soils using remote sensing (Ben-Dor and Banin, 1994; Drake, 1995; Ben-Dor *et al*., 2002; Dehaan and Taylor, 2002; Dehaan and Taylor, 2003; Tamas and Lenart, 2006). It is anticipated that hyperspectral remote sensing will yield more accurate detection of salinized soils. Encouraging findings on the value of hyperspectral remote sensing for studying salinized soils have been reported. Farifteh *et al.* (2007) used PLSR and obtained prediction R^2 values between 0.78 and 0.98 using experimental soil sample data, which in each sample was treated with different salts (namely $MgSO_4$, KCl, $NaCl$, and $MgCl_2$). Weng *et al.* (2010) developed a univariate regression model to estimate soil salt content using a soil salinity index. The index was constructed from continuum-removed reflectance at 2052 and 2203 nm. Their model was applied to a Hyperion reflectance image and was successfully validated (R^2 =

0.63). Zhang *et al*., (2011) investigated the relationship between vegetation spectra and soil salinity. They derived vegetation indices from the recorded hyperspectra and then evaluated their predictive power for salinity. Subsequently, they employed a univariate linear correlogram as well as multivariate PLSR analysis to investigate the sensitive bands. They concluded that there is potential to monitor soil salinity with the hyperspectra of salt-sensitive and halophyte plants. Wang *et al*. (2012) developed an exponent reflectance model to estimate soil salt contents under various soil moisture conditions based on a control laboratory experiment on the two factors (soil salinity and soil moisture) to soil reflectance. They examined Na_2SO_4, NaCl, Na_2CO_3 with wide soil salinity (0% to 20%) and soil moisture (1.75% to 20%) (In weight base) levels for their effects on soil reflectance through a model based approach. They identified moisture resistant but salt sensitive bands of reflected spectra for the model before applying them to inversely estimate soil salt content. They found high R^2 of 0.87, 0.79, and 0.66, and low means relative error of 16.42%, 21.17%, and 27.16% for NaCl, Na_2SO_4 and Na_2CO_3, respectively.

The use of hyperspectral remote sensing for studying soil salinity is not yet fully established. More investigations are still needed to uncover techniques to enhance the detection of salinized and/or salt-affected soils.

2.3 SPECTROSCOPY AND CHEMOMETRIC MODELING OF SALT-AFFECTED SOILS

2.3.1 Spectroscopy of salt-affected soils

Using traditional wet chemistry techniques for soil salinity analysis may be restrictive due to high costs and labour when large amounts of samples have to be analysed. It is accepted that near infrared (NIR) spectroscopy and mid-infrared (MIR) spectroscopy are among less expensive and user-friendly techniques for quantitative soil analysis (Shepherd and Walsh, 2002; Brown *et al*., 2006; Bellon-Maurel *et al*., 2010; Bilgili *et al*., 2010; Bellon-Maurel and McBratney, 2011). It is perhaps the benefits regarding costs and less labour that makes the use of spectroscopic methods attractive, particularly because land that is under agriculture is massive these days. However, the soil sample preparation requirements for NIR and MIR spectroscopic analysis are not the same. MIR spectroscopy requires more sample preparation in order to optimize the light interaction while at the same time showing better specificity and reproducibility than NIR spectroscopy

(Bellon-Maurel and McBratney, 2011). On the other hand, NIR spectroscopy is reported to be easy to use and requires less sample preparation (Viscarra Rossel 2006; Bellon-Maurel *et al*., 2010; Bellon-Maurel and McBratney, 2011).

While not many studies that used NIR spectroscopy for studying soil salinity were found in the literature, it appears that the adoption of spectroscopic techniques for soil analysis is gaining momentum nowadays. For example, Bilgili *et al*. (2010) evaluated visible-near infrared reflectance (VNIR) spectroscopy for prediction of diverse soil properties related to four different soil series of the entisol soil group within a single field in northern Turkey. They obtained strong correlations for exchangeable Ca, Mg, cation exchange capacity, organic matter, clay, sand, and $CaCO_3$ contents. Bellon-Maurel *et al*. (2010) investigated the critical aspects to be conscious of when assessing NIR spectroscopy measurements for soil analysis. They concluded that attention should be paid to reducing bias. Moreover, they found that because the standard deviation (SD) does not describe correctly the spread of the population in skewed reference values for soil studies, the use of ratio of prediction to deviation (RPD) and its thresholds for model evaluation may be misleading. They proposed a new index called the ratio of prediction to inter-quartile distance (RPIQ). This index uses inter-quartile distance instead of the SD. Research on the index is still in progress. Viscarra Rossel *et al*. (2006) reviewed the literature comparing quantitative predictions of various soil attributes using a multivariate statistical technique and spectral response in the ultra violet (UV), visible (VIS), NIR and mid MIR regions of the electromagnetic spectrum. They tabulated the soil attributes studied, spectral regions, spectral range, R^2, multivariate methods used, root mean squares error (RMSE) and the number of validation and calibration samples used. Their work presents a comprehensive base of what can be achieved with respect to spectroscopy of soils.

2.3.2 Chemometric modeling of soil chemical variables

A variety of statistical methods are used by researchers to extract soil attributes from the spectra. The statistical treatments that are used to enhance the extraction of soil attribute information from spectra include amongst others principal component regression (PCR), multiple regression analysis (MRA), stepwise multiple linear regression (SMLR), bagging PLSR and multivariate adaptive regression splines (MARS). Spectral transformations (mathematical treatments) are also applied to the spectra to maximize the extraction of information from spectra. The mathematical

spectral treatments include first and second derivatives, straight line subtraction, vector normalization, and multiplicative scattering correction, to mention a few. It appears that the use of statistical methods and spectral transformation frequently have a favourable result for enhancing the extracting of soil information from spectra. For example, Janik *et al*. (2009) compared the performance of PLSR analysis for the prediction of a variety of soil chemical and physical properties from their MIR spectra using a combination of PLSR and neural networks (NN). While their study established that the PLSR-NN method outperformed the PLSR for the prediction of some soil properties they cautioned that the use of PLSR-NN over the PLSR should be questioned against the backdrop of the trade-off of limited improvement and the added computational complexity. Cozzolino and Moron (2003) used modified partial least squares regression (MPLS) and first derivative transformation of the reciprocal reflectance to analyse soil samples for silt, sand, clay, Ca, K, Mg, Cu and Fe. They used cross validation to avoid over fitting of models. They obtained R^2 values of 0.84, 0.80, 0.90, 0.95, 0.80, 0.90, 0.86 and 0.92 for silt, sand, clay, Ca, K, Mg, Cu and Fe respectively. Primarily, PLSR is the most commonly used statistical spectral treatment technique for soil analysis. Bilgilli *et al*. (2010) asserts that this is mainly because PLSR is superior to traditional methods in dealing with high dimensional multi-collinearity in the data.

2.4 TERRAIN ATTRIBUTE-BASED MAPPING OF SOIL PROPERTIES

Land components, hydrological parameters and terrain attributes can be useful for understanding soil properties. According to Moller *et al*. (2008) and Jenny (1941) landforms and landscape circumstances are crucial for revealing the processes of soil genesis and soil formation in the spatial domain. Thus, digital terrain analysis is beneficial to the establishment of quantitative variables that reveal geomorphic, soil properties, climatic and hydrologic processes (McKenzie and Ryan, 1999). Undoubtedly, DEM-derived attributes can be useful to study the dynamics of soil salinization in the landscape. Evidently, land components possess enormous significance for studying soil properties. Also, accurate hydrological parameters can be crucial for accurately modeling soil salinization in the landscape using hydrological models.

2.4.1 Land components

Land component borders often correspond with environmental land properties such as soil, climate and biology (Speight, 1997; MacMillan *et al*., 2004). According to Hengl and Reuter (2008) and McBratney *et al*. (2003) soil landform units are expected to be relatively homogeneous in terms of the main factors including parent material. This means that accurately delineated land components can be useful for studying soil salinization. Owing to the use of conventional approaches (for example studying topographical maps, interpreting aerial photographs and making field measurements) to delineating land components (Speight, 1977; Graff and Usery, 1993; Drăgut and Blashke, 2006), it is not surprising that these methods are often time consuming, subjective and expensive (Speight, 1977; Argialas, 1995; Adediran *et al*., 2004; Drăgut and Blaschke, 2006; Van Niekerk, 2010).

The increasing availability of DEMs has promoted the use of computers for deriving terrain properties. Pixel-based classification techniques have been commonly used for land component classification. For example, Prima *et al*. (2006) proposed a generic landform classification method using a supervised classifier of four morphometric parameters from DEM-derived slope and topographic openness. They concluded that the constructional and erosional landforms and their evolutional stages were distinguishable using standard deviation ellipses, however, a post-processing procedure to remove noisy cells was recommended. Wuest and Zhang (2009) reported that obtaining high levels of classification accuracy using certain pixel based methods such as neural networks or fuzzy based classifiers can be time consuming. Furthermore, pixel-based classification yield problems of mixed pixels, commonly known as the 'salt and pepper effects' (Laliberte *et al*., 2004; Bhaskaran *et al*., 2010; Saha *et al*., In press).

Most recently, concerns have been raised about the efficacy of the conventional pixel-based analysis methods to represent real world objects (Smith *et al*., 2007; Wulder *et al*., 2008). Although terrain information is represented through points, objects have been proposed as alternatives, mostly for soil-landscape modelling purposes (Blaschke and Strobl, 2003; Deng, 2007; Drăgut *et al*., 2010). The benefit of the object-oriented approach is that it offers new potential for image analysis since image objects can be characterized incorporating spectral values, texture, shape and context (neighbourhood) relationships (Bock *et al*., 2005). Object-

based image analysis methods make use of segmentation algorithms to segment images into non-overlapping homogeneous objects.

Research on the use of DEMs for terrain and soil mapping has been steadily increasing. For instance, Drăgut *et al.* (2010) presented a technique for estimating the scale parameter in image segmentation of remotely sensed data with eCognition software. They proposed a tool, called estimation of scale parameter (ESP), which builds on the idea of local variance (LV) of object heterogeneity within a scene. They argued that the ESP tool enables fast and objective parameterization when performing image segmentation and possess great potential for object based image analysis applications. Drăgut and Eisank (2011) explored the relationships between object delineation and classification or regionalization in the framework of differences between general and specific geomorphometry. They concluded that discrete geomorphometry would apply to and describe land-surface divisions defined solely by the criteria of homogeneity in respect to a given land-surface parameter or a combination of several parameters.

No studies that investigated the value of land components for studying soil salinity were found in the literature. For a start, it would be beneficial to compare the accuracy of land components derived from different DEMs. Accurate land components derived in a non-subjective and cost effective manner will be indispensable for studying soil properties because they are theoretically uniform in terms of parent material and terrain attributes.

2.4.2 DEMs and hydrological parameters

Accurate hydrological parameters (streamlines and catchment boundaries) are crucial for hydrological studies. These parameters are used in hydrological models to model salinity at catchment scales. According to O'Callaghan and Mark (1984), Martz and De Jong (1998) Renssen and Knoop (2000), Turcotte *et al.* (2001), Vogt *et al.* (2003) and Li and Wong (2010) DEMs play an important role for delineating catchments, identifying sub-basins and deriving streamlines. Due to the recent increasing availability of DEMs at high to medium-resolution near-global DEM (e.g. the SRTM DEM and the ASTER GDEM), new opportunities for carrying out hydrological analyses on regional or national levels are available. Delineating hydrological parameters from DEMs will most likely yield precise delineation. It is also highly likely to be useful, fast and cost effective.

DEM's from different sources have been evaluated by a variety of researchers for hydrological studies. Terrain complexity, algorithms, resolution and the accuracy of DEMs have been reported to affect the quality of hydrological parameters. Areas of low terrain complexity yield less accurate hydrological parameters compared to areas of moderate to high terrain complexity (Wang and Yin, 1998; Vogt *et al*., 2003). Research has shown that algorithms considerably affect the delineation of hydrological features from DEMs (Gyasi-Agyei *et al*., 1995; Tarboton, 1997; Jones, 1998; Barker *et al*., 2006; Wise, 2007). For example, Callow *et al*. (2007) investigated three algorithms (namely Stream burning, Agree and ANUDEM) for modifying a DEM to reveal known hydrology. They concluded that different methods yield non-convergent results for catchment parameters (such as catchment boundaries, stream position and length). Seyler *et al*. (2009) recorded that the D8 algorithm does not always give the correct delineation of sub-catchments corresponding to the gauging stations when using the GTOPO30 DEM.

Most recently, availability of medium- (90 m) and high-resolution (30 m) near-global DEMs has opened up new possibilities for hydrological analyses at national and regional scales (Wang *et al*., 2011; Zeilhofer *et al*., 2011; Weepener *et al*., 2012). Researchers use these DEMs for hydrological studies, largely because they are freely available (Wang *et al*., 2011; Gichamo *et al*., 2012; Weepener *et al*., 2012). However, the value of the products that are derived from these DEMs is not known. It is desirable to compare the spatial accuracy of catchment boundaries and streamlines derived from DEMs. Consequently, the value of DEM-extracted hydrological parameters for modeling soil salinity at catchment scales will be exposed.

2.4.3 Soil salinization and topography

Topography is an essential soil forming factor. The discharge of salts from groundwater to the surface is also dependent on topography. Thus, elevation and its derivatives would be essential for studying soil salinity. According to Dowling *et al*. (2003) most of the parameters associated to salinization processes are related to elevation and landscape position. Nevertheless, Clarke *et al*. (1998) and Barret-Lenard and Nulsen (1989) recorded that landscape alone is not adequate to recognize saline areas. According to Malins and Metternicht (2006) and Salama *et al*. (1993), the water table is closer to the surface at low areas in the landscape, and further away from the surface at high points in the landscape. It is documented that ground water discharge occurs at low lying areas in the landscape (Freeze and Cherry, 1979; Evans *et al*., 1990; Williamson, 1998;

Fetter, 2001). Not many studies investigated the relationship of EC with topography. A study by de Clercq *et al*. (2010) established that there is an inverse relationship of EC with topography.

It is anticipated that saline prone areas in the landscape can be identifiable based on terrain attributes. It is projected that high areas in the landscape will be less susceptible to salinity, while low lying areas are expected to be prone to soil salinity. Thus, it is important to evaluate the value of elevation and its derivatives to predict the risk of soil salinity in the landscape.

2.5 CONCLUSIONS

The review of literature presented in this chapter outlined the use of remote sensing, spectroscopy and terrain attributes for quantitative analysis of soil salinity, mapping and modeling. There is consensus in the literature that traditional methods for studying salt-affected soils are not adequate due to the high costs and labour required. Cost effective and timely approaches are needed to enhance the quantitative analysis of soil salinity, mapping and modeling for improved monitoring.

The literature revealed that the use of broadband sensors for studying salt-affected soils is not satisfactory due to limited spectral information. Hyperspectral remote sensing is promising to overcome the limitations of broadband sensors. The use of hyperspectral remote sensing for monitoring salt-affected soils still warrants more investigations.

In the case of spectroscopy, notwithstanding challenges, the literature revealed that there is potential for these techniques to be used as alternative soil analysis methods in the laboratory and in the field. Nevertheless, their repeatability should be improved. The literature exposed that MIR spectroscopy provides more robust predictions than NIR spectroscopy; however, it is more expensive than NIR spectroscopy and requires more sample preparations. Skewed distributions mostly associated with soil chemicals, puts into question the applicability of RPD and its thresholds for evaluating quantitative prediction models for soil studies. An alternative index, namely the RPIQ is proposed and is still being investigated. However, RPD remains one of the main indicators for validating quantitative predictive models. The literature showed that bias is part of SEP or RMSE and cannot be reduced by averaging. But, composite sampling (making replications) can help reduce calibration SEP.

It was exposed in the literature that DEMs are useful to delineate hydrological parameters and land components. The potential use of DEMs for mapping soils properties was evident in the literature. The use of land components delineated by object based image analysis from DEM's has good potential for mapping saline soils. The integration of land components and spectral reflectance of vegetation/crops to study saline soils as soil properties are theoretically uniform within land components is promising.

Finally, the literature revealed that there is a connection of soil salinity with terrain attributes. The literature exposed that wet areas, low lying points and breaks of slope are vulnerable to soil salinity as they are liable to groundwater discharge. Clearly, investigations on the value of using terrain attributes to identify saline prone areas should be conducted.

CHAPTER 3

MODEL-BASED HYPERSPECTRAL QUANTIFICATION OF SALT-AFFECTED SOILS: A SOUTH AFRICAN CASE STUDY[1]

3.1 INTRODUCTION

South Africa is a vast country. Millions of South African rands have been invested in building large irrigation infrastructure. Soil salinity often builds up in these schemes due to incorrect management practices. It is very difficult to monitor salinization in these schemes because current monitoring methods are ground based and the costs of laboratory analysis are high.

Remote sensing is an attractive alternative to ground-based methods due to its relatively low costs and the ability to rapidly provide spatial information covering large areas. The use of remote sensing for soil salinity monitoring in South Africa is, however, not well established. Little is known about how South African conditions influence the spectra of salt-affected soils.

Soil salinization is a world-wide land degradation process that occurs in arid and semi-arid regions. Salts accumulate in the soil due to natural or man-made processes, *e.g.*, irrigation. Although statistics about the extent of salt-affected soils differ according to authors, Szabolcs (1994) and Metternicht and Zinck (2003) agree that about 1 billion hectares of land in the world are affected by salts. According to Nell (2009), nearly 60% of soils in South African are non-

---

[1]This work was published in the Pedosphere Journal and the Water Research Commission Report: Mashimbye ZE, Cho MA, Nell JP, de Clercq JP, van Niekerk A, Turner DP. 2012. Model-based integrated methods for quantitative estimation of soil salinity from hyperspectral remote sensing data: a case study of selected South African soils. *Pedosphere* 22: 640-649. de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013. The work was also presented at the Combined Congress, 17 – 21 January 2011, Pretoria, South Africa.

saline, 23% slightly saline, 5.1% saline, 1.4% moderately saline, 0.4% strongly saline, 3.8% saline-sodic (non-alkaline), 6.3% saline-sodic (alkaline), and 0.4% can be considered as sodic. Nell (2009) used analytical and morphological data derived from soil survey reports, environmental planning and the land type database (LTD) survey undertaken by the Agricultural Research Council-Institute for Soil, Climate and Water (ARC-ISCW) to quantify primary salinity status of South Africa. He then used elementary statistical techniques to identify relationships between the soil, water, climate, topography, vegetation, and salt parameters. Despite the awareness of the negative effects that excess salts in the soil have on agricultural yields, it is reported that the problem is increasing rather than decreasing (Szabolcs, 1994; Metternicht and Zinck, 2003).

According to Metternicht and Zinck (2003), a variety of remote sensing data, *e.g.*, aerial photographs, video images, infrared thermography, visible and infrared multispectral and microwave images, have been used for identifying and monitoring salt-affected soils. Hitherto, broadband remote sensing data have been generally used for monitoring salt-affected soils (Rao *et al*., 1991; Dwivedi, 1992; Verma *et al*., 1994; Sharma and Bhargarva, 1988; Mashimbye, 2005). However, because of their low spectral resolution and the use of conventional classification methods, these multispectral sensors (*e.g.*, SPOT, Landsat MSS, and Landsat ETM+) are reported to have limited value for studying soil properties (Dehaan and Taylor, 2003; Tamas and Lenart, 2006). Notwithstanding, these sensors have been successful in distinguishing severely salt-affected from non-affected soils (Farifteh *et al*., 2006; Weng *et al*., 2010).

Imaging spectroscopy (hyperspectral remote sensing) does provide near-laboratory quality reflectance spectra for each pixel. According to Bertel *et al.* (2006), each picture element contains a unique spectrum which can be used for detecting earth's surface materials. Hyperspectral remote sensing allows the discrimination of subtle differences between materials, permitting investigation of phenomena and concepts that greatly extend the scope of traditional remote sensing (Chang, 2003; Lillesand *et al*., 2004; Campbell, 2007). This is achievable because of the contiguous nature of the spectral profile of a hyperspectral signature.

Hyperspectral remote sensing has been widely used to study salt-affected soils (Ben-Dor and Banin, 1994; Drake, 1995; Ben-Dor *et al*., 2002; Dehaan and Taylor, 2002, 2003; Tamas and Lenart, 2006; Farifteh, 2007). Al-Khaier (2003) achieved an accurate (R^2 = 0.86) detection of

soil salinity by a normalized salinity index in bare agricultural soils using ASTER bands 4 (near-infrared) and 5 (short-wave infrared). Additionally, Khan *et al.* (2005) successfully used a normalized difference salinity index (NDSI) (using the near-infrared and red bands of the Indian Remote Sensing LISS-II sensor) to map soil salinity. No studies that used a hyperspectral NDSI to map soil salinity could be found.

Weng *et al.* (2010) developed a univariate regression model to estimate soil salt content using a soil salinity index. The index was constructed from continuum-removed reflectance at 2052 and 2203 nm. Their model was applied to a Hyperion reflectance image and was successfully validated (R^2= 0.627). Farifteh *et al.* (2007) used partial least squares regression (PLSR) and obtained prediction R^2 values between 0.78 and 0.98 using experimental soil sample data, which in each sample was treated with different salts (namely, $MgSO_4$, KCl, $NaCl$, and $MgCl_2$). Viscarra Rossel (2007) showed that bagging PLSR predictive models provided more robust predictions of organic carbon than PLSR predictive models alone.

The aim of this study was to evaluate the utility of hyperspectral remote sensing data for predicting soil salinity. Hyperspectral individual bands, an NDSI, PLSR, and bagging PLSR were investigated. A NDSI is a soil salinity index developed according to the principle of the normalized difference vegetation index (NDVI) commonly used in vegetation studies. Hyperspectral remote sensing bands located at 1410 and 2040 nm were used to develop the NDSI. To the best of our knowledge, a NDSI developed using hyperspectral data has never been tested for soil salinity prediction. Predictive models were developed using a training dataset. An independent validation dataset which was not included in the training was used to validate the models. In addition, field-based regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were developed using soil samples which were collected from within the Sandspruit catchment and the nearby Langgewens Experimental Farm near Cape Town in the Western Cape Province of South Africa.

3.2 MATERIALS AND METHODS

3.2.1 Soil samples

Two South African soil databases, namely, the LTD and *ad hoc* data held by the Agricultural Research Council (ARC), were used as sources for establishing a suitable set of soil samples for

this study. The LTD arose from the 1:250 000 scale soil mapping program, carried out over a period of 30 years (1972 - 2002) by the ARC-ISCW. From the early 1990s this information was systematically transferred to a geographical information system (GIS), along with the composition of each of the more than 7000 unique land type mapping units, as well as a supporting database containing the soil profile information. The ARC soil samples selected were collected on a monthly or bi-monthly basis over a period of 14 years from fixed sites southeast of Johannesburg in the Gauteng Province. More information about soils of South Africa can be found on the Agricultural Geo-Referenced Information Systems (AGIS) website at http://www.agis.agric.za. An additional 23 top soil samples were collected from within the Sandspruit catchment near Cape Town in the Western Cape Province of South Africa. These samples were not ground and sieved to simulate field conditions. The samples were used to develop regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4.

Most salts in South Africa are of sea origin imbedded in the geology. The LTD soil samples used in this study were from the following geological formations: Adelaide, Beaufort, Barbeton, Bokkeveld, Bushmanland, Drakensberg, Dwyka, Ecca, Kalahari, Meinhardskraal, Nama, Soutpansberg, Table Mountain, Tarkastad, and Zululand (Figure 3.11a). Natural organic carbon of the soil samples ranged from 0.01 to 0.28 g kg^{-1}. The distribution of the samples with different quantities of natural organic carbon is depicted in Figure 3.1b. The soils were found to be saline sodic, moderately saline, non-alkaline sodic, and slightly alkaline soils (Figure 3.1c). The soil and terrain digital database (SOTER) soil units covered by the samples are: A4, AR, C1, E1, G1, and H1 (Figure 3.1d).
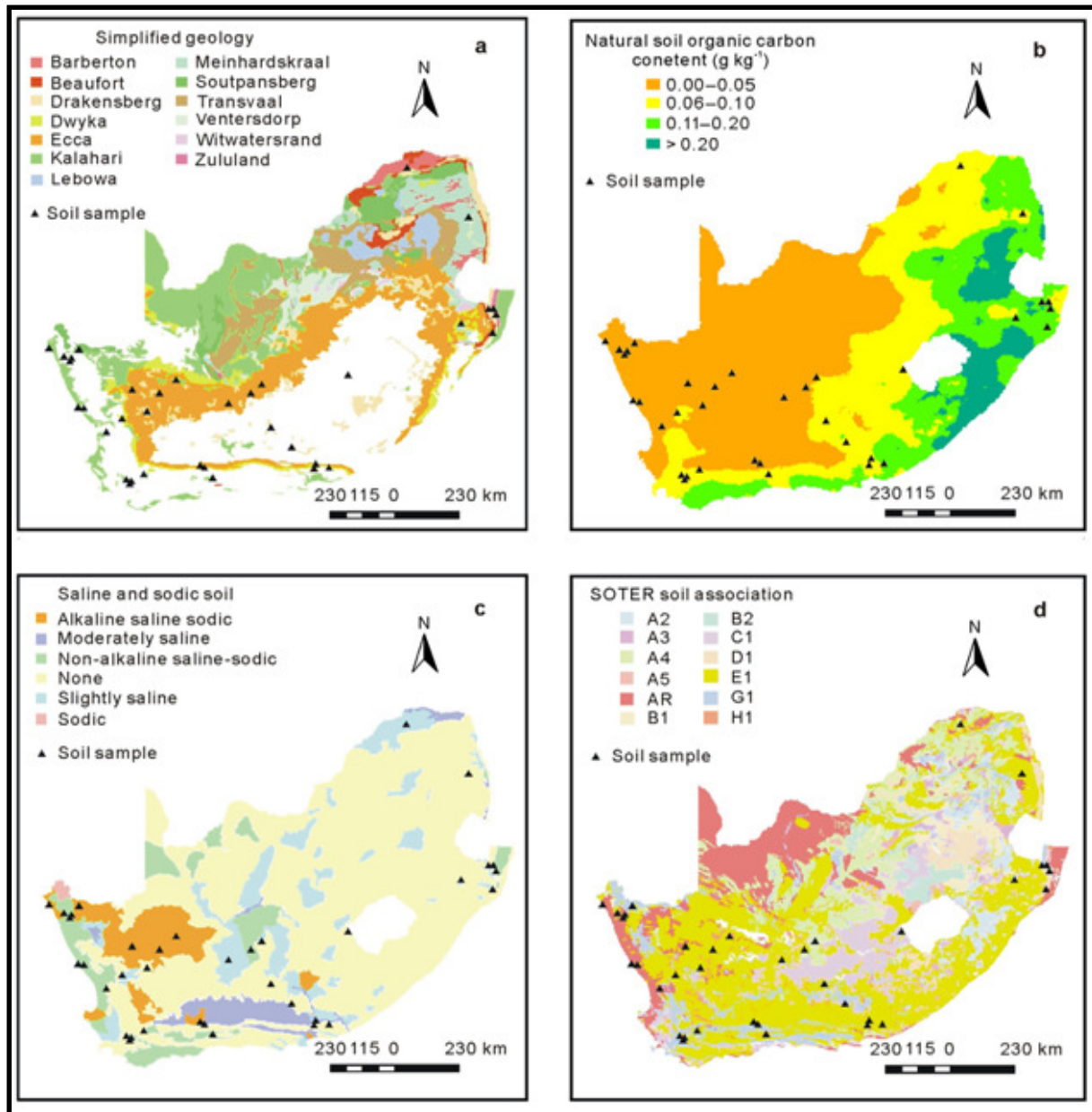
Figure 3.1 Simplified geology (a), natural organic carbon content (b) saline and sodic soils (c) and SOTER soil classification (d) in South Africa

Source: AGIS (http://www.agis.agric.za)

In total, 118 soil samples were used for this investigation. The samples were selected from the two databases using a stratified random sampling technique (Brus and Gruijter, 1997; Christofides, 2003; Kim *et al*., 2007) to ensure an even distribution within the five saline classes; non-saline (0 - 200 mS m^{-1}), slightly saline (200 - 400 mS m^{-1}), moderately saline (400 - 800 mS

m^{-1}), strongly saline (800 - 1600 mS m^{-1}), and extremely saline (> 1600 mS m^{-1}). Additional soil samples (n = 23) were collected from the Sandspruit catchment and the Langgewens Experimental Farm. All the soil samples but the 23 collected from the Sandspruit catchment and adjoining Langgewens Experimental Farm were dried, ground, and put through a 2-mm sieve to remove large particles and plant remains. The Sandspruit and the nearby Langgewens Experimental Farm samples were kept in the original form from the field, i.e. they were not ground and sieved. The samples were analysed for electrical conductivity (EC), organic carbon, texture, pH, soluble Ca, Mg, Na, Cl and SO$_4$. EC was measured by a 1:5 saturated extract.

3.2.2 Spectral data collection

Firstly, an analytical spectral device (ASD) FieldSpec spectrometer was used to acquire spectral signatures of the soil samples in a darkroom to ensure stable atmospheric and uniform illumination conditions. The instrument covers the visible to short-wave infrared wavelength range (350 - 2500 nm). The spectrometer has a sampling interval of 1.4 nm for the region 350 to 1000 nm and 2 nm for the region 1000 to 2500 nm with a spectral resolution of 3 and 10 nm, respectively. Darkroom conditions were used to eliminate diffuse light conditions and to ensure that light conditions are similar to allow comparison. Diffuse lighting conditions will be considered in a separate part of the study as the influence thereof is required for calibrating the remotely sensed information. A halogen lamp (Lowel Light Pro, JCV 14.5V-50WC) was used as a source of light. The lamp was fixed at a nadir position 20 cm above the target. To prevent contamination of one sample by another, each sample was placed on a separate black plastic background before making spectral signature measurements. A sufficient amount of soil for each sample was spread on the plate to completely cover the plate's surface. The soil was flattened on top to form an even surface. Reflectance calibration was done using a white reference. The white reference is a calibrated white Spectralon with a near 100% diffuse (Lambertian) reference reflectance panel made from a sintered poly-tetra-flourethylene based material. Calibration was done before taking measurements of each of the samples. Spectral signatures were taken at a height of approximately 15 cm above the target at approximately 15º off nadir to minimize the effect of bidirectional reflectance.

Secondly, spectral signatures of the soil samples (n = 23) collected from the Sandspruit catchment and the neighbouring Langgewens Experimental Farm were measured using the ASD

spectrometer. These samples were not ground and sieved, and the spectral signatures were measured on a clear sky day to simulate field conditions. The sensor was mounted on a Laboratory Stand at nadir approximately 15 cm above the soil sample (Figure 3.2). Each sample was spread over a plate which was covered by a black plastic background and was rotated five times when the spectra were measured to minimize bidirectional reflectance effects.

Figure 3.2 Photograph showing the experimental setup

3.3 DATA ANALYSIS

Salinity models were computed using untransformed individual reflectance, first derivative individual reflectance (FDR), a NDSI, PLSR, and bagging PLSR. Individual bands were selected based on the correlograms between EC and reflectance. Regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 using soil samples that were not ground or sieved and using the sun as a source of light were computed using the Sandspruit and Langgewens soil samples. Soil reflectance data in the wavelength range between 400 and 2500 nm were used for the analysis. R^2 values were computed for each of the models.

PLSR and bagging PLSR were computed using the ParLeS version 3.1 software (Viscarra Rossel, 2007, 2008). PLSR is a method that specifies a linear relationship between a set of dependent variables, *Y*, and a set of predictor variables, *X* (Farifteh *et al*., 2007). The general idea of the PLSR is to extract the orthogonal or latent predictor variables, accounting for as much of the variation of the dependent variables as possible. The bagging PLSR is a bootstrap technique that leaves out about 37% of the data in the course of re-sampling (Viscarra Rossel, 2007, 2008). The bootstrap automatically calculates the R^2, adjusted R^2 (R^2_{adj}), root mean squares error (RMSE), mean error (ME), ratio of prediction to deviation (RPD), and standard deviation of the error distribution (SDE). The performance of each of the models was evaluated using the calibration R^2 and the validation R^2. The R^2 values indicate the strength of statistical correlation between measured and predicted values (Farifteh *et al*., 2007). Additionally, the PLSR models were evaluated using the RPD, and R^2_{adj}. The R^2_{adj} measures the proportion of the variation in the response that may be attributed to the model rather than to random error, which makes it more comparable across models with different numbers of parameters (Viscarra Rossel, 2007). The RPD measures the ratio of percentage deviation to the RMSE. RPD values of less than 1.5 indicate very poor model predictions, between 1.5 and 2.0 poor model predictions, between 2.0 and 2.5 good model predictions, and greater than 2.5 very good model predictions (Viscarra Rossel, 2007).

3.3.1 Individual bands

A distribution fitting curve using untransformed EC values revealed that the training samples were not normally distributed ($P < 0.05$, Shapiro-Wilk's W test) (Figure 3.3a). A second distribution fitting curve computed using the natural logarithmic values of EC resulted in a normal distributed ($P > 0.05$, Shapiro-Wilk's W test) sample (Figure 3.3b). The analysis was consequently conducted using the natural logarithmic values of EC. Pearson's correlation analyses of original soil spectra and FDR with EC were carried out and the bands that yielded the highest correlations with EC were identified. For individual band analysis, only bands that occur outside the major water absorption bands (1340-1480 and 1770-1970 nm) (Herold *et al*., 2004) were considered for analysis. Consequently, regression models that explained the most degree of variation of EC using spectral reflectance were computed using these bands only. A total of 95 samples were used (training samples = 63 and validation samples = 32).
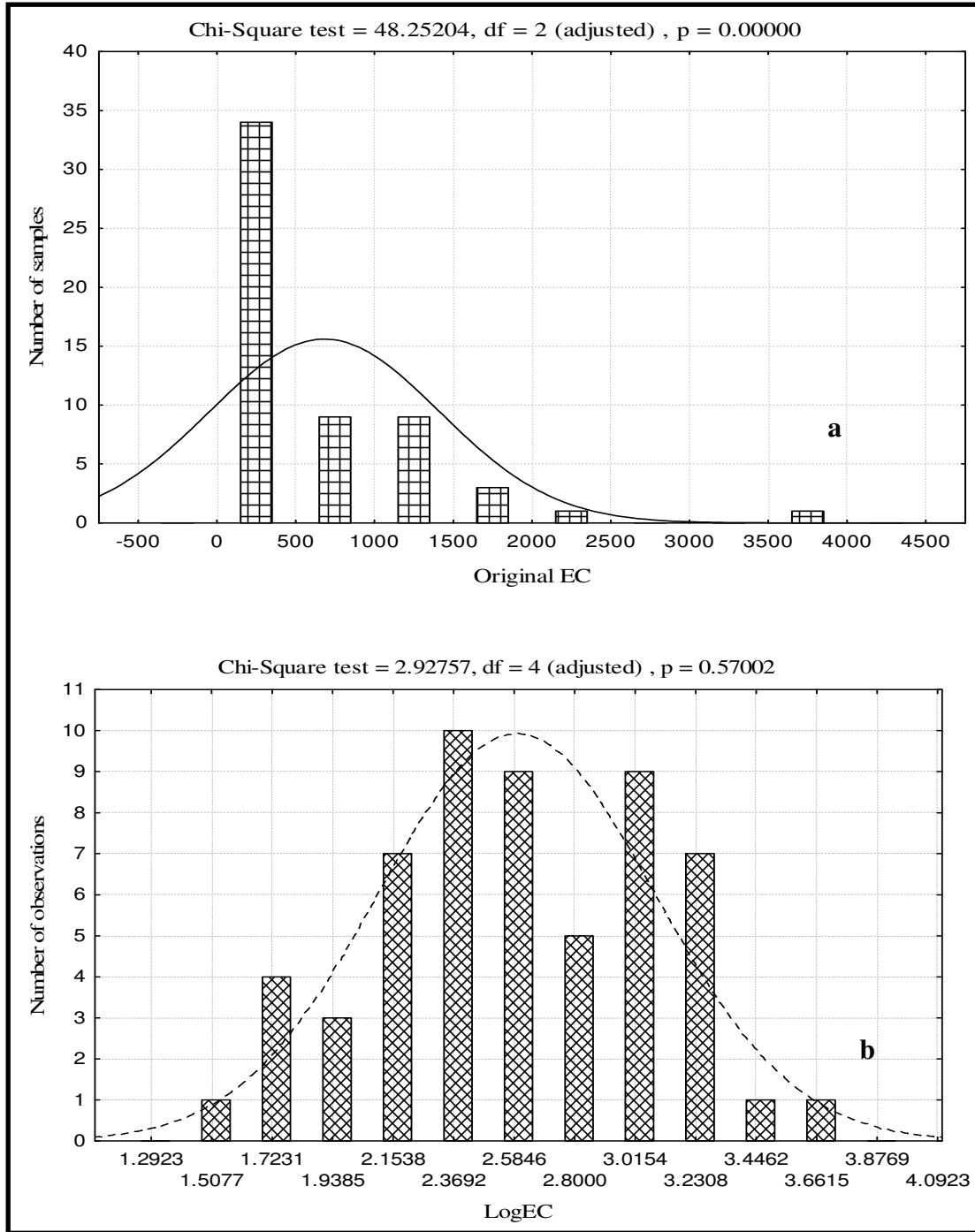
Figure 3.3 Training sample distribution-fitting curve of (a) original EC values and (b)LogEC values

3.3.2 Normalized difference salinity index (NDSI)

An analysis was carried out to develop a NDSI that predicts EC in soils. Candidate NDSI for any two bands i and j for a sample n, $NDSI_{i,j,n}$, was calculated according to the principle of the normalized difference vegetation index (NDVI) used in vegetation studies (Equation 3.1):

$$NDSI_{i,j,n} = (R_{i,n}-R_{j,n})/(R_{i,n}+R_{j,n}) \qquad\qquad 3.1$$

where $R_{i,n}$ and $R_{j,n}$ are the reflectance of any two bands i and j for a sample n.

The candidate NDSI was derived from all possible two-band combinations involving the bands in the 400-2500 nm range, sampled at 10 nm resolution. Only the training sample set was used for this purpose. This resulted in 44 100 (*i.e.*, 210 x 210) candidates. The NDSI was regressed with EC and the best bands were identified. The number of soil samples used for this analysis is 95.

3.3.3 Partial least squares regression (PLSR)

PLSR is a bilinear calibration method using data compression by reducing the large number of measured collinear spectral variables to a few non-correlated latent variables or factors (Hansen and Schjoerring, 2003; Cho *et al.*, 2007). PLSR specifies a linear relationship between a set of dependent variables (*Y*) and a set of predictor variables (*X*), thereby extracting the orthogonal or latent predictor variables accounting for as much of the variation of the dependent variables as possible (Cho *et al.*, 2007; Farifteh *et al*, 2007). The linear equation derived from the PLSR is:

$$Y = Xb + E, \qquad\qquad 3.2$$

where *Y* is the mean-centred matrix containing the response variables, *X* the mean-centred matrix containing the predictor variables (spectral bands in this study), b the matrix containing the regression coefficients and *E* is the matrix of residuals. PLSR of untransformed and first derivative reflectance with EC was conducted using the ParLeS version 3.1 software (Viscarra

Rossel, 2008). As with individual bands and a NDSI evaluated above, 95 samples were used for this analysis.

3.3.4 Bagging PLSR

Bootstrapping performs sampling within a sample. It is a technique that may be used to estimate the cumulative distribution function (CDF) of a population, its moments and their uncertainty by re-sampling with replacement (Viscarra Rossel, 2007). The bootstrap assumes that the CDF of the data is sufficiently similar to that of the original population, and that multiple realizations of the population can be replicated from a single dataset. The bagging PLSR function of the ParLes version 3.1 software was used to conduct automatic bootstrapping consisting of 50 iterations for the bagging PLSR. Although a bootstrap may have duplicate data, it leaves out about 37% of the data in the course of re-sampling for validation statistics (Viscarra Rossel, 2007). These statistics were analysed to assess the performance of the various models.

3.3.5 Field-based soil salinity regression predictive models

An investigation was conducted to evaluate the possibility of mapping soil salinity using airborne/satellite hyperspectral remote sensing data. Pearson's correlation analysis of original and transformed spectra of dry soil with EC, pH, soluble Ca, Mg, Na, Cl and SO_4 was conducted. The bands that yielded the highest *r* were used to develop regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4. Only samples (n = 23) from the Sandspruit catchment and the Langgewens Experimental Farm were used for this investigation. These samples were not ground and sieved, and the sun was used as a source of light when measuring spectral signatures to emulate field conditions.

3.4 RESULTS

3.4.1 Regression between EC and individual bands

Pearson correlation coefficient values of EC with untransformed saline soil spectra increased from the visible through to the short-wave infrared region of the spectrum (Figure 3.4).The raw reflectance data at 2257 nm and FDR at 991 nm showed the highest Pearson correlation coefficient (*r* = -0.59 for 2257 nm and *r* = -0.73 for FDR at 991 nm) with EC among the spectral bands from 400 to 2500 nm. The above bands were subsequently used to derive predictive

regression models for soil EC. Fig. 3.5a indicate that for the untransformed reflectance (at 2257 nm), a quadratic regression model provided a better representation (R^2 = 0.31) of the EC of the training sample set than a linear model (R^2 = 0.25). Despite yielding a lower calibration R^2, the linear predictive model yielded a slightly higher prediction R^2 than the quadratic predictive model (Figure 3.5b, c) compared to the validation sample set. For the FDR (at 991 nm), both the linear and quadratic models yielded similar calibration and prediction R^2 values (Figure 3.5d, e, f).
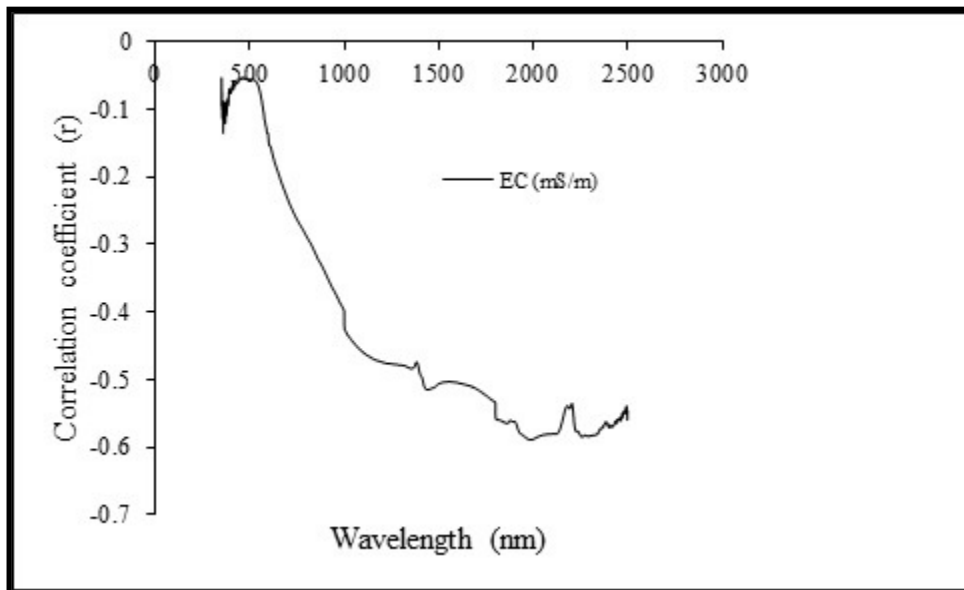
Figure 3.4 The relationship of EC with untransformed reflectance of dry saline soil
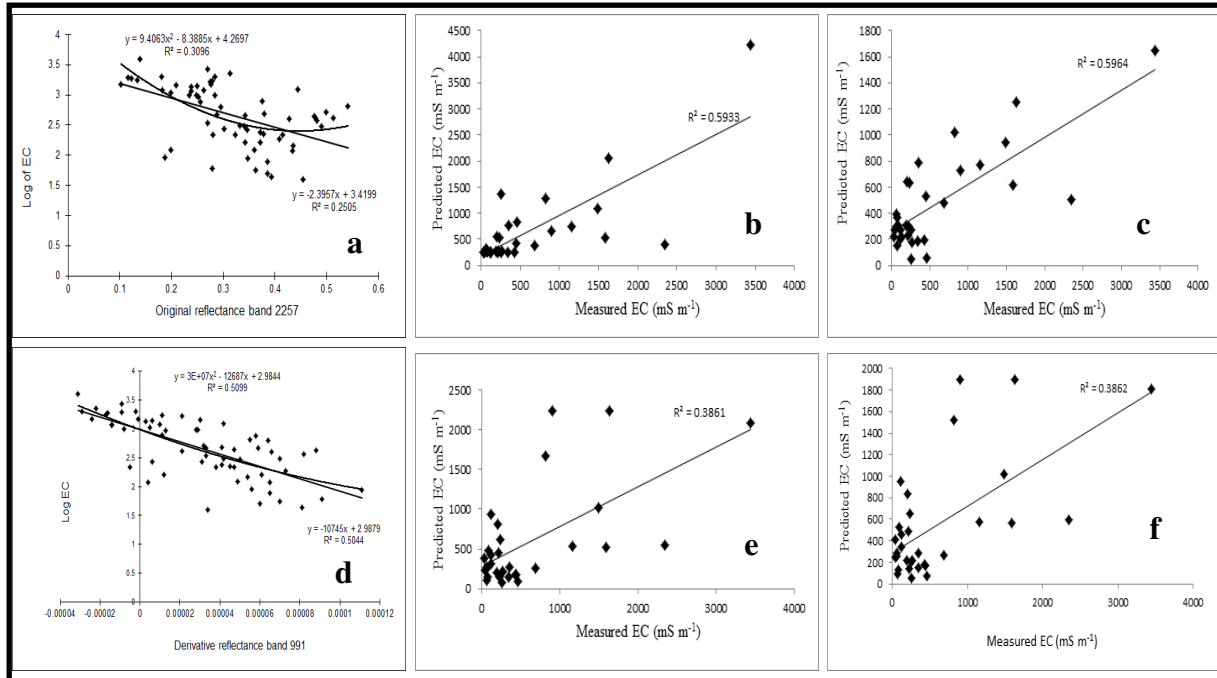
Figure 3.5 Untransformed individual band (at 2257 nm) soil EC predictive models (a), quadratic untransformed individual band soil EC predictive model validation (b), linear untransformed individual band soil EC predictive model validation (c), FDR individual band (at 991 nm) soil EC predictive models (d), quadratic FDR individual band soil EC predictive model validation (e), and linear FDR individual band soil EC predictive model validation (f)

3.4.2 NDSI

Linear regression analyses were performed comparing each candidate NDSI with EC. A contour plot of R^2 of the results is shown in Figure 3.6. The 2040 and 1410 nm wavelengths were identified as the most promising for developing a NDSI. Consequently, a NDSI using the corresponding bands was created and subsequently assessed for its predictive capability using the independent validation dataset.
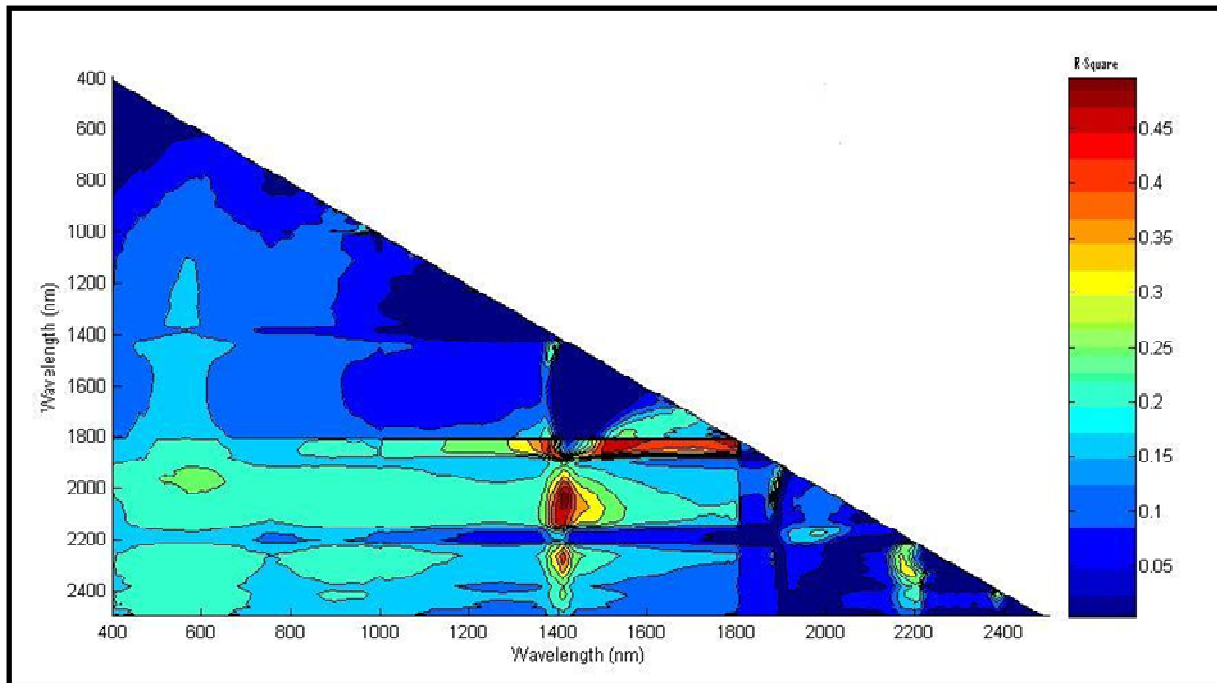
Figure 3.6 Contour plot of R^2 with wavelength (nm)

Although the NDSI quadratic and linear regression predictive models yielded similar calibration R^2 (Figure 3.7a), the NDSI quadratic predictive model yielded a higher prediction R^2 than the NDSI linear predictive model, with the prediction R^2 being 0.65 and 0.57 for the NDSI quadratic predictive model and the NDSI linear predictive model, respectively (Figure 3.7b, c). Compared to the individual band predictive models (Figure 3.5b, c, e, f), the NDSI quadratic predictive model yielded higher calibration and prediction R^2 values.
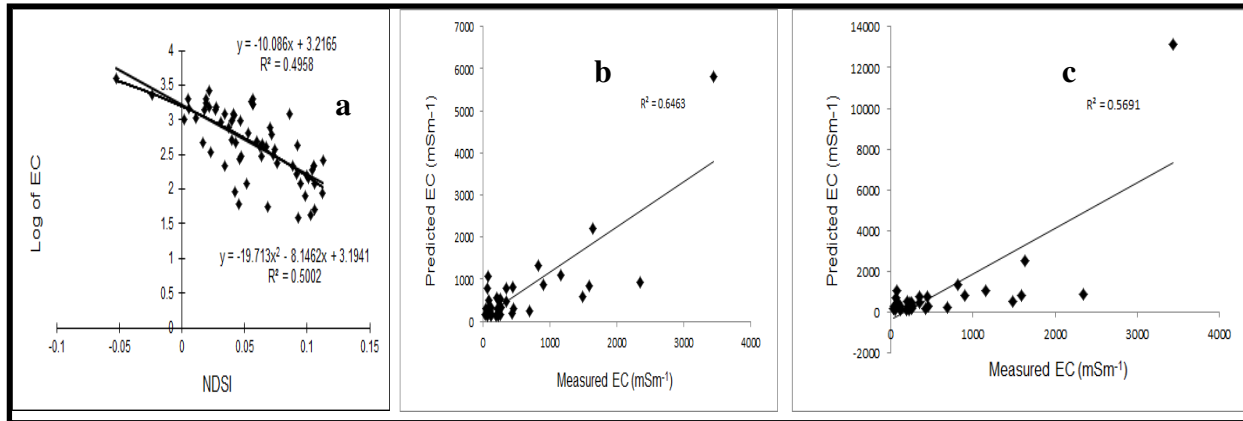
Figure 3.7 NDSI soil salinity predictive models (a), quadratic NDSI soil salinity predictive model validation (b), and linear NDSI soil salinity predictive model validation (c).

3.4.3 PLSR

The results show that the R^2 values for the untransformed and FDR PLSR predictive models were 0.68 and 0.72, respectively (Table 3.1), while the RPD values were less than 1.5 in both cases. According to Farifteh *et al*. (2007), predictive models with RPD values less than 1.5 and calibration R^2 values between 0.66 and 0.81 can be regarded as poor predictive models. In addition, the high RMSE values (0.39 and 0.41 for untransformed spectra and FDR, respectively) were indicative of high prediction errors. Although the R^2 value of the FDR PLSR predictive model was slightly higher than the untransformed reflectance value, the former yielded a significantly lower prediction R^2 (Figure 3.8a, b). The first five factors of the untransformed reflectance PLSR predictive model contained about 68% of the information on soil EC, while the first factor of the FDR PLSR predictive model contained almost 72% of the information on soil EC.

Table 3.1 PLSR soil salinity predictive models calibration statistics

Statistics[a]	Untransformed reflectance	First derivative reflectance
R^2	0.68	0.72
R^2 adj.	0.47	0.41
RMSE	0.39	0.41
RPD	1.35	1.27
Number of factors	5	1

[a] R^2_{adj} = adjusted R^2; RMSE = root mean square error; RPD = ratio of prediction to deviation.
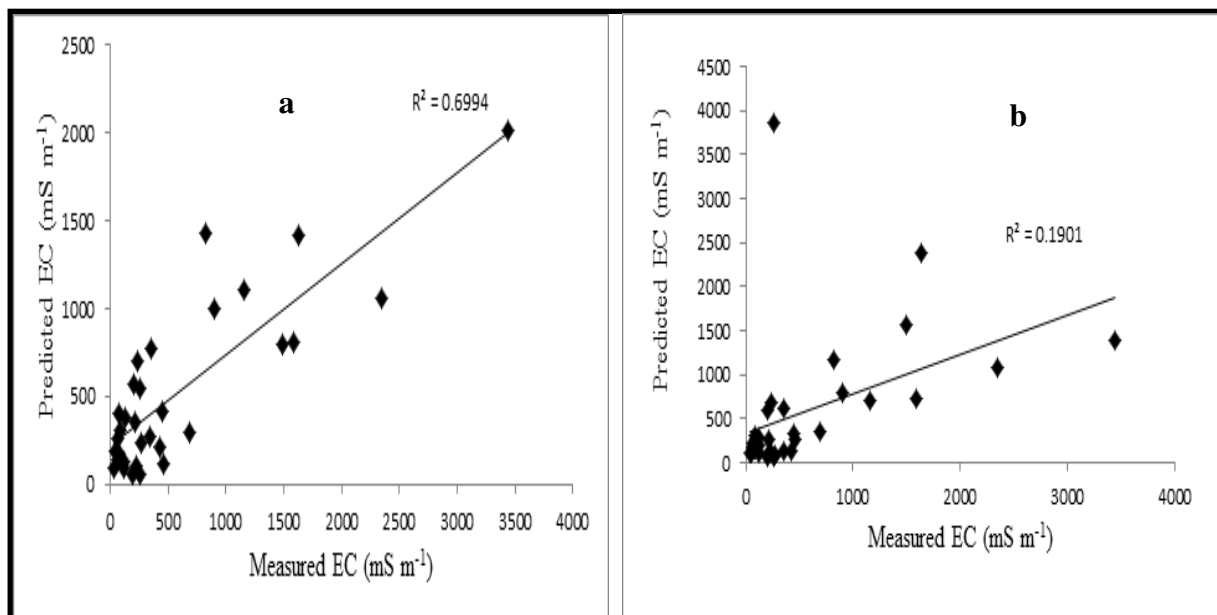
Figure 3.8 Untransformed spectra PLSR soil salinity predictive model validation (a) and the FDR PLSR soil salinity predictive model validation (b).

3.4.4 Bagging PLSR

As with PLSR, the calibration R^2 values were between 0.66 and 0.81 (R^2 = 0.69 for untransformed reflectance and R^2 = 0.67 for derivative reflectance). However, the RPD values are higher than 1.5 (Table 3.2). Additionally, the bagging PLSR presented lower prediction errors when compared to PLSR (Tables 3.1 and 3.2). Amongst all the predictive models evaluated in this work, the bagging PLSR model using FDR yielded the highest prediction R^2 (Figure 3.9b).

Table 3.2 Bagging PLSR soil salinity predictive models calibration statistics

Statistics[a]	Untransformed reflectance	First derivative reflectance
R^2	0.69	0.67
R^2 adj.	0.69	0.66
RMSE	0.29	0.29
RPD	1.81	1.73
Number of factors	8	2
Number of bootstraps	50	50

[a]R^2_{adj} = adjusted R^2; RMSE = root mean square error; RPD = ratio of prediction to deviation.
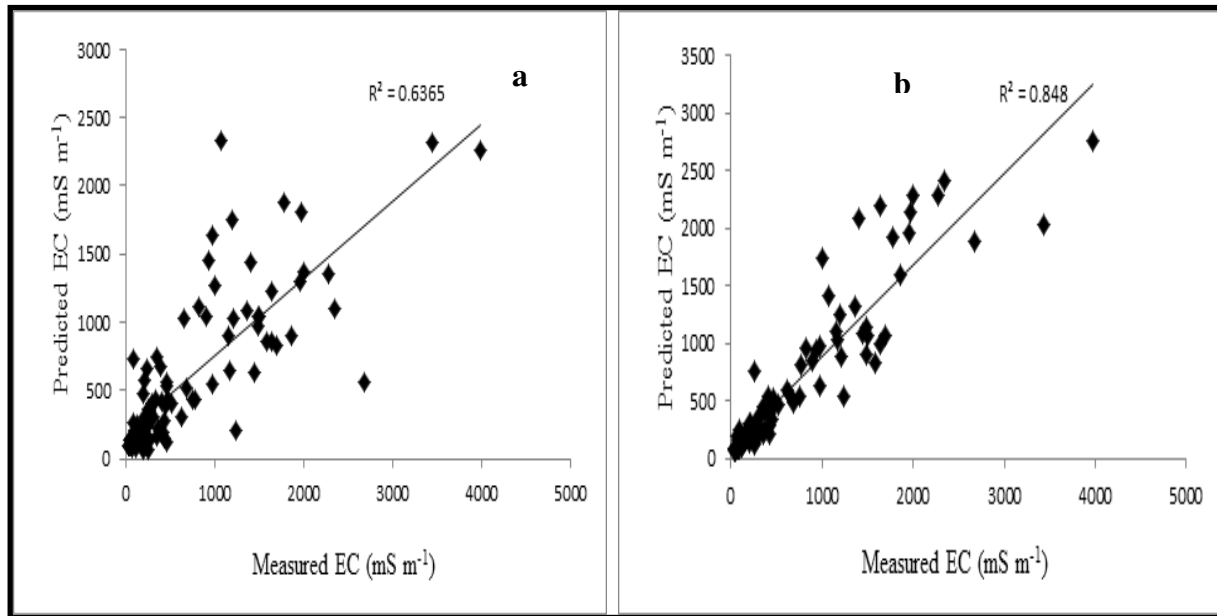
Figure 3.9 Untransformed spectra bagging PLSR soil salinity predictive model validation (a) and FDR bagging PLSR soil salinity predictive model validation (b)

4.4.5 Field-based soil salinity regression predictive models

The relationship of soil spectra with EC is shown in Figure 3.10. Spectral interferences due to atmospheric moisture are observed around 1800 nm (Figure 3.10). First derivative reflectance (FDR) yielded higher r values for all the soil chemicals investigated in this study (Table 3.3). The relationships of EC, pH, soluble Ca, Mg, Na, Cl and SO$_4$ with FDR are shown in Figure 3.11. Owing to higher r values, regression predictive models were computed with FDR for all the soil chemicals investigated in this study.
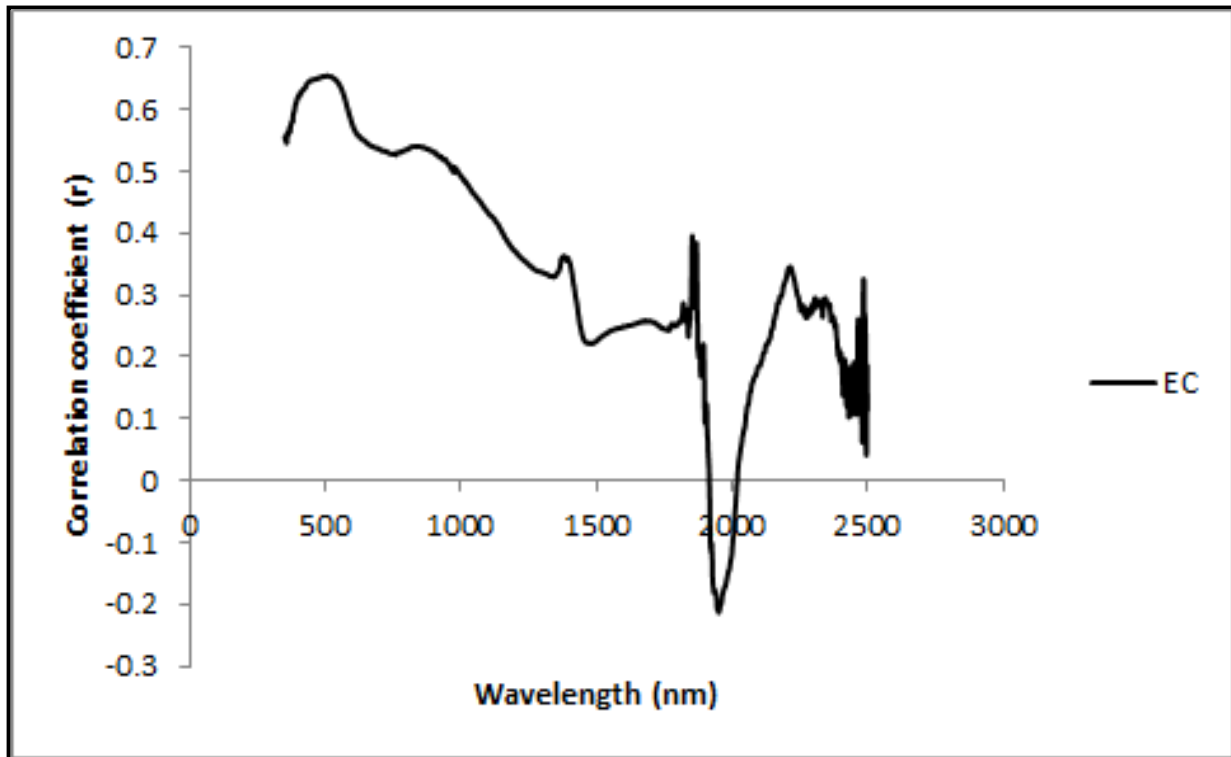
Figure 3.10 The relationship of EC with wavelength based on field samples

Table 3.3 Field-based correlation coefficients for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 based on untransformed and first derivative reflectance

Variable	Highest r (untransformed band where it occurs)	Highest r (FDR band where it occurs)
EC	0.66 (506 nm)	0.88 (2051 nm)
pH	0.43 (441 nm)	0.71 (2205 nm)
Ca	-0.42 (1950 nm)	0.72 (2048 nm)
Mg	0.57 (528 nm)	0.83 (2049 nm)
Na	0.67 (494 nm)	0.83 (2051 nm)
Cl	0.68 (494 nm)	0.84 (2051 nm)
SO_4	0.44 (1950 nm)	0.67 (2048 nm)
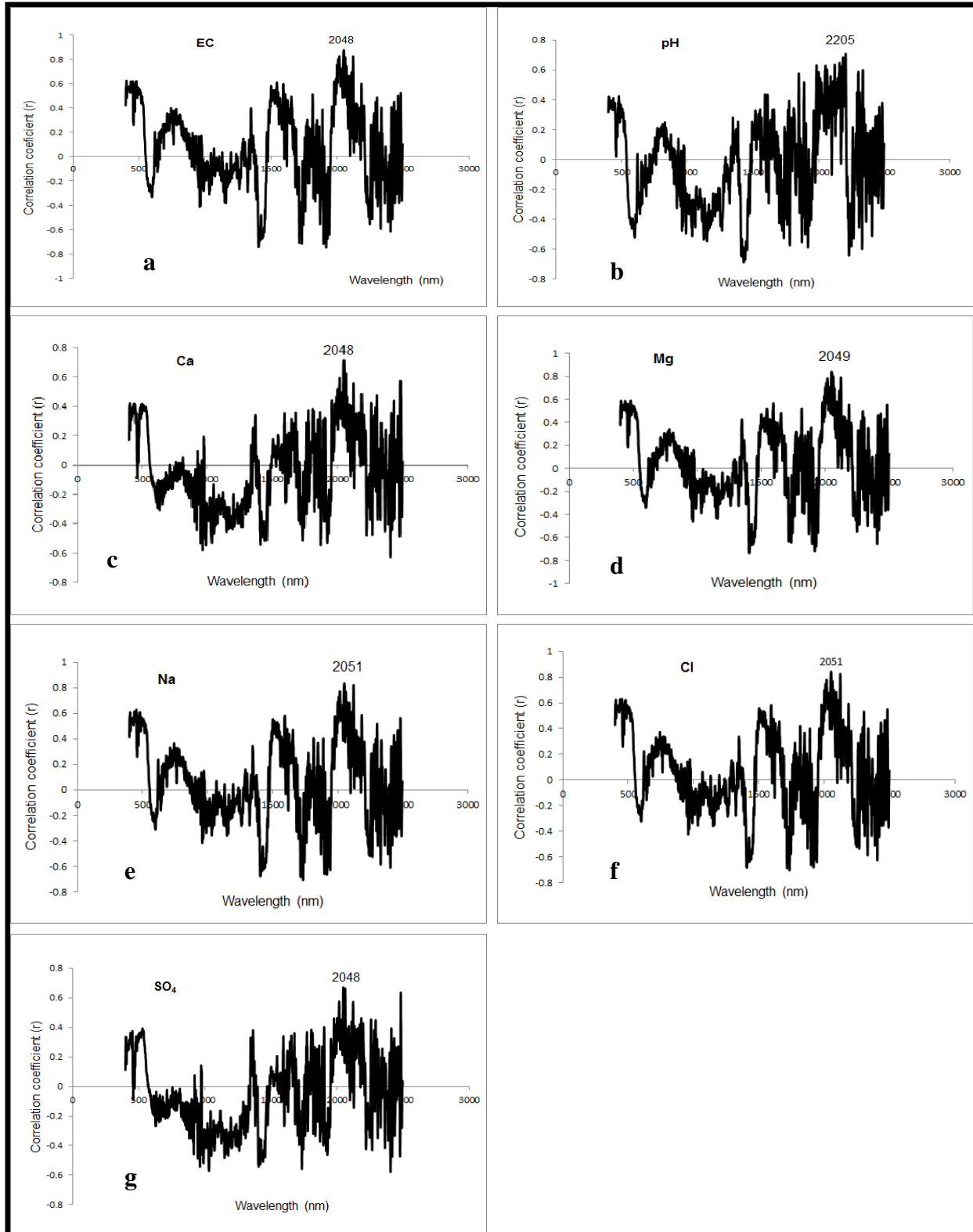
Figure 3.11 The relationship of (a) EC, (b) pH, soluble (c) Ca, (d) Mg, (e) Na, (f) Cl and (g) SO4 with first derivative reflectance showing the band that yielded the highest *r*

Regression predictive models using FDR for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 are depicted in Figure 3.12. Moderate to accurate predictive models for all the soil chemicals investigated were obtained in this investigation. The R^2 values are 0.85, 0.84, 0.81, 0.79, 0.65, 0.58 and 0.50 for the EC, Mg, Cl, Na, Ca, SO_4 and pH regression predictive models respectively. All the soils chemicals but pH could be predicted using a quadratic equation (Figure 3.12). A linear model was found suitable for predicting soil pH.
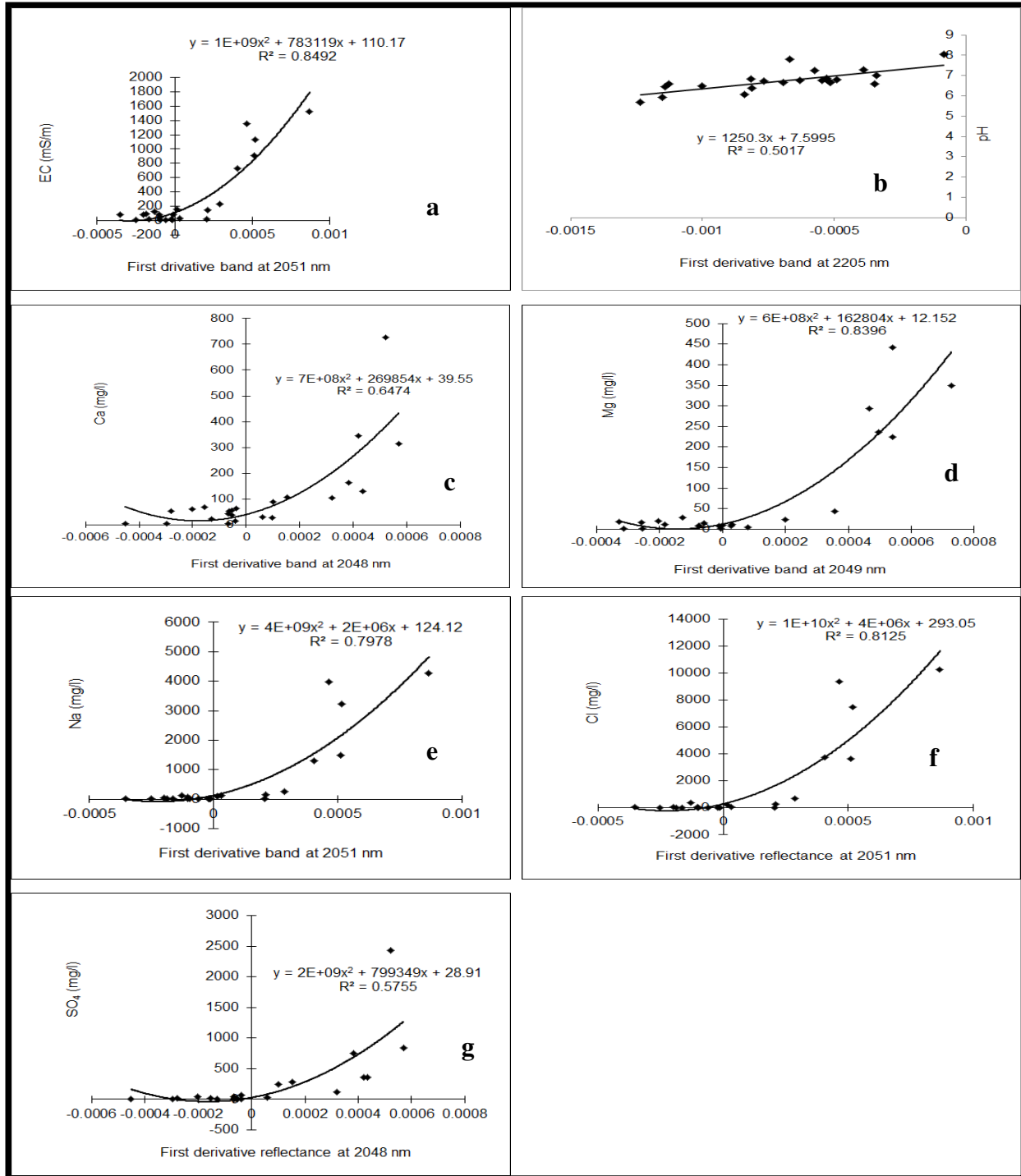
Figure 3.12 Field-based regression predictive models for (a) EC, (b) pH, soluble (c) Ca, (d) Mg, (e) Na, (f) Cl and (g) SO_4 based on first derivative reflectance

The relationship of untransformed reflectance with EC, Na, and Cl looked similar throughout the spectrum (Figure 3.13). A similar trend was observed with FDR (Figure 3.11). Untransformed reflectance displayed a stronger relationship with EC, Na and Cl in the visible and near-infrared regions of the spectrum. The relationship was weaker in the short wave infrared (SWIR) region. On the other hand, FDR yielded a very strong relationship with EC, Na and Cl in the SWIR. The highest *r* of FDR with EC, Na, and Cl occurred at FDR band at 2051 nm.
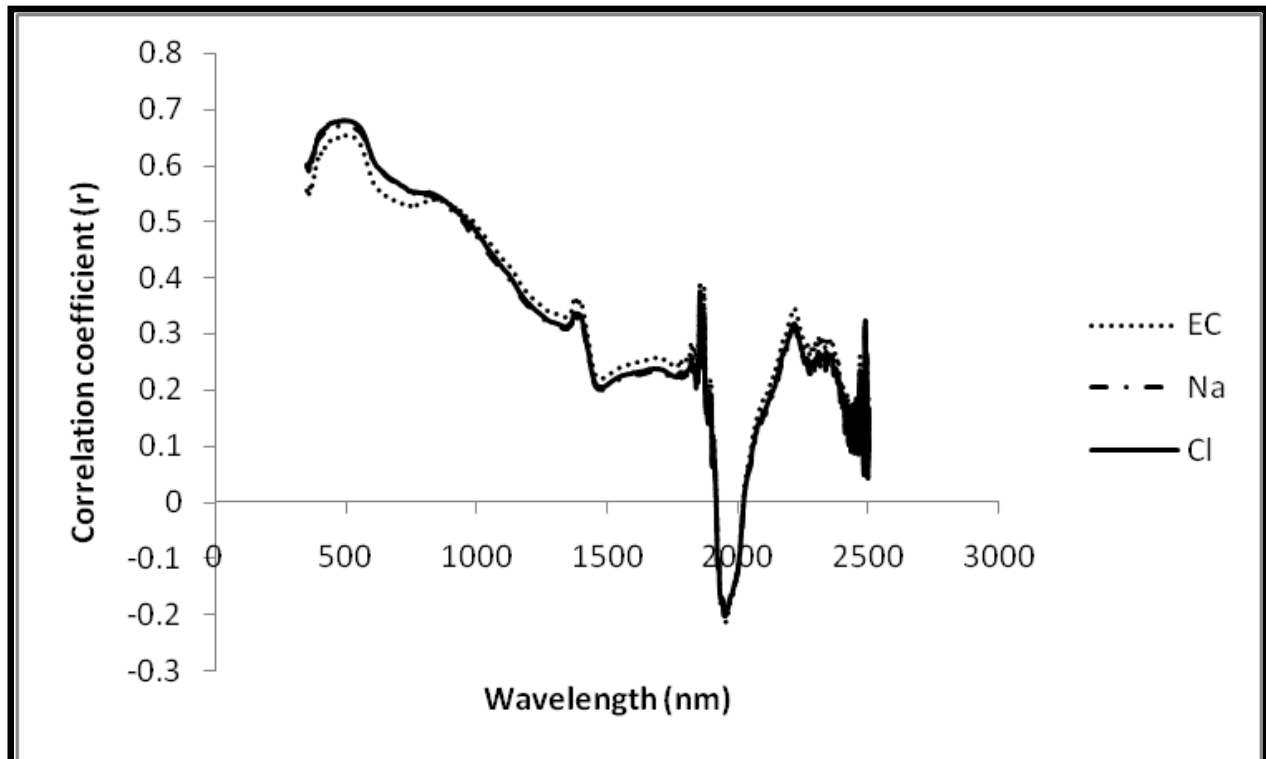
Figure 3.13 The relationship of untransformed reflectance with EC, soluble Na and Cl

3.5 DISCUSSION

This work confirmed the utility of bagging PLSR predictive models in soil studies. Bagging PLSR predictive models produced higher prediction R^2 than PLSR, NDSI and individual bands predictive models. These results support the findings by Viscarra Rossel (2007) which showed that bagging PLSR predictive models provided more robust predictions of organic carbon than PLSR predictive models alone. This is because bagging PLSR incorporates a bootstrap sampling into the construction of the model, which stabilizes the modelling while still allowing for the identification of important relationships in the data (Viscarra Rossel, 2007).

We found that the PLSR predictive models performed better than using a two-band NDSI or individual bands. This is likely due to higher information content of the multiple bands used in PLSR. The prediction R^2 obtained in this research is generally lower than that of Farifteh *et al.* (2007). Presumably, this is because the EC of the samples used in this work is the combined effect of a number of naturally occurring soluble salts. The results of untransformed PLSR analysis suggest that PLSR can provide useful estimates of soil EC.

The NSDI predictive models could explain up to about 50% of the variation in soil EC. Normalized indices have been found to be useful in many studies, including vegetation studies (Gao, 1996; Al-Khaier, 2003; Jins and Sader, 2004; Delbart *et al*., 2005; Khan *et al*., 2005; George *et al*., 2006; Cho *et al*., 2007; Inoue *et al*., 2007). One of the bands (located at 1410 nm) used to compute the NDSI is found in the water absorption region. Consequently, the NDSI can only be applied to dry soils. No studies that linked the other band used in the NDSI (band at 2040 nm) to soil EC were found.

The relationship of EC with the saline soil spectra increased from the visible through to the short wave infrared (SWIR) portion of the spectrum. The highest correlations of untransformed saline soil spectra with EC occurred in the near-infrared (NIR) and SWIR. This is likely due to saline soils having distinct spectral features in the visible and near-infrared portions of the spectrum, which allow the recognition of minerals such as gypsum, bassanite and polyhalite (Dehaan and Taylor, 2002; Metternicht and Zinck, 2003). Also, Farifteh *et al. (*2007*)* found that the best performing bands for field-scale data sets, experimental-scale and image-scale data sets were in the NIR and SWIR regions of the spectrum. The untransformed individual band located at 2257 nm presents possibilities for estimating soil EC for dry soils by a linear predictive model. No other studies were found linking the band at 2257 nm to soil EC.

Because the study was based on dry soils, the influence of water on the soil spectra would be minimal. According to Weng *et al.* (2010, citing Baumgardner *et al.,* 1970), organic carbon content hardly affects the reflectance of soil when it is lower than 20 g kg^{-1}. Hence, organic carbon could not have affected the spectral reflectance of soils in this study because the highest measured organic carbon for the samples was 0.28 g kg^{-1}, while the average organic carbon content was 0.03 g kg^{-1}.

Regarding field-based predictive models, we found that reliable predictions of EC, pH, soluble Ca, Mg, Na, Cl and SO_4 could be made based on FDR. Untransformed reflectance yielded weaker relationships with all the soil chemicals investigated in this work. This is perhaps due to the different particle sizes of the soil samples. The soil samples were not ground and sieved to emulate field conditions. FDR transformation seems to minimize the effect of scattering due to particle size. The best performing bands for field-based predictions of EC, pH, soluble Ca, Mg, Na, Cl, and SO_4 are in the SWIR. The bands that yielded the highest r for EC, pH, soluble Ca, Mg, Na, Cl and SO4 are 2051, 2205, 2048, 2049, 2051 and 2048, respectively. This is consistent with the work of Farifteh *et al*. (2007) who found that best performing bands for salt-affected soils are in the NIR and SWIR.

The techniques applied in this study have not been tested on digital hyperspectral airborne or satellite images. They will be tested using digital airborne hyperspectral data in a selected study site where further work is currently being conducted. Constraints such as atmospheric attenuation are envisaged when airborne or satellite images are used (Ben-Dor *et al*., 2009). Good atmospheric correction methods will have to be used. Other challenges include soil texture and bidirectional reflectance distribution effects. Additionally, the soil surface is not always fully exposed. Litter, vegetation cover and remains, rocky outcrops and other surface features might contribute to creating spectral confusion with salt reflectance properties (Metternicht and Zinck, 2003; Ben-Dor *et al*., 2009). Further research is needed to verify the findings of this work.

3.6 CONCLUSIONS

Under controlled environment, this work suggest that individual bands, a NDSI, PLSR and bagging PLSR present opportunities for mapping salinity during dry seasons. The study also affirmed that bagging PLSR produces more robust predictive models than PLSR alone. Of all the techniques evaluated under controlled conditions, bagging PLSR using FDR is the most effective method of predicting soil EC. In addition, a NDSI and the untransformed band at 2257 nm can potentially predict soil EC under dry conditions. These techniques present possible solutions for estimating soil EC using remotely-sensed imagery during dry seasons. The work also revealed that soil EC can be explained by a linear predictive spectral model. Furthermore, this work presents opportunities for estimating EC using hand-held spectrometers in the laboratory (where minimum soil preparation will be required) and in-situ. With respect to field-based investigation,

the study established that reliable predictions of EC, pH, soluble Ca, Mg, Na, Cl and SO_4 are possible based on FDR. This study confirmed that spectral transformations could minimize spectral interferences such as particle size. More research is needed to evaluate these techniques under field conditions, different soil types and different geological conditions.

CHAPTER 4

NEAR-INFRARED SPECTROSCCOPY AND CHEMOMETRIC MODELLING OF SALT-AFFECTED SOILS[2]

4.1 INTRODUCTION

South Africa being a semi-arid country, monitoring soil salinity is crucial for natural resources management. Salinization of soils is a worldwide problem that negatively affects the productivity of soils, thus leading to the reduction of agricultural yields. Saline soils are characterized by elevated levels of electrical conductivity (EC) which affects the productivity of soils. This problem is common in soils of arid and semi-arid regions of the world (Fitzpatrick, 1993; Rowel, 1994). While incorrect irrigation is a major contributor to soil salinization (Lenny *et al*., 1996; Katerji *et al*., 1996; Slavisch *et al*., 2002; Utset and Borroto, 2001), the removal of deep-rooted plants also contributes to soil salinization in areas where the water table is closer to the surface (Allison *et al*., 1990; Dowling *et al*.; 2003; Xu and Shao, 1990). In the arid and semi-arid regions of the world, including South Africa, soil salinity often builds up as a result of poor irrigation methods and the clearing of vegetation for crop production. It is estimated that about 1 billion hectares of land in the world are affected by salts (Metternicht and Zinck, 2003; Szabolcs, 1994). Despite the awareness of the adverse effects that excess salts in the soil have on agricultural yields, it is reported that the problem is increasing rather than decreasing (Greiner,

---

[2] The contents of this chapter was published in the Water Research Commission Report: de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013. The work was also presented at the 2nd South African Chemometrix Society Conference, 07 – 09 May 2012, Irene, South Africa. The work is being prepared for submission for publication in a suitable peer-reviewed scientific journal.

1998; Metternicht and Zinck, 2003; Bennet *et al*., 2009). For instance, Gao *et al*. (2011) investigated soil salinization temporal and spatial dynamic changes in the upper stream of the Tarim River in China using remote sensing, global positioning systems and other data sets. Their study established that the total area of salinized land had increased. Also, Fey and de Clercq (2004) found credible evidence of dryland salinity in the Berg River catchment in the Western Cape Province of South Africa.

South Africa would benefit from the development of less labour-intensive, more cost-effective and reliable methods for soil salinity monitoring, because conventional wet chemistry methods are tedious and expensive as they require the use of chemicals (Ben-Dor and Banin, 1995; Viscarra Rossel and McBratney, 1998; Shepherd and Walsh, 2002; Islam *et al*., 2003; Viscarra Rossel *et al*., 2009). Speedy assessment of soil salinity is essential for management purposes (de Clercq *et al*., 2009). This will allow early diagnosis of salt accumulation in the soil so that reclamation strategies can be implemented when the problem is still manageable. Furthermore, Shepherd and Walsh (2002) noted that soil-testing laboratories in Africa are closing at the time when they should be gearing to meet the challenges of agricultural development. They proposed a diagnostic surveillance framework modelled on the medical diagnostic approaches for evidence-based management of agriculture and environment. Non-destructive near-infrared (NIR) (14 286 – 4000 cm^{-1}) and mid-infrared (MIR) (4000 – 400 cm^{-1}) spectroscopic measurements are among the economical and user-friendly substitute or complementary methods for soil chemical analysis (Brown *et al*., 2006; Viscarra Rossel *et al*., 2006; Gomez *et al*., 2008; Mulder *et al*., 2011). Bellon-Maurel and McBratney (2011) undertook a review of research on NIR/MIR spectroscopy for soil studies, particularly for determining carbon (C) content. Their objective was to determine which acquisition method (NIR, MIR, in the field or in the laboratory) might be recommended to do C reserves measurement for carbon credit trading. Their study established that MIR spectroscopy performs better than NIR spectroscopy for C measurement. Mashimbye *et al*. (2012) investigated hyperspectral individual bands, a normalized difference salinity index (NDSI), partial least squares regression (PLSR) and bagging PLSR for predicting soil salinity. They established good soil salinity predictions based on PLSR using untransformed spectra (R^2 = 0.70), bagging PLSR using first derivative reflectance (FDR) (R^2 = 0.85), NDSI (R^2 = 0.65) and the untransformed band situated at 2257 nm (R^2 = 0.60). Also, Furthermore, Ge *et al*. (2011) reported that VNIR diffuse reflectance spectroscopy calibration

models are highly dependent on instrument and scanning environment, and that their extent of applicability could be limited. The shift towards fully exploiting the information-rich signal would be beneficial for agricultural development and will also reduce the amount of chemical waste generated by wet chemistry methods.

The adoption of NIR/MIR spectroscopy as a routine laboratory technique for soil analysis has been sluggish. This is partly due to the challenges of repeatability of NIR/MIR results because of different instruments and the scanning environment. The use of NIR/MIR spectroscopy for soil salinity analysis is limited (Janik *et al*., 1998; Shibusawa *et al*., 2001;Islam *et al*., 2003; Viscarra Rossel *et al*., 2006; Wang *et al*., 2012). Most NIR spectroscopy studies for soil have focused on carbon (Ben-Dor and Banin, 1995; Vagen *et al*., 2006; Sankey *et al*., 2008; Bellon-Maurel *et al*., 2010; Bilgili *et al*., 2010; Viscarra Rossel and Behrens, 2010; Ge *et al*., 2011; Fuentes *et al*., 2012). No studies that used NIR spectroscopy for detecting soil salinity were found in the literature. Previous studies on soil EC mainly used MIR, VIS-NIR and ultraviolet visible near-infrared (UV-VIS-NIR) spectroscopy (Janik *et al*., 1998; Shibusawa *et al*., 2001; Islam *et al*., 2003; Farifteh, 2007; Mashimbye *et al*., 2012; Weng *et al*., 2012). Owing to the correlation of pH, cations and anions with EC (which is a major indicator of soil salinity) (Farifteh, 2007; Farifteh *et al*., 2007; Farifteh *et al*., 2008; Yao and Yang, 2010), this study aimed to investigate the utility of NIR spectroscopy to quantify EC, pH, soluble Ca, Mg, Na, Cl and SO_4.

4.2 MATERIALS AND METHODS

4.2.1 Sites and soil sampling

A total of 49 soil samples were used for this study. Twenty-three (23) topsoil samples were collected in the Sandspruit River catchment ($33^o17'12"$S and $18^o46'15"$E) and the neighbouring Langgewens experimental farm north of Cape Town in South Africa. The Sandspruit catchment is the main contributor to the salinity of the water in the Berg River catchment and has a Department of Water Affairs and Forestry monitoring weir with a long-term record of salt and water discharge into the Berg River. An additional 26 samples were obtained from *ad hoc* data held by the Agricultural Research Council-Institute for Soil, Climate and Water (ARC-ISCW). The ARC-ISCW soil samples were collected for salinity monitoring on a monthly basis over a period of 14 years from fixed sites south-east of Johannesburg in Gauteng province, South

Africa. The soil samples were air-dried, ground with a mortar and passed through a 2-mm sieve to remove large stone particles and plant remains. EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were analysed in the laboratory. EC was measured by a 1:5 saturated paste extract.

4.2.2 Spectral data measurement

Near-infrared spectra were measured using a Bruker multi-purpose analyser (MPA) spectrometer (http://www.bruker.com) (wavelength range of 3595 to 12 489 cm^{-1}, equivalent to 800 to 2800 nm). In each of the reflectance measurements, 128 scans were averaged. Owing to soil being a highly heterogeneous medium, composite sampling was implemented to alleviate sampling bias-related errors (Bellon-Maurel *et al*., 2010; Esbensen and Paasch-Mortensen, 2010) and to improve the resultant PLSR calibration models. Composite sampling was implemented by spreading the soil samples in a petri dish and subsequently sampling the soil with a spatula throughout the petri dish. Three independent composite samples were extracted from each soil sample. The soil was constantly mixed with a spatula to improve sample representativeness. Each composite sample was placed in an aluminium cup to completely cover the surface of the cylinder. The base of the cylinder is made of high-quality quartz glass with a diameter of 51 mm. Spectra for the soil samples were measured in reflectance mode because soil is a very diffusive and absorptive medium (Bellon-Maurel and McBratney, 2011). The instrument was calibrated with an in-house spectrometer control solution after measuring every six soil samples to monitor the performance of the instrument (Stellenbosch University Institute for Wine Biotechnology, 2012). The spectrometer control solution is a 1-l mixture of 125 ml of absolute ethanol (12% v/v), 2.5 g tartaric acid and distilled water (pH was wet at 3.5).

4.3 DATA ANALYSIS

4.3.1 PLSR model calibration

Saline soils contain varying amounts of solubleble cations and anions which contribute to the EC of the soil. Also, the pH, cations and anions are correlated to soil EC. All the soil chemical measurements were log transformed before the analysis was conducted because of the non-Gaussian distribution of the values. Thus, the study evaluated the potential of NIR spectroscopy to predict EC, pH, soluble Ca, Mg, Na, Cl and SO_4.

OPUS 6.5 software (http://www.bruker.com) was used to develop PLSR predictive models for each soil chemical property investigated. PLSR is a method that specifies a linear relationship between a set of dependent variables, *Y*, and set of predictor variables, *X* (Farifteh *et al*., 2007). PLSR extracts the orthogonal or latent predictor variables accounting for as much of the variation of the dependent variables as possible (Bilgili *et al*., 2010). Due to the limited number of samples for this study, "leave one out" cross validation (LOOCV) was used. According to Martens and Dardenne (1998) independent test validation yields over-optimistic assessment of predictive models while on the other hand LOOCV gave better predictive performance while yielding little over-optimistic assessment of the models predictive performance when a small number of samples is used for calibration.

Mathematical treatment of the spectral data available in the OPUS 6.5 software involve no spectral data pre-processing, constant offset elimination, straight line subtraction, vector normalization (SNV), min-max normalization, multiplicative scattering correction (MSC), first derivative reflectance (FDR), second derivative reflectance (SDR), FDR + straight line subtraction, FDR + SNV and FDR + MSC. The gaps over which the derivatives were taken were 17 data points. Pre-processing treatments are selected prior to model development and the OPUS software selects the best pre-processing technique for a particular PLSR predictive model. The software generates several models and ranks them according to the root mean square error of prediction (RMSEP) or the root mean square error of cross validation (RMSECV). More than 300 models were generated for each soil chemical property that was investigated.

4.3.2 PLSR predictive model validation

The performance of each model was evaluated using the R^2, RMSECV, ratio of performance to deviation (RPD) and the ratio of performance to inter-quartile distance (RPIQ). The R^2 values indicate the power of statistical correlation between measured and predicted values (Farifteh *et al*., 2007). RMSECV is an error based on *n* calibration samples. RMSCEV is computed by equation 4.1 (Collell *et al*., 2011):

$$RMSECV = \sum_{i=1}^{m} \frac{(\hat{y}_i - y_i)^2}{n} \hspace{4cm} 4.1$$

where $\hat{y}_i$ is the NIR predicted value of the response, y_i is the value of the response measured by a reference method of sample i, and n is the number of samples.

RPD measures the ratio of percentage deviation from the RMSECV. According to (Bellon-Maurel *et al*., 2010), the three categories for RPD model reliability are: (1) excellent models (RPD > 2), (2) fair models (1.4 < RPD < 2) and (3) non-reliable models (RPD < 1.4). RPD is computed by equation 2:

$$RPD = \frac{SD}{SEP} \qquad\qquad\qquad 4.2$$

where *SD* is the standard deviation and *SEP* is the standard error of prediction.

RPIQ is an index recently proposed by Bellon-Maurel *et al*. (2010). They argued that owing to the use of SD in computing RPD, RPD may be misleading concerning the strength of PLSR predictive models for log-normally distributed samples as mostly is the case with soil chemicals. The RPD thresholds are also not grounded in a statistical basis and different thresholds were suggested by different researchers (Farifteh *et al*., 2007; Fuentes *et al*., 2012). RPIQ is based on quartiles which represent the spread of the population better than SD (Cozzolino *et al.*, 2011). RPIQ is calculated by equation 4.3 (Bellon-Maurel *et al*., 2010):

$$RPIQ = \frac{(Q3-Q1)}{SEP} \qquad\qquad\qquad 4.3$$

where *Q1* is the value below which 25% of the samples are found, *Q3* is the value below which 75% of the samples are found and *SEP* is the standard error of prediction.

4.4 RESULTS

4.4.1 Soil chemical properties

The statistics and the Pearson correlation coefficients for the soil chemical properties are given in Tables 4.1 and 4.2 respectively. Soil samples from both study sites display high values for EC (Table 4.1). It is clear from Table 4.1 that Sandspruit soil samples show high variations in Na and Cl. The soil chemical properties yielded moderate to very high correlations with each other. While pH was merely moderately correlated with EC, the other soil chemical properties were

highly correlated with EC. Soil samples from the eastern Johannesburg site have higher pH values (Figure 4.1). Also, the Gauteng site yielded slightly higher than average SO_4 readings. On the other hand, the Sandspruit catchment soil samples yielded lower than average values for Na, Ca, Mg and Cl (Figure 4.1). Figure 4.1 shows that that SO_4, Na, Cl, Mg and Ca are highly correlated to EC.

Table 4.1 Summary statistics of the soil chemical properties

Soil chemical (n)	Mean (all samples)	Range[1]	Range[2]	SD[1]	SD[2]
EC (49)	250.8	4.0 - 1 520.0	37 – 1 234	535.9	232.9
pH (49)	7.1	5.7 – 8.1	5.7 – 8.0	0.6	0.5
Ca (49)	128.3	3.1 – 786.8	330- 528	161.3	125.8
Mg (49)	71.1	1.2 – 442.2	7.33 - 534	144.9	102.4
Na (49)	425.6	3.5 – 4263.3	19.9 – 1776.9	1 479.9	335.7
Cl (49)	1 586.1	4.2 – 10 322.1	14.7 – 27 76. 5	3 568.4	536.9
SO_4 (48)	257.3	1.4 – 2 450.1	30.2 – 4 281.9	689.44	887.9

[1]= Sandspruit soil samples, [2] = Johannesburg soil samples, n = total number of samples, SD = standard deviation

Table 4.2 Pearson correlation coefficients for soil chemical properties of all the soil samples

Soil chemical	EC	pH	Ca	Mg	Na	Cl	SO$_4$
EC	1.0						
pH	0.3	1.0					
Ca	0.8	0.3	1.0				
Mg	0.9	0.3	0.8	1.0			
Na	0.9	0.3	0.7	0.9	1.0		
Cl	0.9	0.2	0.6	0.8	0.9	1.0	
SO$_4$	0.6	0.3	0.9	0.8	0.5	0.4	1.0
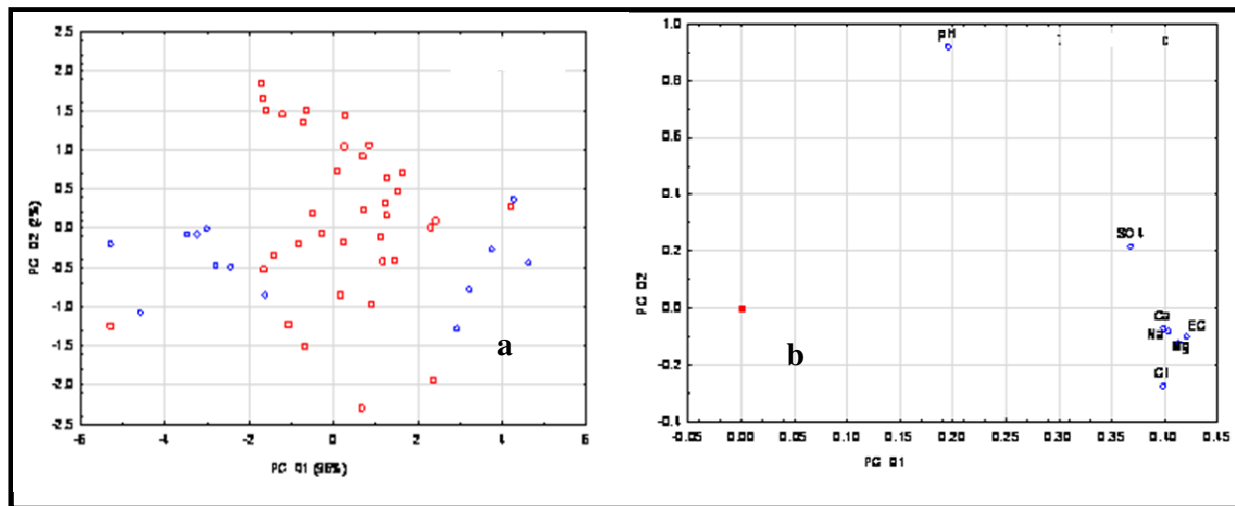
Figure 4.1 Principal component analysis (PCA) plot for (a) the two sites and (b) the soil chemicals PCA X-loadings plot for the two study sites (Blue is Sandspruit and red is Gauteng soil samples)

4.4.2 Spectral features

The spectral features of all the soil samples are depicted in Figure 4.2. The spectral reflectance is typical of soil spectra. The spectra show prominent absorption features which are associated with the bending and stretching of the O-H bonds of free water at 7142, 5128 and 4546 cm^{-1} (equivalent to 1400, 1950 and lattice minerals around 2200 nm respectively) (Viscarra Rossel *et al.*, 2006).
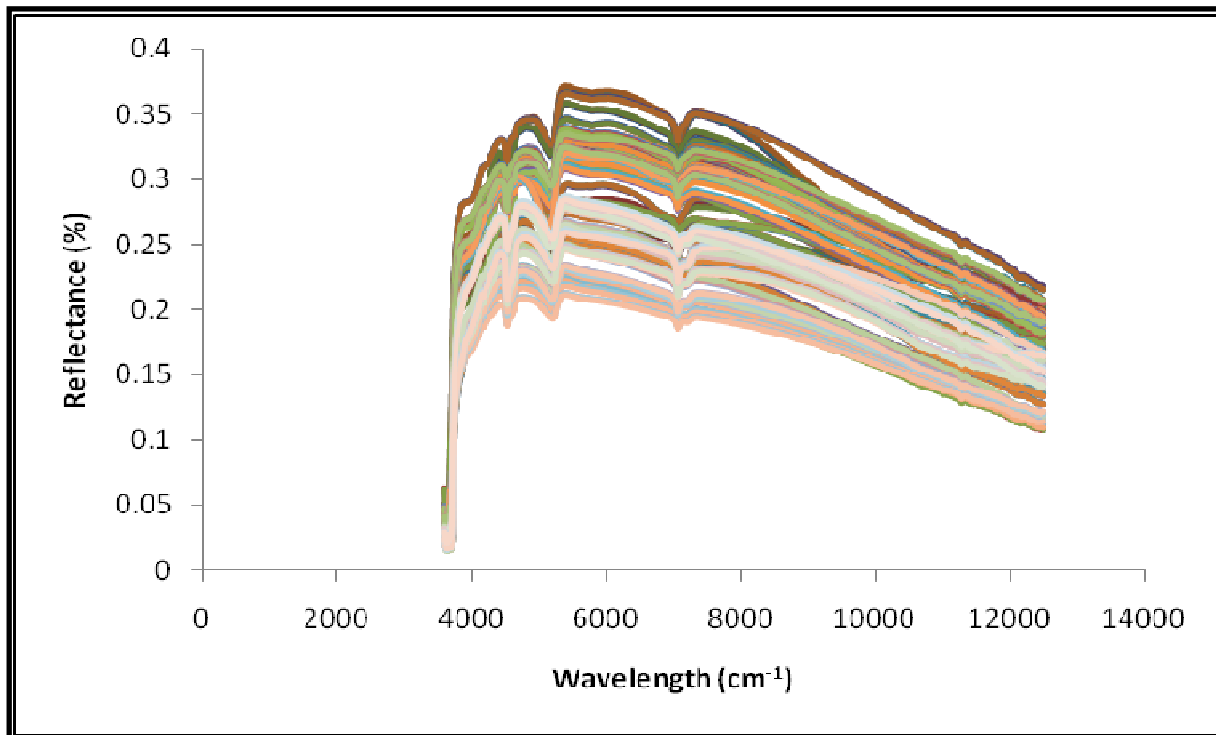
Figure 4.2 Untransformed spectral reflectance of all the soil samples

4.4.3 PLSR modelling

LOOCV PLSR validation statistics, the spectral region used for model development and the OPUS software-determined pre-processing techniques for each soil chemical variable are given in Table 4.3. The LOOCV R^2 values for all the soil chemical properties ranged from 0.62 to 0.87. While soluble Ca yielded the lowest R^2, EC yielded the highest R^2 value. RMSECV for all the soil chemical properties ranged from 0.22 to 0.51. Whereas EC yielded the lowest RMSECV, SO$_4$ yielded a high RMSECV (Table 4.3). RPD values for all the soil chemical PLSR were higher than 1.5. This means that the models can yield fairly accurate to very precise predictions.

According to (Bellon-Maurel *et al*., 2010), predictive models with RPD values less than 1.4 are non-reliable, RPD values greater than 1.4 and less than 2 are fairly accurate and those models with RPD values greater than 2 can yield excellent predictions. RPIQ values for all the soil chemical predictive models were higher than 5. This means that the models can predict accurately. RPIQ is a recent index, so its use on model evaluation is not yet widespread. The spectral regions used to develop the models and the OPUS software-determined pre-processing methods for each soil chemical property is also given in Table 4.3.

Table 4.3 LOOCV PLSR statistics, spectral regions and pre-processing methods

Soil chemical	R^2	RMSECV	RPD	RPIQ	Spectral region (cm^{-1})	Pre-processing
EC	0.87	0.22	2.78	32	5176-4242	SDR
pH	0.66	0.34	1.68	10	7505-4242	SDR
Ca	0.62	0.29	1.61	5	7505-6793, 5453-5021	Min-Man normalization
Mg	0.78	0.29	2.11	19	7505-6093, 5453-5021	Min-Man normalization
Na	0.86	0.29	2.66	26	5453-4597	SNV
Cl	0.85	0.33	2.59	23	5453-4597	Straight line subtraction
SO_4	0.65	0.51	1.7	8	7505-4597	SNV

RMSECV = root mean square error of cross validation, RPD = ratio of prediction to deviation, RPIQ = ratio of performance to inter-quartile distance, SDR = second derivative reflectance, SNV = vector normalization

Scatter plots of measured versus predicted values for all the soil chemical properties are depicted in Figure 4.3. Without doubt, the relationship between the measured and the predicted values for all the soil chemical properties investigated are high (Figure 4.3).
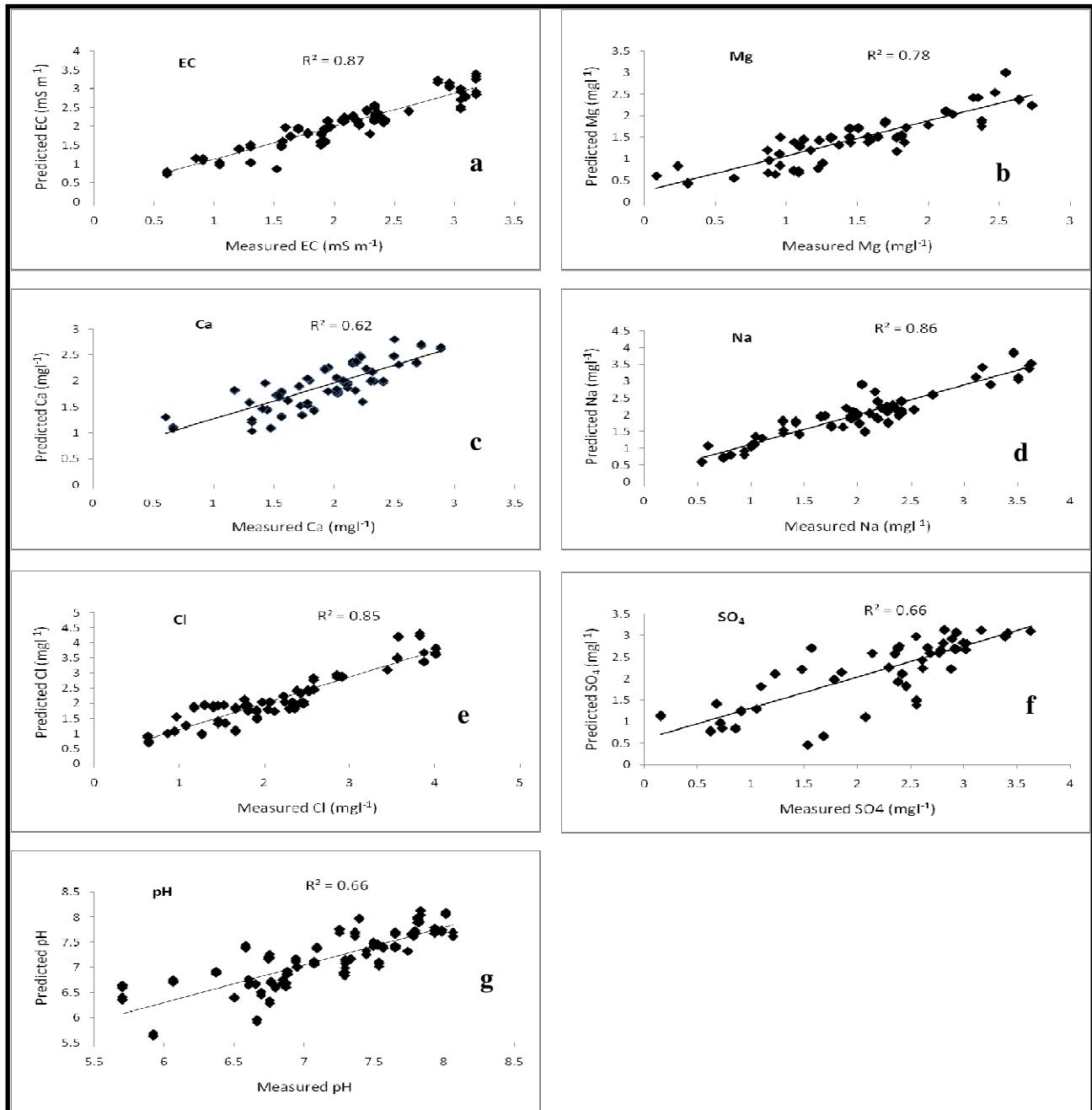
Figure 4.3 LOOCV predicted versus measured scatter diagrams for (a) EC, (b)Mg, (c) Ca, (d)Na, (d) Cl, (f) SO4 and (g) pH

4.5 DISCUSSION

The study found that reliable predictions of soil EC, pH, Ca, Mg, SO_4, Na and Cl could be made based on PLSR using NIR spectroscopy. We found that the PLSR R^2 for EC obtained in this study is comparable to previous studies. While Mashimbye *et al.* (2012) reported R^2 values

between 0.60 and 0.85 for EC, Farifteh *et al*. (2007) found R^2 values between 0.78 and 0.98 for EC using PLSR. A FT Bruker MPA NIR spectrometer and LOOCV were used in this study. Mashimbye *et al*. (2012) and Farifteh *et al*. (2007) used independent test validation and an analytical spectral device (ASD) VIS-NIR spectrometer. The R^2 for EC in this study is higher than that reported by Janik *et al*. (1998), Shibusawa *et al*. (2001) and Islam *et al*. (2003). Presumably, this is due to different regions of spectrum used and the statistical method used. While Janik *et al*. (1998) used PLSR and MIR spectroscopy, Shibusawa *et al*. (2001) and Islam *et al*. (2003) used VIS-NIR and UV-VIS-NIR spectroscopy respectively. Furthermore, Shibusawa *et al*. (2001) and Islam *et al*. (2003) used stepwise multiple linear regressions (SMLR) and principal component regression (PCR) respectively. While the RPD for EC obtained in this study is comparable to that obtained by Farifteh *et al*. (2007), it is higher than of Mashimbye *et al*. (2012). This is most likely due to different statistical calibration techniques used and the chemical composition of the soil. This study used LOOCV while Mashimbye *et al*. (2012) used independent test validation with the EC of the soil being the result of naturally occurring cations and anions, while Farifteh *et al*. (2007) used independent test validation, the EC of their samples was based on irrigating soil samples with different soil samples in the laboratory.

With respect to soluble Mg, this study found that the R^2 using PLSR was comparable to that obtained by Janik *et al*. (1998) using MIR spectroscopy and PLSR. The R^2 for exchangeable Mg in this study is 0.78 while that of Janik *et al*. (1998) was 0.76. On the other hand, a higher R^2 was recorded by Shepherd and Walsh (2002) for exchangeable Mg using MARS based on VIS-NIR spectroscopy. Chang *et al*. (2001) reported a lower R^2 (0.68) using PCR based on VIS-NIR spectroscopy. These differences are presumed to be due to the use of different statistical techniques and the spectral region. No study that used NIR spectroscopy for Mg was found in the literature.

While the R^2 for Ca obtained in this study is lower than that obtained by Shepherd and Walsh (2002) using VIS-NIR region, it is comparable to that of Chang *et al*. (2001) using VIS-NIR region ($R^2 = 0.75$). Chang *et al*. (2001) used a comparable number of samples for calibration. This is highly likely due to different statistical techniques, region of spectrum and the number of samples used. While this study used NIR region and 49 samples for calibration, Chang *et al*.

(2001) calibrated with 30 samples using VIS-NIR region. Shepherd and Walsh (2002) used the VIS-NIR region and 493 samples for calibration.

In the case of soluble Ca, we found an R^2 that was lower than reported by Shepherd and Walsh (2002) ($R^2 = 0.88$) and Chang *et al*. (2001) ($R^2 = 0.75$). The number of samples used for model calibration, the spectral region and the statistical methods used are likely the cause of the differences in the prediction capacity of these methods. Also, Shepherd and Walsh (2002) and Chang *et al*. (2001) used independent test validation while we used LOOCV in this study.

The PLSR predictive model for Na using NIR in this study was more accurate than previously recorded. While we found an R^2 of 0.86 for exchangeable Na, Bikindou *et al*. (2012) reported an R^2 of 0.12 using NIR spectroscopy using PLSR, whereas an R^2 of 0.33 for exchangeable Na was reported by Janik *et al*. (1998) using MIR and PLSR. In addition, an R^2 of 0.09 was reported by Chang *et al*. (2001) using PCR and VIS-NIR for exchangeable Na. It is probable that the higher R^2 obtained in this study is due to the use of composite sampling, the number of scans measured and different instrumentation. Whereas Bikindou *et al*. (2012) made two replications, measured 16 scans per sample and used a Foss NIRystems 5000 spectrometer (Silver Spring, MD, USA) with wavelength range between 1100 and 2500 (equivalent to 9090 to 4000 cm^{-1}) at a 2-nm interval, we used a Bruker MPA spectrometer with a wavelength range of 12 500 to 4000 cm^{-1} (equivalent to 800 to 2500 nm) and a 1-nm frequency. Also, we extracted three replicates from the soil samples and measured 128 scans were done for each replicate.

NIR PLSR results for pH are comparable with those reported by Islam *et al*. (2003) and Reeves III *et al*. (1999) using test set validation. Dong *et al*. (2011) and He *et al*. (2007) also recorded correlation coefficients of 0.89 and 0.91 respectively for pH using NIR spectroscopy. As regards the RMSECV for pH, it was lower than previously recorded by Shepherd and Walsh (2002) and Chang *et al*. (2001). These studies used multivariate adaptive regression splines (MARS) and principal component regressions (PCR) respectively.

 Regarding Cl and SO_4, we that found reliable predictive models could be developed based on PLSR using NIR spectroscopy. No studies that reported the use of MIR, VIS-NIR, NIR and UV-VIS-NIR of Cl and SO_4 could be found.

RPIQ values for all the PLSR predictive models in this study were higher than 5. This suggests that the models are reliable. However, no studies that evaluated the robustness of the RPIQ as a measure of the predictive capability of models were found. RPIQ is a fairly recent validation technique. No studies on the application of this index for soil EC, pH, Ca, Mg, Na, Cl and SO_4 appear to exist. Applications of the RPIQ for model validation have been reported for soil organic carbon and fruit juice by Cecillon *et al.* (2012), Cozzolino and Cynkar (2011) and Sarkhot *et al.* (2011).

Generally, we found that fairly accurate to very accurate predictive models for EC, pH, Ca, Mg, Na, Cl and SO_4 based on PLSR using NIR spectroscopy could be developed. While the relationship of measured and predicted values for EC, Mg, Na and Cl was high, it was moderate for pH, Ca and SO_4. The RMSECV for SO_4 was also high, thus this model may not be as reliable. The differences in the predictive abilities of the models in this study with previous studies are mainly attributed to different instruments, the abundance of soil chemicals, conditions under which the spectra are measured and the statistical methods used. According to Dong *et al.* (2011), the abundance and distributions of the various organic functional groups may influence both the biochemical properties and the NIR spectra. The findings of this study are consistent with previous work.

4.6 CONCLUSIONS

Spectral reflectance for dried, crushed and sieved soil samples were measured using a laboratory spectrometer. LOOCV was used to develop PLSR predictive models for EC, pH, Ca, Mg, Na, Cl and SO_4. We conclude that EC, Ca, Mg, Na and Cl can be reliably predicted using PLSR based on NIR spectroscopy. The pH, Ca and SO_4 predictive models are not as accurate. Since some of the results recorded were better than those reported in previous studies, the use of composite sampling very likely contributed to the superior PLSR predictive models. Although MIR spectroscopy is reported to perform better than NIR spectroscopy for quantifying organic C, the NIR PLSR predictive models used in this study were comparable and in certain instances more accurate than previously reported MIR results. Certainly, NIR spectroscopy can be useful as a routine procedure for analysing saline soils in the laboratory.

CHAPTER 5

ASSESSING THE INFLUENCE OF DEM SOURCE ON DERIVED STREAMLINE AND CATCHMENT BOUNDARY ACCURACY[3]

5.1 INTRODUCTION

Digital elevation model (DEM)-derived catchment boundaries, subbasins and streamlines play an important role in hydrological studies (O'Callaghan and Mark, 1984; Martz and De Jong, 1998; Renssen and Knoop, 2000; Turcotte *et al*., 2001; Vogt *et al*., 2003; Li and Wong, 2010). In the past, hydrological analyses were mainly confined to relatively small areas such as single catchments or irrigation schemes. However, the increasing availability of high-resolution, near-global DEMs, such as the shuttle radar topographic mission (SRTM) DEM and the advanced spaceborne thermal emission and reflection radiometer (ASTER) global digital elevation model (GDEM), offers new opportunities for carrying out hydrological analyses on regional or national levels. DEMs are offered at a variety of resolutions ranging from very high (0.1-5 m) to low (1 km) (Behrens *et al*., 2010; Tarekegn *et al*., 2010). Very high-resolution DEMs, as derived from airborne light detection and ranging (LIDAR) data, are often only available for small areas, particularly in developing countries such as South Africa where this technology is still prohibitively expensive. Consequently, other sources of DEMs must be considered for hydrological studies at national or regional scale.

---

[3] The contents of this chapter were published in the Water Research Commission Report: de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013. The work was also presented at the Geography Doctoral and Post-Doctoral Conference, 11th of November 2011, Stellenbosch, South Africa. The work is also being prepared for publication in a suitable peer-reviewed scientific journal.

Various studies investigated the value of DEMs for hydrological analysis. For instance, Weepener *et al*. (2012) developed a hydrologically improved DEM for South Africa from the SRTM DEM by filling voids using 20-m 1:50 000 contours and ASTER GDEM data. They found that useful riverlines and catchment boundaries can be delineated from the hydrologically improved SRTM DEM. Li and Wong (2010) compared stream networks extracted from the National Elevation Dataset, SRTM DEM, and LIDAR with stream networks extracted from the National Hydrography Dataset. They also compared flood simulations using the stream networks delineated from the different DEMs. They concluded that higher-resolution DEMs can derive more accurate river networks, and that the spatial resolution of a DEM has only minor effects on flood simulation-results. Callow *et al*. (2007) evaluated the effect of commonly used hydrological correction methods (namely stream burning, Agree.aml and ANUDEM v4.6.3 and ANUDEM v5.1) on the overall nature of a DEM. They found that different methods produce non-convergent results for catchment boundaries, stream position and length and that these techniques differentially compromise secondary terrain analysis. Their study also concluded that while hydrological correction methods successfully improved calculation of catchment area, stream position and length, they all increased catchment slope. Vogt *et al*. (2003) presented an approach of integrating medium-resolution digital elevation data (250-m grid cell size) with climate data, vegetation cover, terrain morphology, soils and lithology to derive river networks and catchments over extended areas. They found that the methodology provided a good agreement of river superimposition and drainage density.

The recent availability of medium- (90 m) and high-resolution (30 m) near-global DEMs has opened up new possibilities for hydrological analyses at national and regional scales (Wang *et al*., 2011; Zeilhofer *et al*., 2011; Weepener *et al*., 2012). Researchers are employing these DEMs for hydrological studies, mainly because they are freely available (; Wang *et al*., 2011; Gichamo *et al*., 2012; Weepener *et al*., 2012). However, little is known about the quality of the products that are derived from these DEMs. This study compares the spatial accuracy of catchment boundaries and streamlines derived from four DEMs that are available at national level in South Africa. The paper first describes the study site, the data used, and the methodology employed for delineating streamlines and catchment boundaries. A detailed account of the results is then provided, followed by a discussion and conclusions.

5.2 MATERIALS AND METHODS

5.2.1 The study site

The study area is the Sandspruit catchment of the Berg River catchment (BRC) in the Western Cape Province of South Africa. Refer to section 1.5 for more information about the study area.

5.2.2 Data used

DEMs, reference streamlines and a reference catchment boundary were used in this study. The DEMs and each of the data sets are described in the succeeding paragraphs.

The DEMs considered in this study are the 90-m SRTM DEM, the 30-m ASTER GDEM2, two versions of the 5-m Stellenbosch University digital elevation model (SUDEM) and a 1.5-m DEM generated from GeoEye images (GEOEYE DEM). The SRTM DEM was completed in 2000 and provides the first medium-resolution DEM data at near-global scale (Farr and Kobrick, 2001; Li and Wong, 2010). The SRTM has an absolute vertical error of less than 16 m and an absolute horizontal accuracy of 20 m (Farr, 2000; Mulder *et al*., 2011). According to the Consultative Group on International Agricultural Research Consortium for Spatial Information (CGIAR-CSI, 2011), the SRTM DEM data has been processed to fill data voids, and can be used by a wide range of potential users. The ASTER GDEM was developed jointly by the Ministry of Economy, Trade and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA). The second version of ASTER GDEM (GDEM2) was released in October 2011 (ASTER GDEM Validation Team, 2011) with the inclusion of 26 000 additional scenes to improve coverage. The new version uses a smaller correlation kernel to yield higher spatial resolution and water masking was also enhanced. ASTER GDEM2 was validated by comparing it to the absolute geodetic references over the conterminous United States (CONUS), the national elevation grids over the US and Japan, the SRTM 1 arc-second DEM over the US and 20 sites around the globe, and global space-borne laser altimeter data. The vertical and horizontal accuracy of the GDEM2 is less than 17 m and 71 m respectively (ASTER GDEM Validation Team, 2011; Mukherjee *et al*., 2013). The SUDEM was developed by the Centre for Geographical Analysis (CGA) at the University of Stellenbosch. Large-scale (1:10 000) contours, spot heights and smaller-scale (1:50 000) contours were used to interpolate the DEM. Small-scale contours were only used in areas where large-scale data was not available. Two

DEM products were produced. The first version (Level 1) only used contours and spot heights, while the second version (Level 2) combined contours, spot heights and the SRTM DEM. The mean absolute vertical error of the Level 1 and Level 2 products was estimated (using LiDAR data as reference) to be 2.1 m and 2.2 m respectively (van Niekerk, 2011). The GEOEYE DEM was created from GeoEye stereo images acquired in July 2011 using the rational polynomial coefficients (RPC) model in the LPS module of ERDAS Imagine software. The GEOEYE DEM was extracted at 1.5-m horizontal interval and was validated using reference points (trigonometric beacons) in the Sandspruit catchment. A mean absolute vertical error of 0.70 m was recorded. The GEOEYE DEM was used to delineate a reference catchment boundary and reference streamlines.

Reference streamlines were digitized at a scale of 1:10 000 from the 1-5 othorectified GeoEye stereo images. The reference streamlines were visually compared to the 1:50 000 national riverlines data set. It was found that, although the two data sets were geometrically aligned, the 1:50 000 streamlines were much more generalised and contained many topological errors (e.g. gaps). A reference catchment boundary, generated from the 1.5-m resolution GEOEYE DEM, was used to validate DEM-delineated catchment boundaries. The reference catchment boundary was validated during several field visits and by visual inspection in ERDAS Stereo Analyst.

5.2.3 Delineation of catchment boundaries and streamlines from DEMs

The Arc Hydro extension for ArcGIS software was used to delineate the Sandspruit catchment boundaries and streamlines from the DEMs. All the data sets were projected to the Universal Transverse Mercator (UTM) projection (Zone 34S). Catchment boundaries and streamlines were extracted at the native resolution of the DEMs. The threshold for stream delineation was set at 1% of the maximum flow accumulation according to Arc Hydro's recommended rule of thumb for stream delineation from DEMs. The GEOEYE DEM was used to calculate reference flow accumulation thresholds for the other DEMs at their respective resolutions. For catchment boundary delineation, outlet (pour) points were selected at the same position. A stream network was extracted from the 2-m GEOEYE DEM to enable comparison with previous studies conducted with high-resolution DEMs (Li and Wong, 2010). Catchment boundaries and streamlines extracted from all the DEMs were converted to raster data sets with a cell size of 5 m for comparison purposes. Cells representing boundaries or streamlines were allocated values of

1. All other cells were defined as having no values (i.e. NODATA). Separate raster data sets were created for catchment boundaries and streamlines.

5.2.4 Validation of DEM-delineated catchment boundaries and streamlines

The catchment boundaries and streamlines extracted from the DEMs were visually compared to the reference data sets. Four measures, the correctness index (C_r), figure of merit index (FMI), mean absolute error (MAE), and root mean square error (RMSE) were used to quantitatively evaluate the delineated catchment boundaries and stream networks. The C_r and FMI were used by Li and Wong (2010) to validate stream networks extracted from DEMs, while MAE and RMSE is proposed in this study as an additional measure of spatial agreement.

The C_r compares two sets of raster cells (A and B) which represent DEM-extracted and reference raster data sets respectively (Li and Wong, 2010). The C_r is calculated by the equation:

$$C_r = \frac{N_{(A \cap B)}}{N_B}$$

5.1

where N_B is the number of cells representing the reference raster and $N_{(A \cap B)}$ is the number of cells of the DEM-extracted raster, but the cells are also available in N_B. Index values range between 0 and 1 and indicate the proportion of the reference raster that is correctly represented by the extracted raster (Li and Wong, 2010). High correctness index values mean high accuracy of extracted streams.

According to Li and Wong (2010), C_r does not reflect how well the extracted raster (representing stream networks in their case) can reproduce the entire actual raster and assert that the FMI offers a better solution. The FMI is the ratio of the intersection of the observed change and predicted change to the union of observed change and predicted change (Klug *et al*., 1992; Perica and Foufoula-Georgiou, 1996; Pontius *et al*., 2008). FMI was computed by the equation:

$$FMI = \frac{N_{(A \cap B)}}{N_{(A \cup B)}}$$

5.2

where $N_{(A \cap B)}$ is the number of unique cells found in rasters A and B, and $N_{(A \cup B)}$ is the total number of cells found in both A and B (overlapping cells are only counted once). FMI values

range between 0 and 1 and a higher FMI value indicates a higher overlap between the two raster data sets, therefore high accuracy.

MAE and RMSE consider the offsets (Euclidean distances) between each cell in the reference raster (of which the cells with values represent the reference streamlines or catchment boundaries) and the closest cell in the candidate raster (of which the cells with values represent streamlines or catchment boundaries extracted from the DEMs under consideration). Euclidean distance is calculated from the centre of the reference raster cell to the centre of the extracted raster cell. Figure 5.1 depicts how ED is calculated for streamlines. The sum of the offsets was used to calculate MAE and RMSE using formulae:

$$\text{MAE} = \frac{\sum ED}{N}$$

5. 3

$$\text{RMSE} = \sqrt{\frac{\sum ED^2}{N}}$$

5.4

where, ED is the Euclidean distance between the reference and candidate cells and N is the total number of cells with values in the reference raster. Relatively low MAE and RMSE values indicate a high accuracy of DEM-extracted raster data sets. RMSE is considered a better indicator of accuracy as it is more sensitive to outliers than MAE, but it is often useful to interpret these measures in combination. High differences between MAE and RMSE indicate high variance in the individual errors.
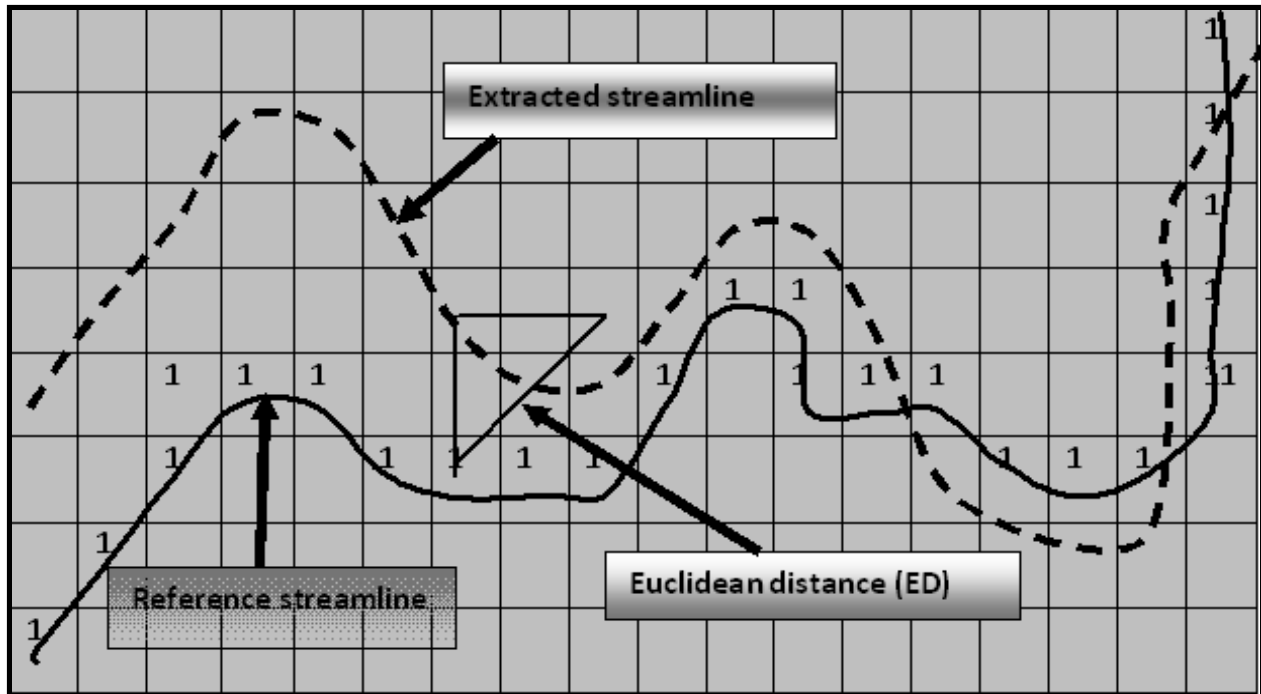
Figure 5.1 Depiction of Euclidean distance calculation using streamlines

5.3 RESULTS

5.3.1 DEM delineated catchments

Overall, the catchment boundaries extracted from all four DEMs seem relatively accurate compared to the reference boundaries (see Figure 5.2a-d and Table 5.1). Visually it seems that the SUDEM L2 delineated the most accurate catchment boundary, although it underestimated the boundary in the south-eastern and north-western parts of the catchment (Figure 5.2a). Some of these errors are attributed to disturbances due to mining activities in the south-eastern section of the catchment. Similar errors are observed in the boundaries derived from the SUDEM L1 product (Figure 5.2b). In addition, a large section of the eastern boundary was incorrectly delineated using the SUDEM L1 product (Figure 5.2b). The ASTER GDEM2 slightly overestimated the catchment boundary at the south-eastern part of the catchment and was also unable to correctly delineate the eastern boundary (Figure 5.2c). The SRTM DEM overestimated the catchment boundary in the south-eastern parts, but performed better than the ASTER DEM and SUDEM L1 in delineating the eastern boundary.

The SUDEM L2 yielded the lowest RMSE and MAE values followed by the SUDEM L1 and ASTER GDEM2 (Table 5.1). In contrast, the SRTM DEM yielded the highest RMSE and MAE values. Similarly, the SRTM DEM yielded the highest ED variance (RMSE – MAE) followed by the ASTER GDEM2, SUDEM L1 and SUDEM L2. This suggests that the SRTM DEM yields high inaccuracies in delineating the catchment boundary than all the DEMs. SUDEM L2 and SUDEM L1 yielded similar C_r and FMI ratios (0.98 and 0.96 respectively). The C_r ratio for ASTER GDEM2 and SRTM DEM is 0.98, while the FMI for SRTM DEM is slightly lower (0.95) than that of ASTER GDEM2 (0.96). From these results it is clear that the SUDEM L2 delineated the most accurate catchment boundary compared to the other DEMs.
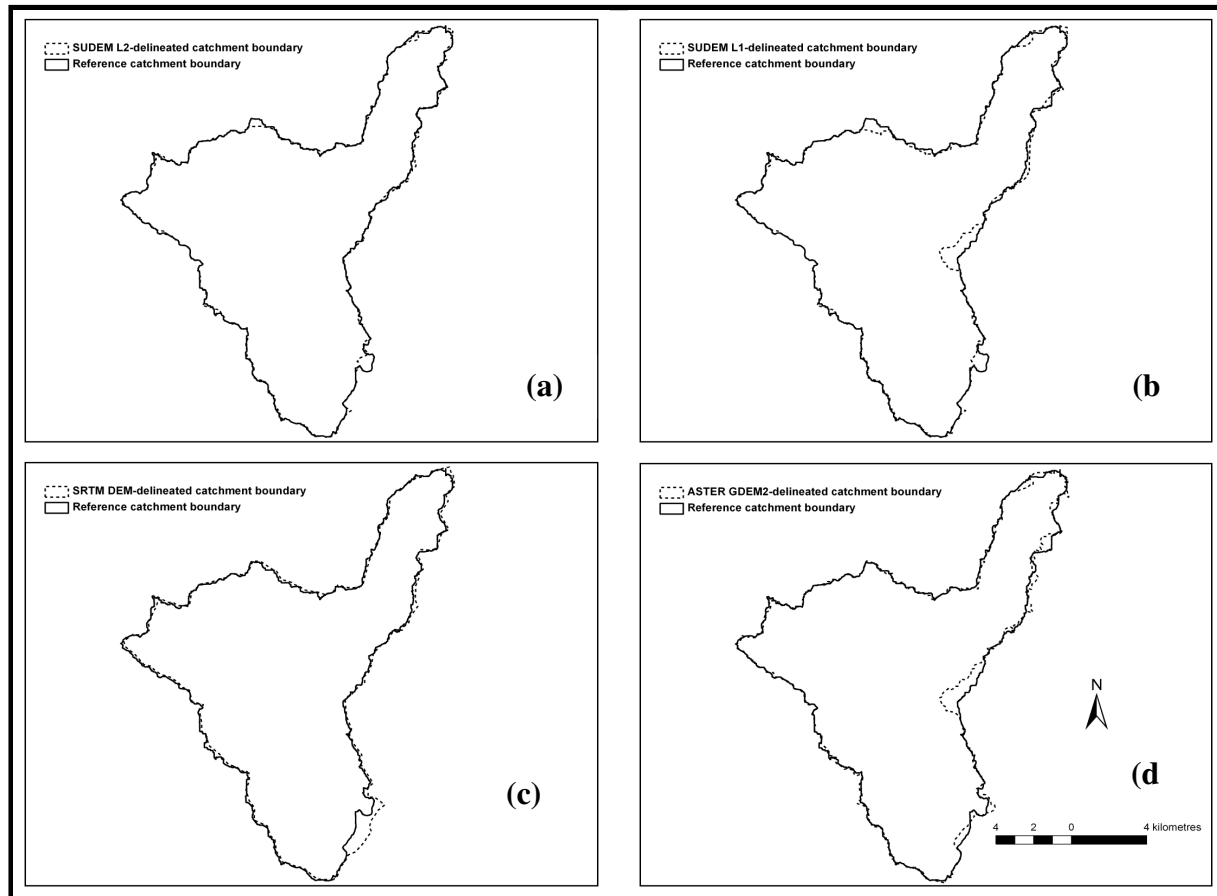
Figure 5.2 DEM-delineated catchment boundaries for (a) SUDEM L2, (b) SUDEM L1, (c) SRTM DEM and (d) ASTER GDEM2

Table 5.1 DEM-extracted catchment sizes, Euclidean distance index, correctness index and figure of merit index

DEM	Supplied Resolution (m)	Catchment size (km^2)	RMSE (ED) (m)	MAE (ED) (m)	Variance (RMSE-MAE)	C_r	FMI
GEOEYE DEM	1.5	153.90					
ASTER GDEM2	30.0	150.21	2.01	0.01	2	0.98	0.96
SRTM	90.0	155.77	5.85	0.06	5.79	0.98	0.95
SUDEML1	5.0	149.18	1.18	0.007	1.17	0.96	0.96
SUDEML2	5.0	152.25	0.98	0.005	0.98	0.98	0.98

RMSE – root mean square error, MAE – mean absolute error, ED – Euclidean distance index, C_r – correctness index, FMI – figure of merit index

5.3.2 DEM extracted stream networks

Streamlines extracted from the DEMs are depicted in Figure 5.3a-e. Generally, the streamlines align well with the reference streamlines, although some alignment distortions were apparent on some areas. The SUDEM L2 yielded the lowest RMSE and MAE values (Table 5.2). Whereas the RMSE and the MAE for the ASTER GDEM2 and SUDEM L1 are comparable, the SRTM DEM yielded the highest RMSE and MAE values. Surprisingly, the RMSE and MAE values of the GEOEYE DEM are higher than those of the SUDEMs (L1 and L2) and ASTER GDEM2. The SRTM DEM yielded the highest ED variance followed by the GEOEYE DEM, SUDEM L1, ASTER GDEM2 and the SUDEM L2 (Table 5.2). The GEOEYE DEM also yielded higher inaccuracies in delineating the streamlines than the ASTER GDEM2, SUDEM L1 and SUDEM L2. While the geometric distortions of SUDEM L2 are minimal, those of GEOEYE DEM, ASTER GDEM2 and SUDEM L1 are comparable (Table 5.2). Clearly, the SRTM DEM yielded high geometric distortion errors in delineating the streamlines. At closer inspection it was

observed that the GEOEYE DEM did not perform well in relatively flat areas (Figure 5.4a-e). However, the GEOEYE DEM performed better than the other DEMs in areas of moderate to complex terrain. This is likely why the GEOEYE DEM yielded the highest C_r and FMI ratios. The C_r and FMI ratios for the two SUDEMs are comparable and slightly lower than those of the GEOEYE DEM, while the SRTM DEM and ASTER GDEM2 yielded significantly lower C_r and FMI ratios.

Table 5.2 DEM-extracted streamlines ED, C_r and FM

DEM	Supplied Resolution (m)	Total stream length (m)	RMSE (ED) (m)	MAE (ED) (m)	Variance (RMSE-MAE)	C_r	FMI
GEOEYE DEM	2.0	116.82	49.52	6.29	43.23	0.21	0.08
ASTER GDEM2	30.0	105.68	41.34	5.09	36.25	0.04	0.02
SRTM DEM	90.0	101.65	95.74	13.88	81.87	0.05	0.02
SUDEML1	5.0	98.81	42.12	4.98	37.14	0.15	0.06
SUDEML2	5.0	101.25	25.29	2.72	22.57	0.14	0.06

RMSE – root mean square error, MAE – mean absolute error, ED – Euclidean distance index, C_r – correctness index, FMI – figure of merit index
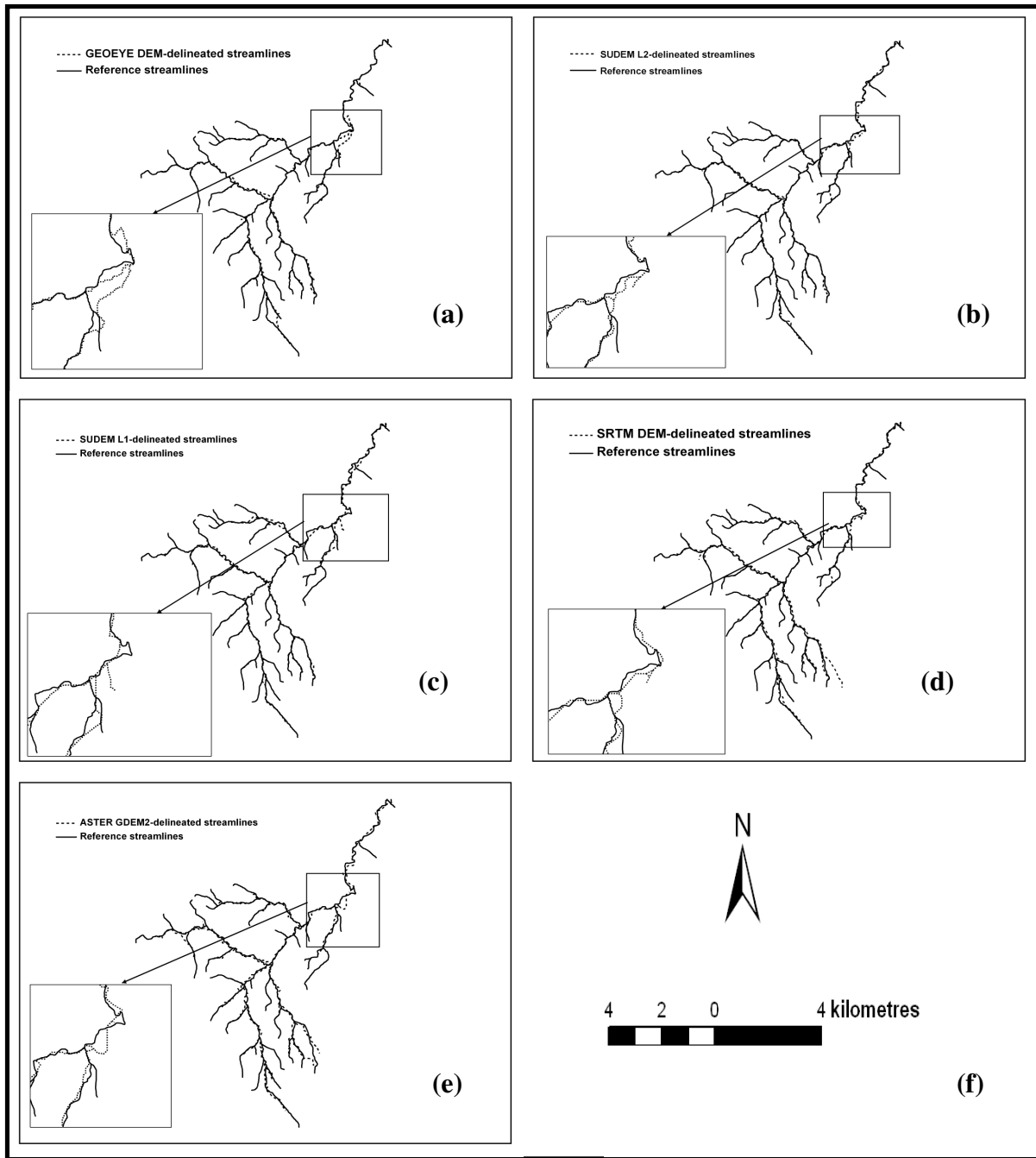
Figure 5.3 DEM-delineated stream networks for GEOEYE DEM (a), SUDEM L2 (b), SUDEM L1(c), SRTM DEM (d) and ASTER GDEM2 (e)
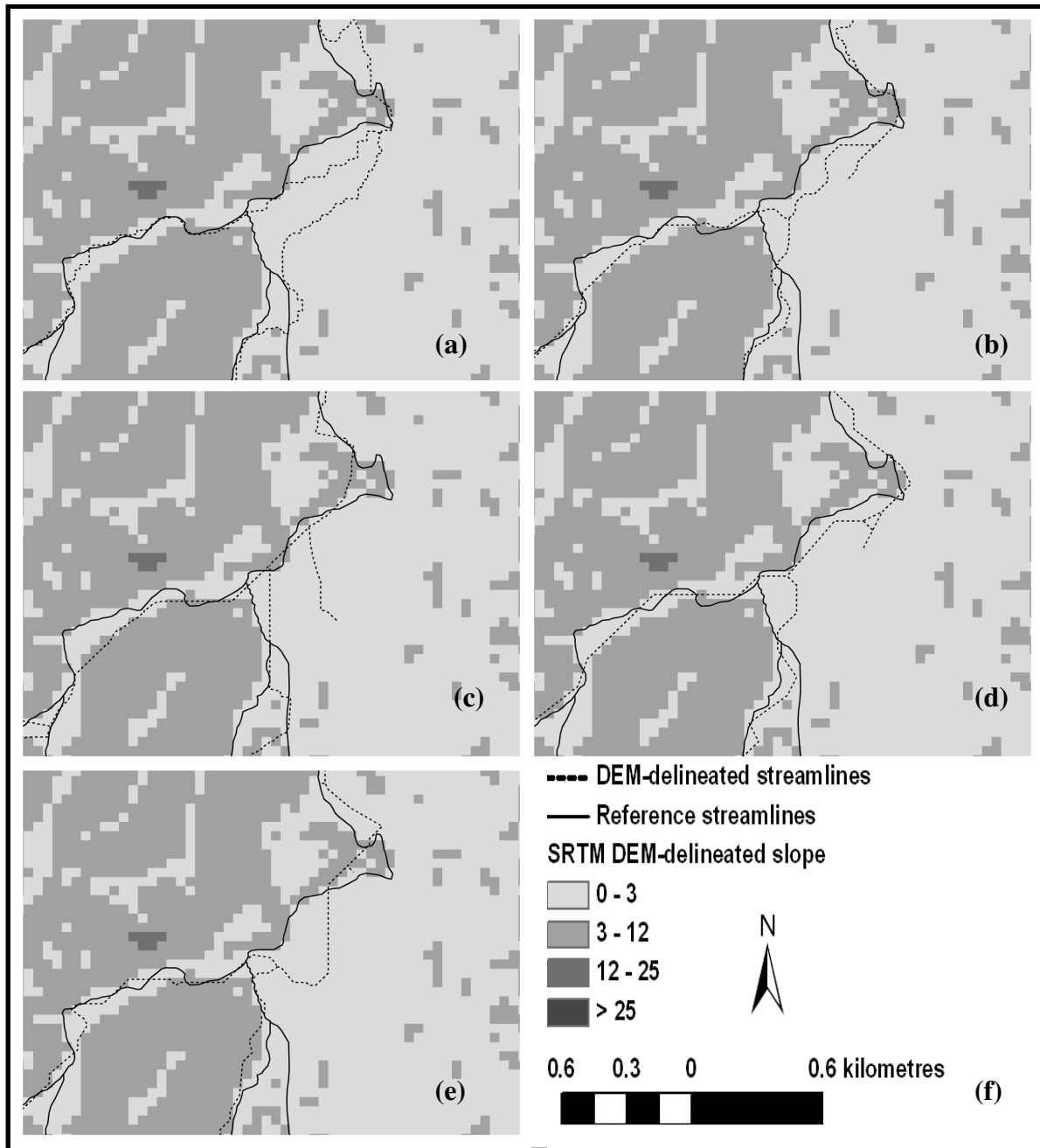
Figure 5.4 Stream networks in a selected area delineated from the (a) GEOEYE DEM, (b) SUDEM L2, (c) SUDEM L1, (d) SRTM DEM and (e) ASTER GDEM2

5.4 DISCUSSION

The results showed that the ASTER GDEM2 yielded a more satisfactory catchment boundary than the SRTM DEM. The SRTM DEM generally overestimated the catchment boundary, while the ASTER GDEM2 underestimated it. This is consistent with the findings of Wang *et al*. (2011) who concluded that the SRTM DEM overestimates valley-bottom elevation values while the ASTER GDEM underestimates the valley-bottom elevation values. The SUDEMs produced the most accurate catchment boundaries, most likely due to their higher resolutions. However, resolution is clearly not the only factor as the SUDEM L2 consistently outperformed the SUDEM L1 which also has a 5 m resolution. The only difference between these two DEMs is that the SUDEM L1 was fused with the SRTM DEM to produce SUDEM L2. According to van Niekerk (2011), this fusion process significantly improves the quality of the DEM in areas of moderate terrain.

In terms of streamline delineation, the GEOEYE DEM outperformed the other DEMs when C_r and FMI ratios are considered. Similarly, the SUDEMs yielded more satisfactory stream networks than both the SRTM DEM and ASTER GDEM2. The higher resolution of GEOEYE DEM and the SUDEMs is likely the main factor in this result. This supports the findings by Vogt *et al*. (2003) who found that the quality of DEM-derived river networks are limited by the spatial resolution and vertical accuracy of the underlying DEMs. In contrast to these findings, the highest resolution DEM considered in this study (i.e. the GEOEYE DEM) yielded the highest RMSE and MAE values in the streamline delineation assessment. This was mainly due to distortions occurring in areas with limited variability in elevation. This confirms the findings of Charrier and Li (2012) who found that higher-resolution DEMs are sensitive to minor topographic features and potentially produce incorrect watershed boundaries, while on the other hand coarser resolution DEMs delineated more accurate catchment boundaries. A possible explanation why the coarser-resolution DEMs (in particular the SUDEM L2) performed better in delineating streamlines in such areas is the way in which these DEM were interpolated. The SUDEMs as well as the 5-m versions of the SRTM DEM and ASTER GDEM2 were interpolated using spline functions. These trend-fitting functions estimate elevations based on trends in the landscape and are consequently known to exaggerate local maxima and minima (e.g. ridges and

valley bottoms). In this study it seems that the interpolators were relatively successful in using these trends to predict where valley bottoms are likely to be, even in very flat areas.

Considering C_r and FMI, the ASTER GDEM2 and SRTM DEM streamlines are comparable. However, the MAE and RMSE values indicate that the ASTER GDEM-extracted streams have fewer geometric distortions than those extracted from the SRTM DEM. The ED variance of the ASTER GDEM is far lower than that of the SRTM DEM. It appears that the ED RMSE and MAE can capture distortions more accurately than the other validation techniques. According to Tarekegn *et al*. (2010), ASTER-based DEMs are relatively accurate in near-flat and smoothly-sloped areas, but they are characterized by large errors in areas covered by forest, snow, steep cliffs and deep valleys. The catchment area in this study is generally flat, which may explain why the ASTER GDEM performed relatively well. The C_r and FMI ratios calculated for the SRTM DEM in this study are lower than those reported by Li and Wong (2010), who recorded C_r and FMI ratios of 0.35 and 0.16 respectively for the 90-m resolution SRTM DEM. The significantly lower C_r and FMI ratios calculated in this study (0.05 and 0.02 respectively) is attributed to the way in which the SRTM DEM was upsampled to 5 m resolution prior to analyses. Upsampling decreased the chances of overlap between the reference and extracted cells. The C_r and FMI ratios are consequently not good indicators of accuracy when DEMs of different resolutions are compared. For such applications, the MAE and RMSE measures are recommended as they are less sensitive to differences in resolution.

5.5 CONCLUSIONS

This study investigated the utility of DEMs for extracting catchment hydrological parameters, namely catchment boundaries and streamlines. The SUDEM L1, SUDEM L2 and two freely-available DEMs (30-m ASTER GDEM2 and the 90-m SRTM DEM) were investigated. The study confirmed that high-resolution DEMs generally produced more accurate parameters, but that other factors such as source data and interpolation algorithm also play a role. It is also evident from the results that the ASTER GDEM2 produced more satisfactory catchment hydrological parameters than the SRTM DEM.

The ED MAE and RMSE proposed in this study can be reliably used to compare reference and DEM-extracted raster data sets of different resolutions and are generally better indicators of

geometrical accuracy than the C_r and FMI ratios. In spite of the relatively lower accuracies of the streamlines and catchment boundaries derived from the SRTM DEM and ASTER GDEM2, the quality of these data sets seems to be acceptable for many applications. Of the available DEMs covering South Africa, the SUDEM L2 is the most suitable product for delineating detailed catchment boundaries and streamlines.

In this study, a catchment with relatively moderate terrain was chosen to assess the quality of the derived data sets. It is, however, expected that the quality of the DEM-derived products will improve as terrain complexity increases, particularly in the case of the SUDEMs as they were mainly interpolated from contours and are as such largely unaffected by distortions caused by view angle and vegetation cover. Contours are also more density distributed in areas of complex terrain, which means that interpolated elevations are generally more accurate in such areas. More research is, however, needed to evaluate how the different DEMs will perform in landscapes with complex terrain and land cover.

CHAPTER 6

AN EVALUATION OF DIGITAL ELEVATION MODELS FOR DELINEATING LAND COMPONENTS[4]

6.1 INTRODUCTION

Terrain is one of the most important soil-forming factors (Behrens *et al*., 2010; Jenny, 1941) and is essential for spoil property mapping (McBratney *et al*., 2003). According to Moller *et al*. (2008), landforms and landscape context are particularly important to understanding the processes of soil genesis and soil formation in the spatial domain. Minar and Evans (2008) and van Niekerk (2008) describe land components as elementary landform elements with a constant value of elevation or having two or more readily interpretable morphometric variables, bordered by lines of discontinuities. Land component borders frequently coincide with environmental land properties such as soil, climate and biology (Speight, 1977; MacMillan *et al*., 2004; Van Niekerk, 2010).

Conventional approaches to delineating land components include studying topographical maps, interpreting aerial photographs and making field measurements (Speight, 1977; Graff and Usery, 1993; Dragut and Blashcke, 2006). However, these methods are often time-consuming, biased and costly (Speight, 1977; Argialas, 1995; Adediran *et al*., 2004; Drăgut and Blaschke, 2006;

---

[4] The contents of this chapter were published in the Water Research Commission Report: de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013.The work was also presented at the Combined Congress, 16 – 19 January 2012, Potchefstroom, South Africa. The work is being prepared for publication in a suitable peer-reviewed scientific journal.

van Niekerk, 2010). The increasing availability of DEMs has promoted the use of computers and image processing techniques for deriving terrain properties. The application of object-based image analysis for land component mapping has gained popularity in recent years (Drăgut and Blascke, 2006; Smith *et al*., 2007; Wulder *et al*., 2008; Drăgut and Eisank, 2011), particularly for soil-landscape modelling purposes (Blaschke and Strobl, 2003; Deng, 2007).

Various researchers have investigated the use of DEMs for digital soil and land component mapping. Van Niekerk (2010) evaluated land component maps delineated from DEMs using three algorithms, namely the automated land component mapper (ALCoM), the iterative self-organizing data analysis technique algorithm (ISODATA) and multiresolution image segmentation (MRS) to determine which technique yields the most homogenous and morphologically representative land components. The three algorithms generated significantly different land component maps and MRS performed better and was more sensitive to morphological discontinuities than the other algorithms. Drăgut and Blaschke (2006) investigated an automated classification system of landform elements based on object-orientated image analysis. Elevation, profile curvature, plan curvature and slope gradient was used to delineate relatively homogeneous objects through image segmentation. This was followed by a classification of objects into landform elements using a relative classification model based on the surface shape and on the altitudinal position of objects. They concluded that the methodology is reproducible and it is readily adaptable for diverse landscapes and data sets. A semi-automated method to recognize and spatially delineate geomorphological units in mountainous forested ecosystems using statistical information extracted from a 1-m resolution digital terrain model (DTM) derived from laser data was proposed by van Asselen and Seijmonsbergen (2006). They determined slope angle and elevation characteristics for each key geomorphological unit occurring in the study area and derived a map of slope classes from the DTM in an expert-driven multilevel object-orientated approach. They concluded that topographical data derived from high-resolution DTMs are useful for the extraction of geomorphological units in mountainous areas.

It has been demonstrated that delineating land components from DEMs is more cost-effective and objective than traditional field-based and visual interpretation methods and that land component mapping is invaluable for landscape characterization and soil mapping (Minar and

Evans, 2008; Moller *et al*., 2008). However, although research has been done on the various algorithms available for segmenting DEMs to produce land components (van Niekerk, 2010), very little has been done to determine how the use of different input DEMs influences the delineation of land components. This paper compares the land components derived from five DEMs, namely the 90-m shuttle radar topography mission DEM (SRTM DEM), the second version of the 30-m advanced spaceborne thermal emission and reflection radiometer global digital elevation model (ASTER GDEM2), two versions of the 5-m Stellenbosch University DEM (SUDEM L1 and L2), and a 5-m DEM (GEOEYE DEM) derived from GeoEye stereo-images. The results are interpreted and evaluated in the context of using land component delineation for mapping and studying soil properties.

6.2 MATERIALS AND METHODS

6.2.1 Study Area

The study area is the Sandspruit catchment of the BRC in the Western Cape Province of South Africa. Refer to section 1.5 for more information about the study area.

6.2.2 Data used

The data used for this study include orthorectified digital aerial photographs and DEMs. Each data set is described in the following paragraphs.

High resolution (0.5 m) orthorectified digital aerial images covering the Sandspruit catchment were obtained from the Chief Directorate National Geo-spatial Information (CDNGI) (http://www.ngi.gov.za). The aerial images were used to digitize terrain morphological discontinuities for assessing the accuracy of the DEM-delineated land components.

The 90-m SRTM DEM, completed in 2000, is the first high-resolution DEM developed at near-global scale (Farr and Kobrick, 2001; Li and Wong, 2010). The SRTM DEM is reported to have a vertical error of less than 16 m (Farr, 2000; Rodriguez *et al*, 2005; van Niekerk, 2008; Mulder *et al*., 2011). According to the Consultative Group on International Agricultural Research Consortium for Spatial Information (CGIAR-CSI) (2011), the latest version of the SRTM DEM has been processed to fill data voids and it is suited to a range of potential users.

ASTER GDEM was developed jointly by the Ministry of Economy, Trade and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA). The full 1.5-million-scene ASTER archive was used to create the DEM. The second version of ASTER GDEM (GDEM2) was released in October 2011 (ASTER GDEM Validation Team, 2011) with the inclusion of 26 000 additional scenes to improve coverage. A smaller correlation kernel was also used to yield higher spatial resolution and enhanced water masking. ASTER GDEM2 was validated by comparing it to the absolute geodetic references over the conterminous United States (CONUS), the national elevation grids over the US and Japan, the SRTM 1 arc-second DEM over the US and 20 sites around the globe, as well as global space-borne laser altimeter data. The vertical and horizontal accuracy of the GDEM2 is less than 17 m and 71 m respectively (ASTER GDEM Validation Team, 2011; Mukherjee *et al*., 2013). The number of voids and artefacts noted in GDEM1 were substantially reduced in GDEM2 and were almost eliminated in some areas (ASTER GDEM Validation Team, 2011).

The GEOEYE DEM was created from GeoEye stereo-images acquired in July 2011. As with the ASTER GDEM, the elevation data that were extracted from the GeoEye imagery included objects above ground (i.e. it is a surface model and not a terrain model). However, because most of the study area is used for cultivation of grains, very few tall objects (e.g. trees and buildings) are present. Moreover, the July images record a time when the crops were at seedling height and thus had very little impact on the extracted elevations. Elevations were extracted at a 5-m horizontal interval using the rational polynomial coefficients (RPC) model in the LPS module of Erdas Imagine software. The resulting GEOEYE DEM was validated using the altitudes at reference points (trigonometric beacons) in the Sandspruit catchment. An absolute vertical accuracy of 0.70 m was achieved. The DEM was smoothed with a 7 x 7 circular median filter to remove artefacts caused by vegetation and crop patterns. Judging by visual inspection of histograms prior to and after the filtering, and the statistics recorded in Table 6.1, the filter did not significantly alter the terrain morphology.

Table 6.1 Attributes of original and filtered 5-m GEOEYE DEM

DEM attributes	Original GEOEYE DEM	Filtered GEOEYE DEM
Minimum elevation (m)	29.0	31.0
Maximum elevation(m)	965.0	956.0
Mean elevation (m)	497.0	493.5
Standard deviation	270.5	267.3

GEOEYE DEM – digital elevation model created from GeoEye stereo images

The SUDEM was developed by the Centre for Geographical Analysis (CGA) at Stellenbosch University, South Africa. Large-scale (1:10 000) contours and spot heights were used to interpolate two DEM products (van Niekerk, 2011) using a combination of interpolation algorithms (e.g. the Topo to Raster and Spline tools in ArcGIS software). The first product (Level 1) only used 5 m vertical interval contours and spot heights as input, whereas the second product (Level 2) combined contours, spot heights and the SRTM DEM. For Level 2, the SRTM DEM was used to supplement the contour and spot height data in areas of low relief (i.e. where contour and spot height density was low).

6.2.3 Data preparations

All the DEMs were projected to the Universal Transverse Mercator projection (Zone 34S). For easier comparison, the SRTM DEM and ASTER GDEM2 were upsampled from their native resolutions (90 m and 30 m respectively) to 5-m resolution. This was achieved by converting the DEMs to points and interpolating new elevation values using the Spline algorithm in ArcGIS 9.3 software. It is clear from Table 6.2 that up-sampling did not significantly alter the data content of the original DEMs. This was confirmed by examining histograms for the up-sampled and original DEMs which showed similar distributions of elevation values prior to and following the up-sampling procedure.

Table 6.2 Attributes of original and upsampled ASTER GDEM and SRTM DEM

DEM attributes	90-m SRTM	5-m SRTM	30-m ASTER GDEM	5-m ASTER GDEM
Minimum elevation (m)	31.0	39.0	21.0	19.0
Maximum elevation (m)	944.0	944.0	957.0	957.0
Mean elevation (m)	460.6	491.5	489.0	488.0
Standard deviation	253.7	261.5	270.5	271.1

ASTER GDEM2 – second version of the 30-m advanced space borne thermal emission and reflection radiometer global digital elevation model, SRTM DEM - the 90-m shuttle radar topography mission digital elevation model

6.2.4 Land component segmentation

Land component segmentation was carried out using the MRS algorithm as implemented in eCognition 8.6 software (http://www.ecognition.com). The MRS algorithm is a bottom-up segmentation algorithm based on a pairwise region-merging technique (Mathieu *et al*., 2007; Blaschke, 2010). According to Trimble (2011), the segmentation procedure starts with single image object of one pixel and repeatedly merges them in several loops in pairs to larger units as long as an upper threshold of homogeneity is not exceeded. In the first step of the procedure the seed looks for its best-fitting neighbour for a potential merger and if best fitting is not mutual, the best candidate image object becomes the new seed image object and is fitted with its best partner. When best fitting is mutual, image objects are merged. In each loop every image object in the image object level is handled once. The loops continue until no further merger is possible (Mancas *et al*., 2005; Thakur and Anand, 2005; Van Niekerk, 2010; Trimble, 2011).

Slope gradient and slope aspect were used as input layers to MRS. Slope aspect was converted to mean vector strength for analysis. A suitable MRS scale factor was determined by experimentation and visual interpretation using hill-shaded DEMs as backdrops (Van Niekerk, 2010; Drăgut *et al*., 2011). A systematic approach was used by increasing the scale factor by one until meaningful objects were obtained (Drăgut *et al*., 2010). The mean slope gradient standard deviation (SGSD) of the objects was used to evaluate the internal homogeneity of the resulting objects. This experimentation with suitable scale factors was carried out on the DEM with the highest detail (i.e. GEOEYE DEM). The MRS algorithm was configured by setting the shape parameter to its minimum value (0.1) and colour was set to its maximum value (0.9) to maximize the internal homogeneity of objects. Both input layers were allocated equal weights in the segmentation.

For the GEOEYE DEM, a scale factor of 12 produced land components that best represented terrain morphology. The scale factors for the other DEMs were adjusted so that their segmentations yielded a similar number of objects to allow comparison. The parameters and the number of objects produced by all the DEMs are summarized in Table 6.3.

Table 6.3 Scale factors and the number of delineated land components for each DEM

DEM	Scale factor	Total number of LC	% difference from GEOEYE DEM LC
ASTER GDEM2	24	21 086	-2.73
GEOEYE DEM	12	21 678	0.0
SUDEM L1	24	21 949	1.25
SUDEM L2	11	21 443	-1.08
SRTM DEM	12	20 670	-4.65

LC –Land components, ASTER GDEM2 – second version of the 30-m advanced spaceborne thermal emission and reflection radiometer global digital elevation model, SRTM DEM - the 90-m shuttle radar topography mission digital elevation model, SUDEM (L1 and L2) – Stellenbosch University digital elevation models (level 1 and 2)

6.2.5 Land component evaluation

Three assessment methods were employed to evaluate the land components delineated from each of the DEMs. Firstly, the land components were visually inspected using hill-shaded DEMs as backdrops. Visual interpretation entailed evaluating how well the land components identified morphological discontinuities (e.g. aspect and slope breaks). Morphological discontinuities to be used in the evaluation were visually delineated on a 0.5-m resolution orthorectified digital aerial photograph covering the study area. The second assessment method evaluated the internal homogeneity of the land components by computing the mean SGSD (van Niekerk, 2010). It was premised that a small SGSD is indicative of high internal homogeneity (i.e. low interclass differences) and that a higher proportion of units with small SGSDs suggests accurate land component delineation (van Niekerk, 2010). The third assessment method employed the mean slope gradient local variance (MSGLV) to determine the effectiveness of the derived land components to detect morphological discontinuities (i.e. high interclass difference). Given that local variance (LV) is the mean of the standard deviation (SD) computed in a small neighbourhood (usually a 3×3 moving window) (Drăgut and Eisank, 2011; Drăgut *et al*., 2011), a satisfactory land component delineation will maximize internal (interclass) homogeneity and minimize external (intraclass) homogeneity. A land component should ideally have a low internal MSGLV and a high MSGLV at its edges. In this study the land component boundaries were defined as being one pixel (5 m) in width and all other pixels were considered internal. Internal and edge MSGLV were calculated for each set of land components derived from each DEM and a MSGLV ratio was computed using the equation:

$$\text{MSGLV ratio} = \frac{Edge\ MSGLV}{Internal\ MSGLV} \qquad\qquad 6.1$$

The MSVLV ratio is a relative measure and attempts to quantify how well land component boundaries coincide with morphological discontinuities.

6.3 RESULTS AND DISCUSSION

A subset of the 0.5-m orthorectified digital aerial photograph showing the delineated test morphological discontinuities and land components generated from the five DEMs is depicted in Figure 6.1a–f. The GEOEYE DEM, SUDEM L2 and the SRTM DEM land components look

similar in shape and are distinctively different from the land components generated from ASTER GDEM2 and SUDEM L1 (Figure 6.1a-f). Closer visual inspection revealed that the GEOEYE DEM very effectively identifies morphological discontinuities (i.e. slope gradient and aspect breaks). Land component boundaries delineated from the GEOEYE DEM and SUDEM L2 mostly coincided with the test morphological discontinuities. The GEOEYE DEM land components were more sensitive to morphological discontinuities than those of SUDEM L2. The GEOEYE DEM land components yielded more detailed morphological discontinuities and incorporated land surface features (for example trees and buildings) in certain areas. This is very likely due to the way the DEMs were created. The GEOEYE DEM was created from stereo-imagery whereas the SUDEM L2 was created from large-scale contour data fused with the SRTM DEM. Consequently, the GEOEYE DEM is a more detailed DEM than the SUDEM L2. The reason why the GEOEYE DEM incorporated land surface features in the delineation in certain areas is because it is a surface model as opposed to the SUDEM L2, which is a terrain model. Despite the SRTM DEM land components looking similar in shape to those of the GEOEYE DEM and SUDEM L2, they were less sensitive to morphological discontinuities. Figure 6.1d shows that the SRTM DEM land components are generalized in certain areas and do not coincide with some significant morphological discontinuities. This is attributed to the lower native resolution (90 m) of the SRTM DEM. The ASTER GDEM2 and SUDEM L1 also failed to identify many significant morphological discontinuities (Figure 6.1e and f). This result confirms those of Gichamo *et al*. (2012), Frey and Paul (2012) and Shafique *et al*. (2011) who found that the ASTER GDEM2 quality is dependent on factors such as quality of the image pair, image acquisition angle and terrain complexity. Contour-interpolated DEMs such as the SUDEM L1 are usually not as accurate as DEMs generated by other means, because DEMs generated from contours suffer from oversampling in steep areas and generalizations in flat terrain (Taud *et al*., 1999; Ardiansyah and Yokoyama, 2002; Xie *et al*., 2003; Wise, 2007; Vaze *et al*., 2010).
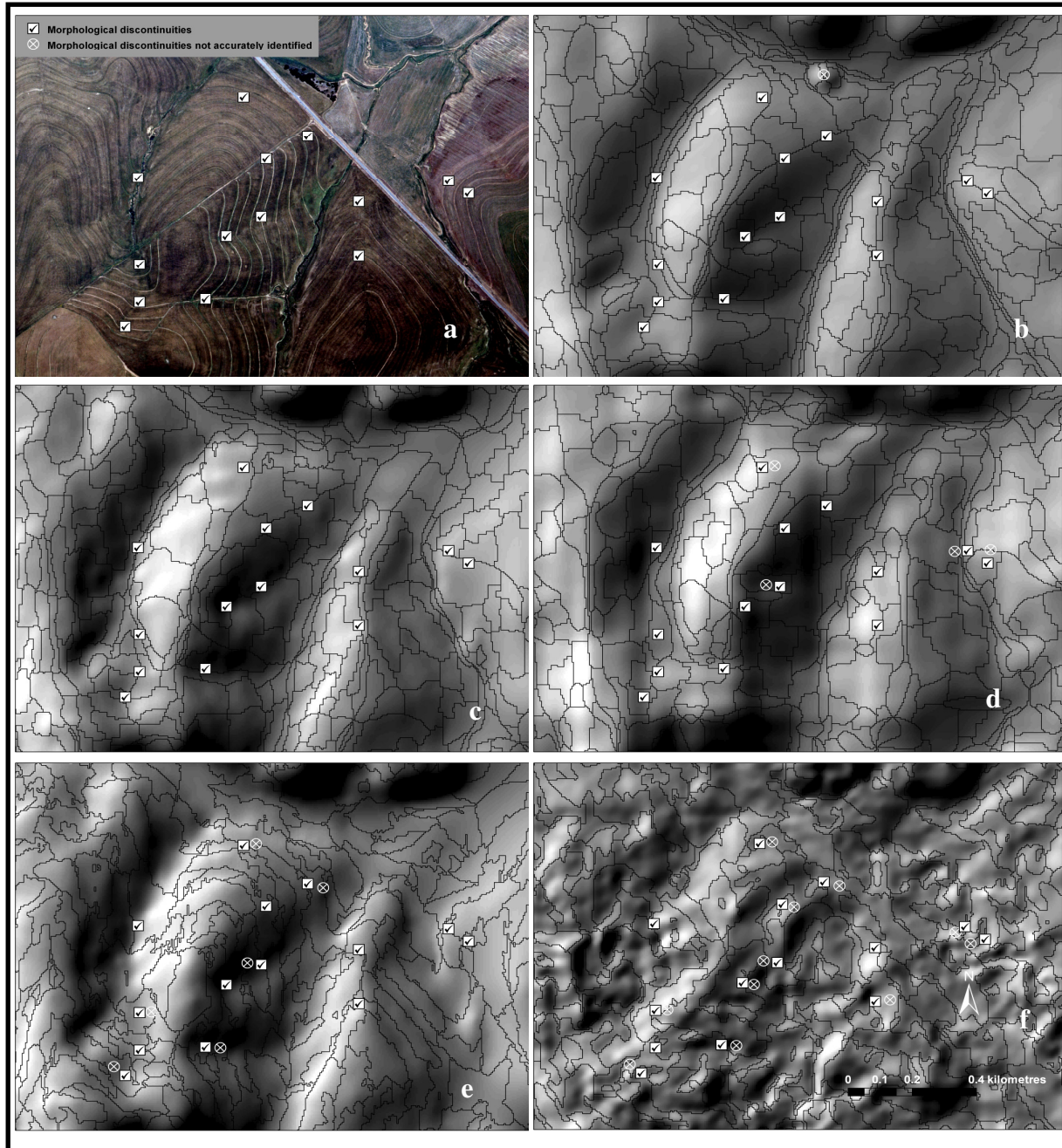
Figure 6.1 Sandspruit orthorectified digital aerial image insert (a), land components delineated from GEOEYE DEM (b), SUDEM L2 (c), SRTM DEM (d), SUDEM L1 (e) and ASTER GDEM2 (f)

The GEOEYE DEM and SRTM DEM yielded the lowest (1.2) overall (mean) SGSD (Table 6.4). This suggests that these products are internally the most homogeneous. The low mean SGSD of the SRTM-delineated land components is attributed to the relatively low resolution of the SRTM

DEM (90 m), which limits the variation within land components. The mean SGSD of the SUDEM L2 product is not significantly higher (1.3) than those of the GEOEYE DEM and SRTM DEM. The ASTER GDEM2 land components are the least homogeneous internally (mean SGSD of 4.4). The histogram of SGSD (Figure 6.2) revealed that, in contrast to the other DEMs, most of the ASTER GDEM2 land components are highly heterogeneous in terms of slope gradient. This result suggests that the ASTER GDEM2 is not suitable for land component mapping.

Table 6.4 Overall SGSDs of digital elevation models

DEM	Mean SGSD
ASTER GDEM2	4.4
GEOEYE DEM	1.2
SUDEM L1	1.5
SUDEM L2	1.3
SRTM DEM	1.2

ASTER GDEM2 – second version of the 30-m advanced spaceborne thermal emission and reflection radiometer global digital elevation model, SRTM DEM - the 90-m shuttle radar topography mission digital elevation model, SUDEM (L1 and L2) – Stellenbosch University digital elevation models (level 1 and 2), LC –Land components, SGSD – slope gradient standard deviation
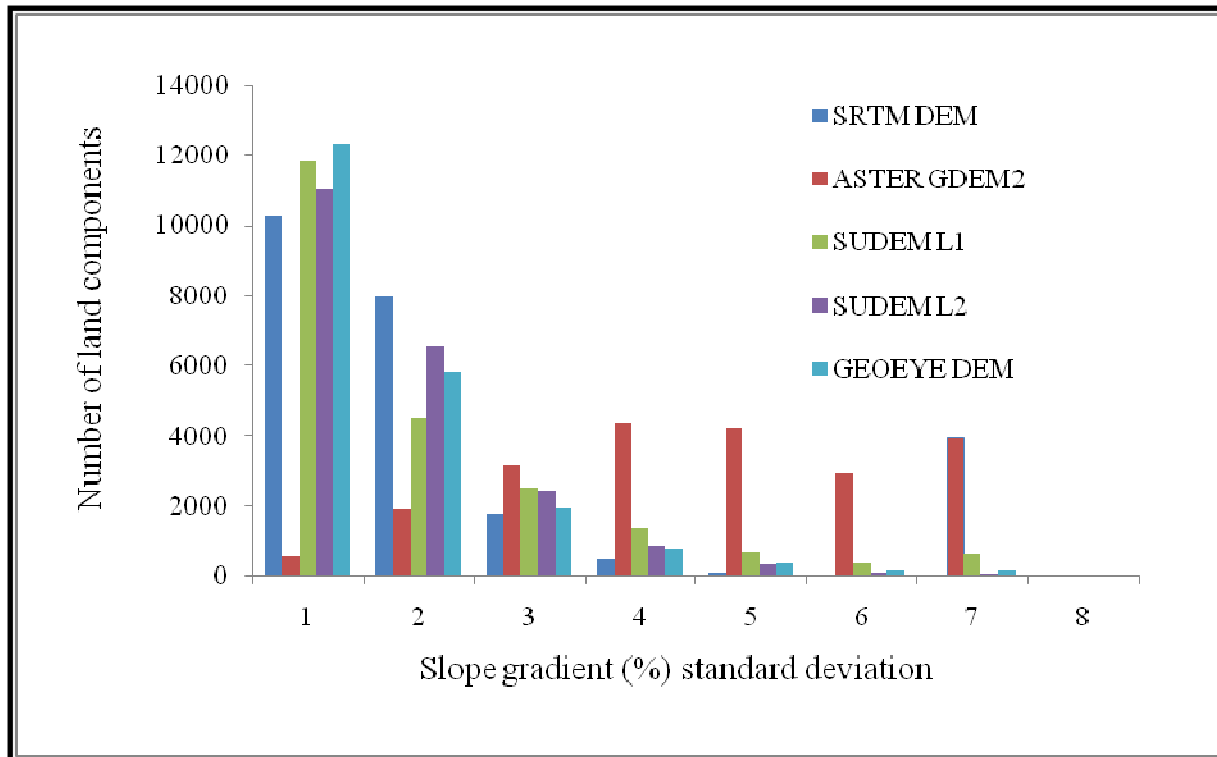
Figure 6.2 Mean slope gradient (%) standard deviation of the land components delineated from different DEMs

The internal MSGLV for the GEOEYE DEM, SUDEM L2, SRTM DEM and ASTER GDEM2 were lower than the edge MSGLV, resulting in a MSGLV ratio of more than 1 (Table 6. 5). This indicates that the internal homogeneity of the land components delineated from these DEMs is maximized, while the homogeneity at the edges is minimized and consequently suggests that land component boundaries coincide with morphological discontinuities. In contrast, the internal and external MSGLV for the SUDEM L1 land components are equal (MSGLV ratio is 1), indicating that morphological discontinuities are not effectively represented by land component boundaries. GEOEYE DEM and SUDEM L2 yielded land components with the highest MSGLV ratio and as such are the most successful in representing terrain transitions. This was confirmed during the visual inspection of the land component boundaries, which revealed that these two DEMs perform equally well in producing land components boundaries that coincide with morphological discontinuities. The high accuracy of delineated land components from SUDEM L2 was unexpected given that SUDEM L1 and SRTM DEM (which was used to develop the SUDEM L2) did not perform as well. This result seems to suggest that the way in which 5-m

vertical contour data were fused with the SRTM DEM in areas of moderate terrain (i.e. where the density of contours is low) optimises the detail of each input DEM (Van Niekerk 2011).

Table 6.5 Land component internal and edge MSGLV for each digital elevation model

DEM	Internal MSGLV	Edge MSGLV	MSGLV Ratio
ASTER GDEM2	1.9	2.4	1.3
GEOEYE DEM	0.6	0.9	1.5
SUDEM L1	0.6	0.6	1.0
SUDEM L2	0.6	0.8	1.5
SRTM DEM	0.5	0.7	1.3

MSGLV – mean slope gradient local variance, ASTER GDEM2 – second version of the 30-m advanced spaceborne thermal emission and reflection radiometer global digital elevation model, SRTM DEM - the 90-m shuttle radar topography mission digital elevation model, SUDEM (L1 and L2) – Stellenbosch University digital elevation models (level 1 and 2)

In spite of its relatively lower MSGLV, SRTM DEM outperformed ASTER GDEM2 regarding the identification of morphological discontinuities as evidenced by visual inspection and the SGSD. This may result from the ability of the microwave energy used to develop the SRTM DEM to penetrate some features, such as trees. In contrast, ASTER GDEM2 includes the height of trees and other objects, and hence is not a true terrain model. This interpretation is consistent with that of Frey and Paul (2012) who found that SRTM DEM yielded slightly more accurate results than ASTER GDEM for the compilation of topographic parameters in glacier inventories. Siart *et al*. (2009) concluded that, despite its coarser resolution, SRTM DEM yielded more satisfactory results than ASTER GDEM for identifying large depressions.

6.4 CONCLUSIONS

This study compared land components delineated from five different DEMs. The GEOEYE DEM (created from GeoEye stereo-images) was the most effective in producing land component boundaries that coincide with morphological discontinuities. The SUDEM L2 (created from contours and SRTM data) produced similar land components to those of the GEOEYE DEM, and it was almost as successful in maximizing internal (interclass) homogeneity and minimizing

external (intraclass) homogeneity. The SRTM DEM appeared to be more suitable for land component mapping than the ASTER GDEM2.

A novel measure, namely the MSGLV ratio, was developed and applied in this study for evaluating how well land component boundaries coincide with morphological discontinuities. The MSGLV ratio measures the relationship between internal homogeneity and external heterogeneity of land components. The ratio complimented the other validation techniques used.

The research demonstrated that a DEM's properties (e.g. resolution, source data, and development method) have significant impacts on the delineation of land components. This has decisive implications for all applications using land components. An example of such an affected application is digital soil mapping which relies on the principle of a strong relationship between terrain and soil properties, and that soil boundaries coincide with land component boundaries. Discrepancies between land component boundaries and terrain transitions will consequently lead to unreliable deductions and inaccurate soil maps.

CHAPTER 7

THE VALUE OF TERRAIN ATTRIBUTES TO MAP SOIL SALINITY[5]

7.1 INTRODUCTION

Soil salinity presents a serious risk to agricultural production and to the environment. It is vital to identify and map areas at risk of salinity to ease the danger of salinity to agricultural production and the environment. Tangible evidence of dryland salinity has been observed in the Berg River catchment (BRC). The emergence of dryland salinity in the BRC is attributed to human activities, particularly the removal of deep-rooted plants as has been the case in Australia and other countries where dryland salinity occurs (Fey and de Clercq, 2004; Dent, 2007; Kingswell and John, 2007; de Clercq *et al*., 2010). According to Dowling *et al*. (2003) dryland salinity can be expressed as salt-affected land or degraded stream water quality. Agricultural crops use less water than the deep-rooted native plants causing variations of hydrological processes and thus the rising of water tables over time (Kingswell and John, 2007).

Remote sensing and geographical information systems (GIS) offer advantages to ground-based methods because they make it possible to objectively map vast areas at risk of soil salinity. It is recognized that salinised land frequently develop in lower valley locations and at breaks of slope (Freeze and Cherry, 1979; Evans *et al*., 1990; Williamson, 1998; Fetter, 2001; de Clercq *et al*., 2010). However, Barrett-Lennard and Nulsen (1989) and de Clercq *et al*. (2010) argued that

---

[5] The contents of this chapter was published in the Water Research Commission Report: de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013. The work is being prepared for submission to a suitable peer-reviewed journal.

topography alone was not adequate to predict the location of all salinised areas. A study by de Clercq *et al*. (2010) demonstrated that there was an inverse relationship of soil salinity with elevation where areas of salinity greater than 60 mS m^{-1} occur at areas less than 200 m. Also, Akramkhanov *et al*. (2011) maintain that terrain indices have low but noteworthy influence on bulk soil salinity.

Despite the awareness of the relationship of soil salinity with terrain attributes, the utility of terrain attributes-based predictive models to map saline prone areas has not yet been fully understood. This study aims to evaluate the utility of mapping saline prone areas using DEMs and their derivatives. Two DEM-based approaches for mapping potentially saline areas are investigated. First, a high-resolution (5 m) DEM and EC of soil samples collected from within the Sandspruit catchment are used to develop soil salinity regression predictive models based on elevation, slope gradient percentage, terrain wetness index (TWI) and curvature using the CurveExpert software (http://www.curveexpert.net). The elevation, slope gradient percentage, TWI and curvature were derived from a 5-m level 2 Stellenbosch University DEM (SUDEM L2). Second, elevation, slope gradient percentage, TWI and curvature which were derived from a 20-m DEM are used to develop soil salinity regression predictive models using the EC of groundwater. The soil salinity predictive models were used to map the risk of dryland salinity in the Sandspruit catchment. The results are evaluated in the context of using DEM-based terrain attributes for enhanced mapping of soil salinity at local and regional scales.

7.2 MATERIALS AND METHODS

7.2.1 The study site

The study area is the Sandspruit catchment of the BRC in the Western Cape Province of South Africa. Refer to section 1.5 for more information about the study area.

7.2.2 Data used

Three data sets were used for this analysis, namely soil samples, groundwater hydro-census data and DEMs. Each data set is explained in the succeeding paragraphs.

Twenty three (23) top soil samples were collected from within the Sandspruit catchment. The samples were analyzed for EC using a 1: 5 saturated paste extract. The EC of the samples were

used to investigate the relationship of EC with terrain attributes in the Sandspruit catchment. The location of the soil samples collected in the Sandspruit catchment is shown in Figure 7.1.
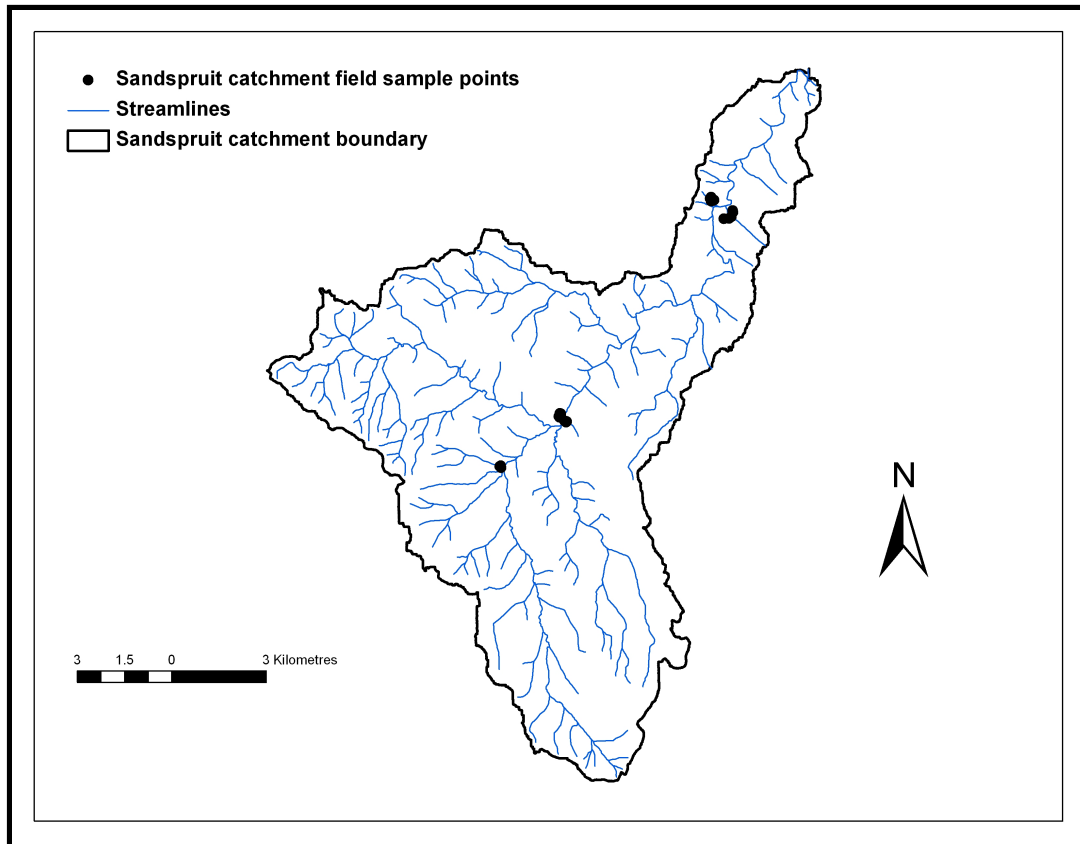
Figure 7.1 Sandspruit catchment field sample points

Hydro-census data was obtained from the National Groundwater Information System of the Department of Water Affairs and Forestry (DWAF) (http://www.dwa.gov.za), South Africa. The data records the EC range of groundwater, annual evapotranspiration, the aridity index and the geographic location. The EC ranges of the groundwater are 0 - 70, 70 -150, 150 - 300, 300 – 500, 500 – 1000 and greater than 1000 mS m^{-1}. The data covered the whole of the Berg River catchment and the points were spaced at approximately 1 000 m from each other. The distribution of the hydro-census data in the Berg River catchment is shown in Figure 7.2. The hydro-census data was used to develop soil salinity predictive models using elevation, slope gradient percentage, TWI and curvature.
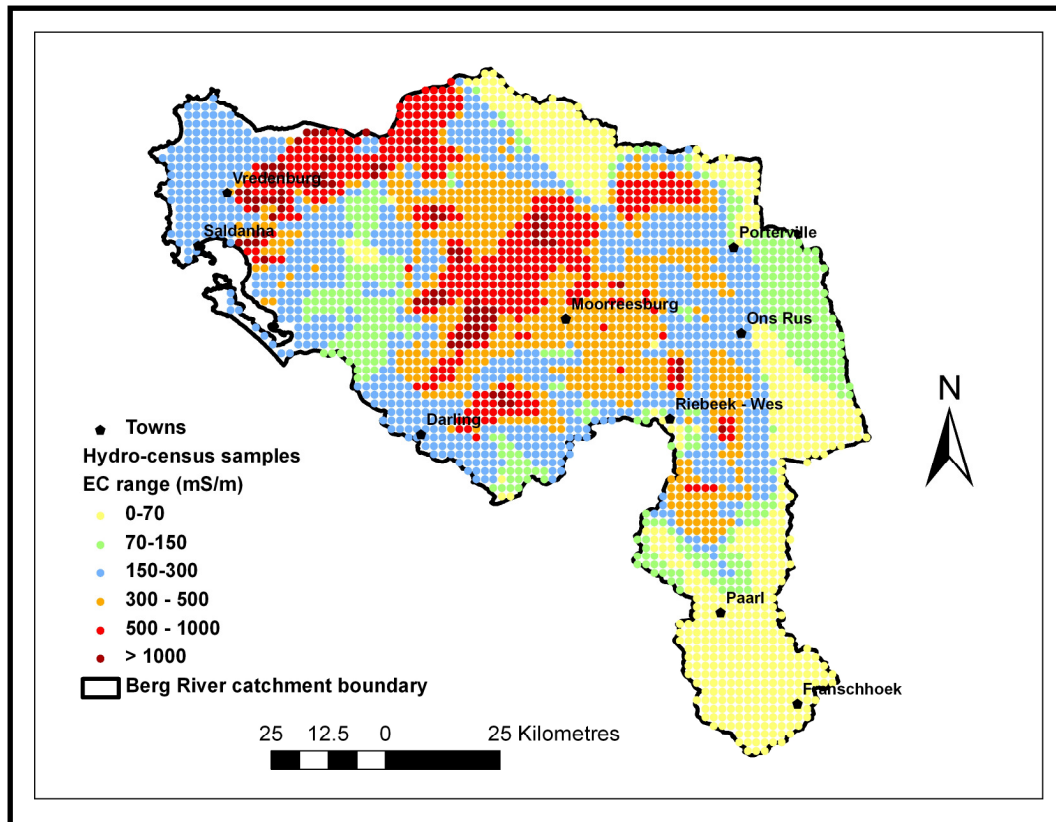
Figure 7.2 Hydro-census samples covering the Berg River catchment

In the case of DEMs, two DEM products were used in this investigation, namely the 5-m Stellenbosch University DEM level 2 (SUDEM L2) product and the 20-m Western Cape DEM (WCDEM). Both DEMs were developed by the Centre for Geographical Analysis (CGA) at Stellenbosch University, South Africa. The 5-m SUDEM L2 was created from large-scale (1:10 000) contours, the shuttle radar topography mission (SRTM) DEM (at areas of low relief) and spot heights using a combination of interpolation algorithms (e.g. the Topo to Raster and Spline tools in ArcGIS software) (van Niekerk, 2011). The 20-m WCDEM was developed by the Stellenbosch University CGA using contours digitized from 1:50 000 national topographic map series (van Niekerk, 2001). The vertical accuracy of the WCDEM was determined by computing mean absolute error (MAE) and root mean squares error (RMSE) using highly accurate elevation data obtained from Chief Directorate National GeoSpatial Information (CDNGI), South Africa (van Niekerk, 2008). It was found to have a MAE and RMSE of 7 and 10 m respectively.

While the SUDEM L2 was used to develop soil salinity predictive models using the EC of the soil samples collected from within the Sandspruit catchment, the WCDEM was used to compute terrain attribute-based soil salinity predictive models using the EC of groundwater hydro-census data. The SUDEM L2 was used because it is a true DEM and was found to perform as well as the GEOEYE. The WCDEM was used because of its higher resolution and is more accurate than the SRTM DEM.

7.3 DATA ANALYSIS

7.3.1 Development of soil salinity regression predictive models

Curve Expert software was used to develop soil salinity regression predictive models based on the EC of the soil and groundwater using elevation, slope gradient percentage, TWI and curvature. Concave and convex curvatures were treated separately as it was challenging to model them together. CurveExpert software models data using a toolbox of linear regression, nonlinear regression models, interpolation or splines. The basic version (CurveExpert 1.4) can be downloaded for free at http://www.curveexpert.net/curveexpert-basic. The software computes several models and ranks the models using the standard error (SE).

The EC of the soil was determined by the 1:5 saturated paste extract. The soil samples were collected from within the Sandspruit catchment. The EC of groundwater was based on the DWAF hydro-census data. Hydro-census calibration samples were randomly selected from the entire Berg River catchment. This was done because it was problematic to model using all the points covering the Berg River catchment as the EC of the groundwater was based on ranges and not actual values. Also, the sampling was done so that the predictive models could be independently validated using the samples that were not used in the calibration. The calibration samples were selected by constructing a 1 000 m buffer around a line feature delineated to cover most of the catchment. The points that intersected with the buffer were extracted. The total number of calibration samples used was 203. The distribution of calibration samples in the Berg River catchment is depicted in Figure 7.3. The calibration samples were used to develop soil salinity regression predictive models using CurveExpert. Soil salinity predictive models that explained the most variation in EC were used to map the risk of soil salinity in the Sandspruit catchment.

In addition, stepwise multiple linear regressions was used to develop soil salinity predictive models based on the natural logarithm of EC, slope, elevation, TWI, annual evapotranspiration and the aridity index. The resultant model was used to map soil salinity in the Sandspruit catchment using Arc Map 10.0.
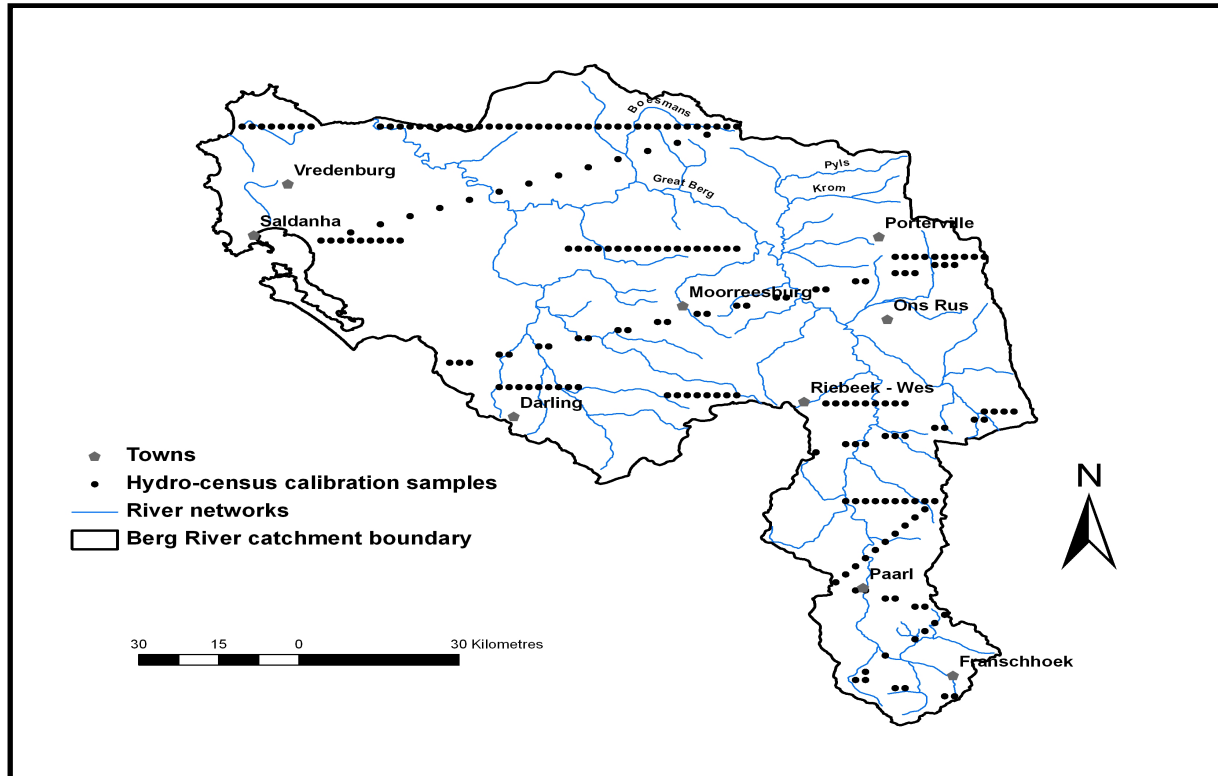
Figure 7.3 Hydro-census calibration samples in the Berg River catchment

7.3.2 Validation

The terrain attribute-based salinity predictive models were used to map dryland salinity risk in the Sandspruit catchment. The soil salinity predictive models developed using CurveExpert software were validated using the correlation coefficient (r) and the SE which are generated by the software. The r indicates the strength of statistical correlation between measured and predicted values. The R^2 was used to validate the model generated by the Statgraphics software. In addition, the soil salinity predictive models were independently validated using 54 hydro-census samples covering the Sandspruit catchment. The distribution of the hydro-census validation samples in the Sandspruit catchment is shown in Figure 7.4. Measured groundwater EC values were compared with the predicted soil EC values. The percentage of correctly

predicted EC range values was computed for each of the terrain-attributes-based soil salinity regression predictive model.
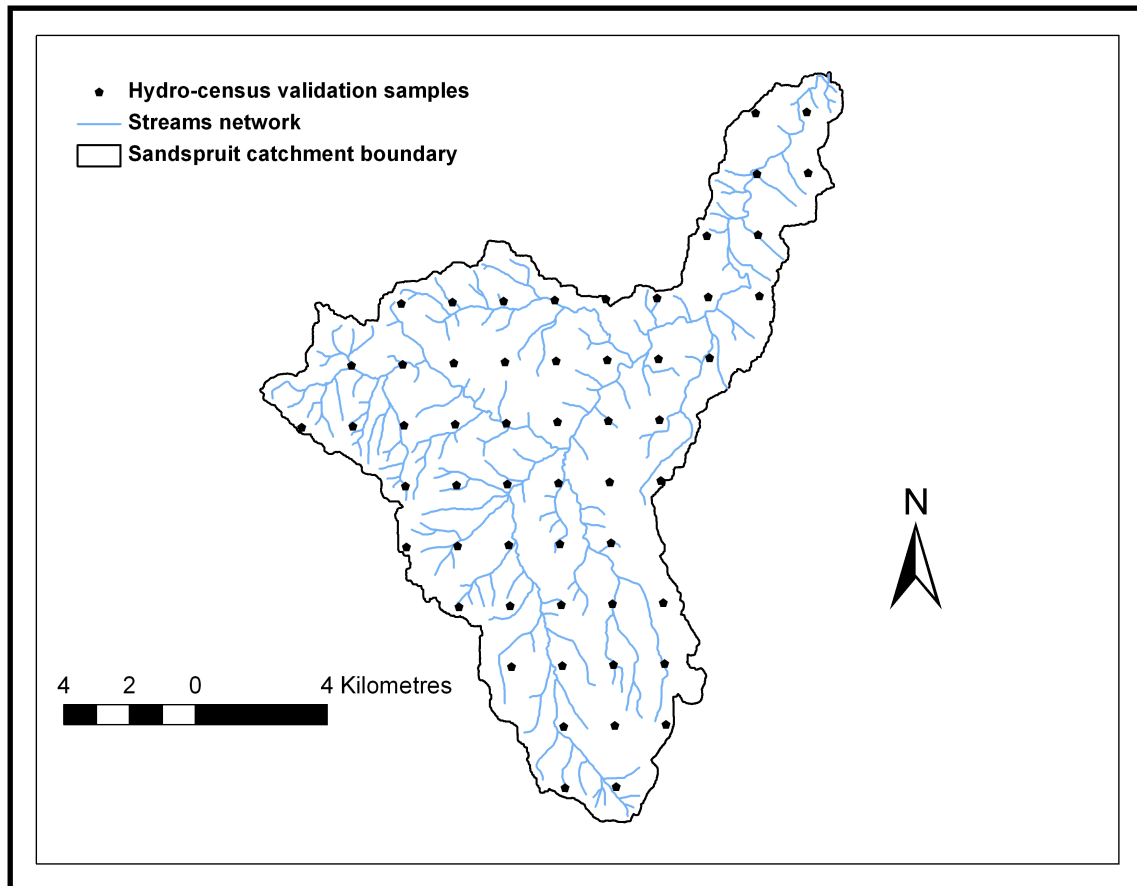
Figure 7.4 Hydro-census validation samples covering the Sandspruit catchment

7.4 RESULTS AND DISCUSSION

7.4.1 Soil EC-based salinity regression predictive models

Terrain-attributes-based soil salinity regression predictive models which were developed based on the EC of the soil are depicted in Figure 7.5. The terrain attributes were derived from the 5-m SUDEL L2. Table 7.1 gives the corresponding equations and statistics of the soil salinity predictive models. An elevation-based soil salinity quadratic regression predictive model yielded the highest *r* followed by the slope gradient percentage-based soil salinity predictive model (Table 7.1). The elevation-based soil salinity predictive model yielded the lowest SE (SE equal to 447.85). Despite this model yielding the highest *r* and a lower SE, it seems unlikely that

it can yield accurate soil salinity predictions. Judging by the model fit, it is obvious that the model can predict negative EC values (Figure 7.5a). EC values may not be negative. This is perhaps due to inadequate number of samples used for model calibration. Also, the soil samples are not evenly spread throughout the catchment. The samples are restricted to areas very close to the river due to problems of accessibility in the catchment.

The r values for concave curvature-, convex curvature- and the TWI-based soil salinity predictive models are low and comparable (Table 7.1). The low r values suggest that the relationship between EC and these attributes is low. SE values for convex curvature-, TWI- and concave curvature-based soil salinity predictive models are 452.08, 486.49 and 678.11 respectively. Due to a weaker relationship of EC with curvature and TWI, these models were not considered for further analysis. Further analysis was conducted with the elevation and slope-based soil salinity predictive models.

Potential soil salinity maps based on elevation and slope gradient percentage predictive models for the Sandspruit catchment were computed using ArcMap 10.0 software (http://www.esri.com). The slope gradient percentage-based soil salinity predictive model shows that areas of low slope and flat areas are highly prone to soil salinity. Potential salinity maps derived from the elevation and slope gradient-based regression predictive models are depicted in Figure 7.6. Despite a high SE and a moderate r for the slope gradient-based soil salinity predictive mode, the results of this model are consistent with the theory that soil salinity would occur at areas of low slope (Figure 7.6a). On the contrary, the elevation-based soil salinity predictive model indicates that elevated areas are more susceptible to salinity than low lying areas (Figure 7.6b). Clearly, this is unlikely. This model would less likely yield reliable soil salinity predictions. This is most likely due to limited data available for model calibration. The samples are also not spread throughout the catchment and thus may not be representative of the conditions in the catchment. It was recorded that soil salinity usually occurs at low lying areas (de Clercq *et al*., 2010).

Table 7.1 Soil EC-based soil salinity regression predictive models equations and statistics

Terrain attribute	Equation	Correlation coefficient (r)	Standard error (SE)
Elevation (m)	$y = 16726.82 - 366.79x + 1.94x^2$	0.66	447.85
Concave curvature	$y = 658.28 + 2834.48x$	0.29	678.11
Convex curvature	$y = 286.71 - 591.24x$	0.27	452.08
Slope (%)	$y = 1131.45e^{-0.31x}$	0.55	486.49
Terrain wetness index	$y = 117.25 + 94.04x$	0.28	578.78
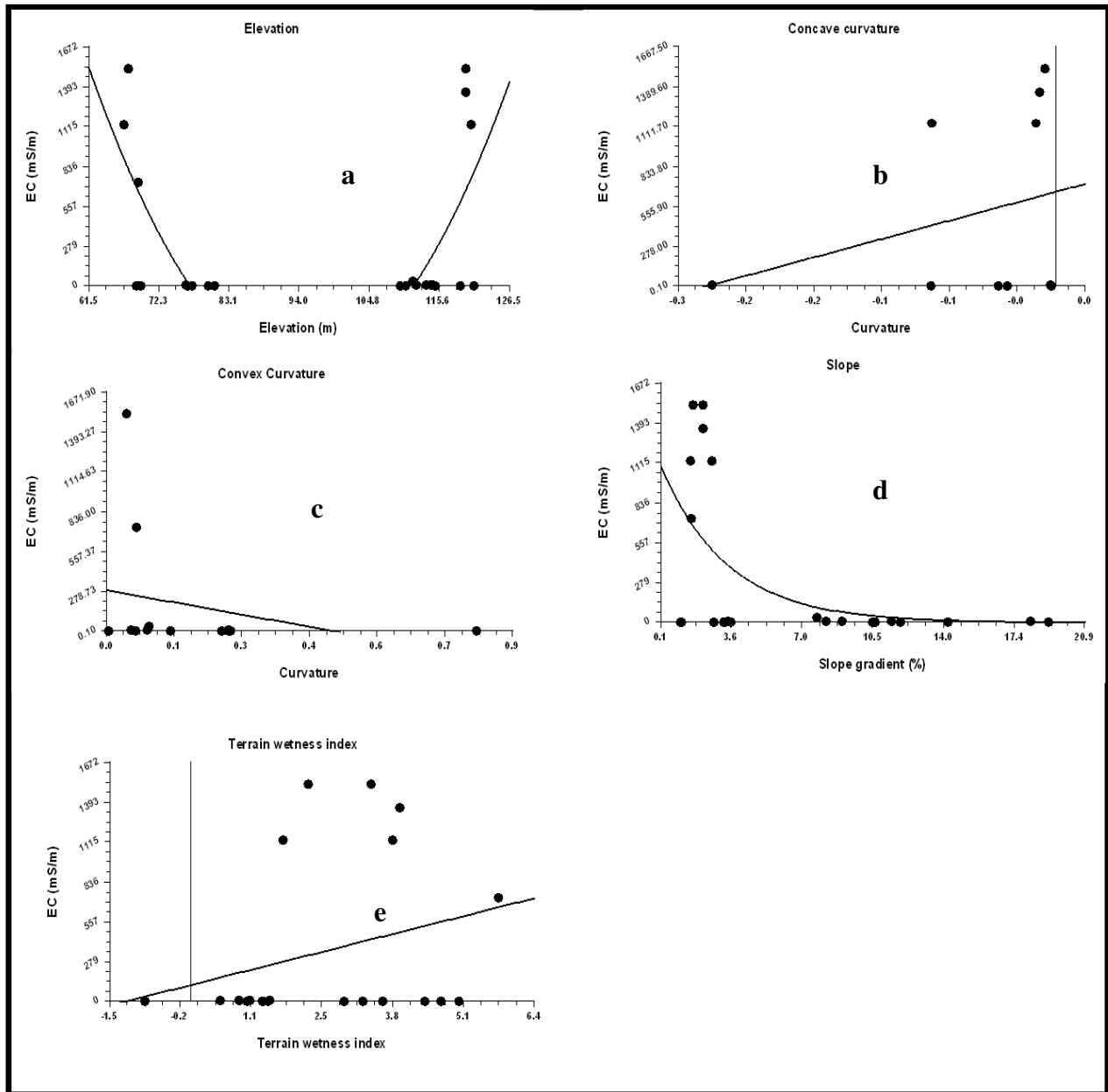
Figure 7.5 Terrain attribute-based soil EC regression predictive models for elevation (a), concave curvature (b), convex curvature (c), slope gradient percentage (d) and terrain wetness index (e)
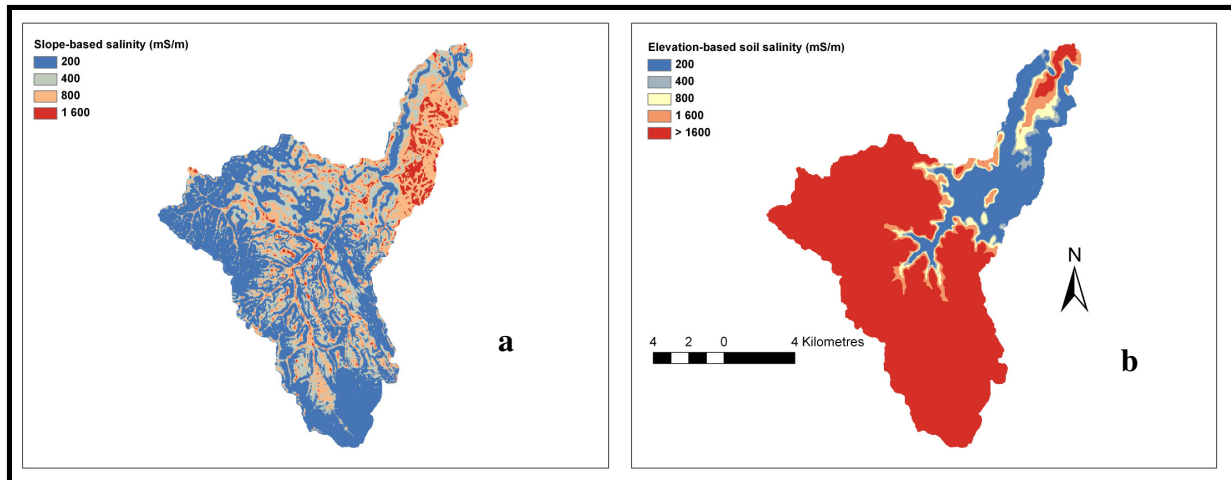
Figure 7.6 Sandspruit catchment potential salinity maps derived from (a) slope-based soil salinity predictive model, (b) elevation-based soil salinity predictive model

Soil salinity maps computed using the elevation- and slope-based soil salinity predictive models were compared with the measured groundwater hydro-census EC values covering the Sandspruit catchment. The percentage of accurately predicted soil EC was computed. The validation of the soil EC-based soil salinity predictive models is given in Table 7.2. While the slope-based soil salinity predictive model yielded a moderately accurate prediction, the elevation-based soil salinity predictive model yielded inaccurate predictions of soil salinity (Table 7.2). This is mostly likely because the elevation-based soil salinity predictive model fit seems to predict negative values. It appears that the elevation data was not adequate to yield reliable a calibration. Contrary to the elevation-based soil salinity predictive model, the slope-based soil salinity predictive model yielded somewhat reasonable soil salinity predictions. The results are consistent with Barrett-Lennard and Nulsen (1989) who recorded that elevation alone is not sufficient to identify saline areas in the landscape

Table 7.2 Soil EC-based soil salinity predictive models accuracy

Soil salinity regression predictive model	Percentage correctly identified (%)
Slope-based soil salinity predictive model	39.13
Elevation-based soil salinity predictive model	8.69

7.4.2 Groundwater EC-based soil salinity regression predictive models

Soil salinity regression predictive models were developed based on the EC of groundwater and terrain attributes. Terrain attributes-based soil salinity regression predictive models are depicted in Figure 7.7. The corresponding equations, correlation coefficient and the SE of the models are given in Table 7.3. While elevation, concave curvature, convex curvature and TWI yielded an exponential relationship with EC, slope yielded a logarithmic relationship with EC (Table 7.3). Although not recorded in Table 7.2, an exponential relationship of soil salinity with slope was also observed. This relationship recorded a slightly lower *r* than the logarithmic relationship reported here. The *r* values for curvature with EC were low (*r* equal to 0.32 and 0.37 for concave and convex curvature respectively). This means that the relationship of curvature with EC is low. Hence the predictive models for curvature were not considered for further analysis. Elevation, slope and TWI-based soil salinity predictive models were considered for further analysis. Whereas *r* for slope was 0.44, elevation and TWI yielded the same *r* value (r = 0.51). The concave curvature soil salinity predictive model yielded the highest standard error (SE) while the elevation-based soil salinity predictive model yielded the lowest SE (Table 7.3). The SE for TWI, convex curvature and slope were comparable (SE equals to 316.76, 317.22 and 321.11 for convex curvature, TWI and slope gradient percentage respectively). SE values for groundwater EC-based soil salinity predictive models are lower than those of soil EC-based regression predictive models. This is most likely due to the number of samples used for model calibration. The moderate *r* values for slope gradient percentage-, elevation- and TWI-based soil salinity predictive models indicates that these models may not be accurate in predicting soil salinity. Perhaps this is due to the fact that the groundwater hydro-census data is course. We observed that low lying areas in the landscape are more prone to soil salinity (Figure 7.7a). Similar to the soil EC-based soil salinity regression predictive model, the groundwater EC-based soil salinity predictive model indicates that flat areas and areas of slopes lower than 4% are highly susceptible to soil salinity. This is consistent with the theory. Regarding TWI, the study found that wet areas are prone to soil salinity. This is most probably because the water table is most likely going to rise under wet areas. Salts will be mobilized to the surface when this occurs.

Table 7.3 Groundwater EC-based soil salinity regression predictive models equations and statistics

| Terrain attribute	Equation	Correlation coefficient (r)	Standard error (SE)
Elevation (m)	$y = 745.53e^{-0.0086x}$	0.51	308.68
Concave curvature	$y = 441.15e^{6.25x}$	0.32	355.99
Convex curvature	$y = 507.78e^{-42.02x}$	0.38	316.76
Slope (%)	$y = 430.71 - 106.19\ln(x)$	0.44	321.11
Terrain wetness index	$y = 93.04e^{0.25x}$	0.51	317.22
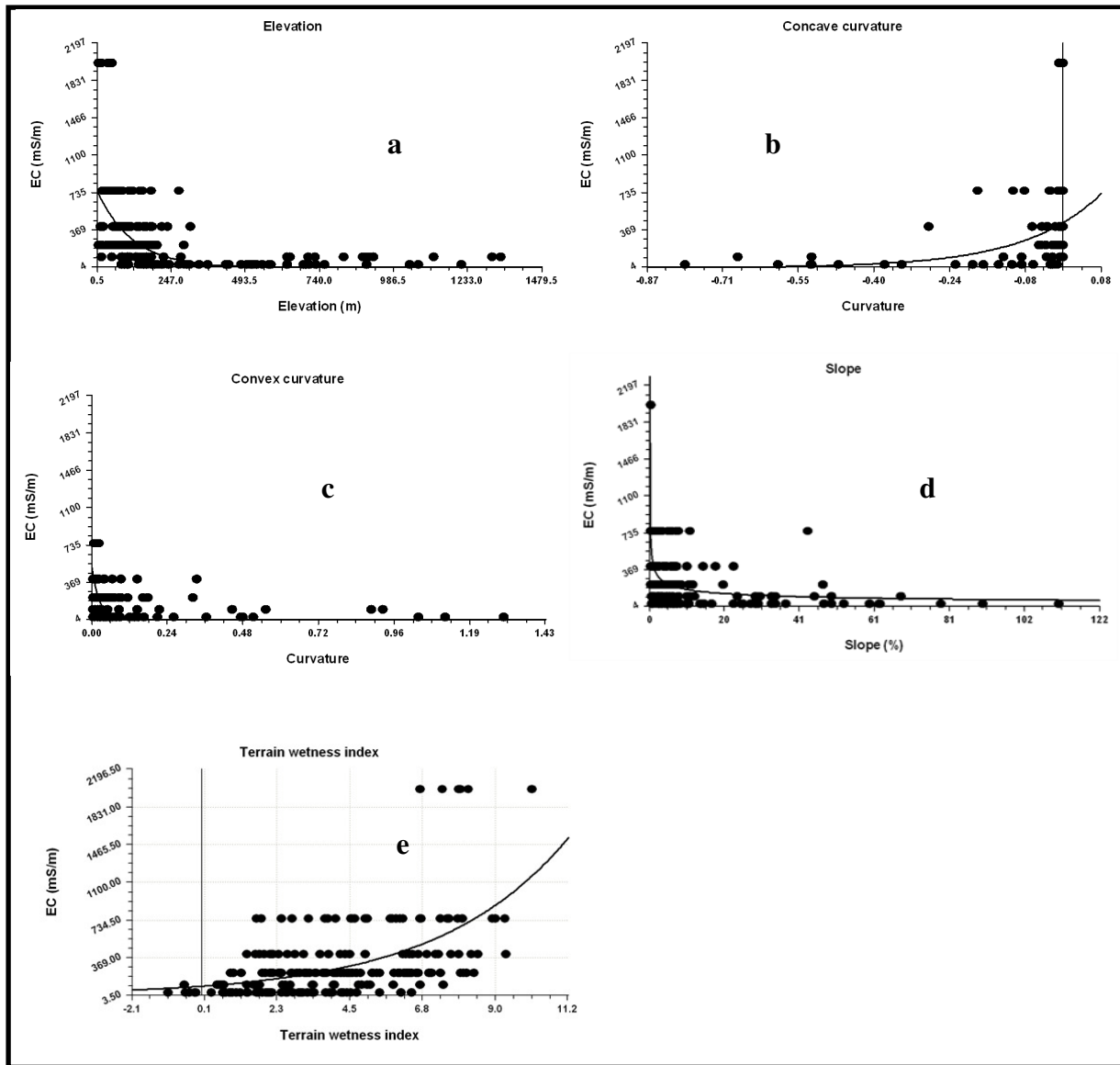
Figure 7.7 Terrain attribute-based EC range regression predictive models for elevation (a), concave curvature (b), convex curvature (c), slope gradient percentage (d) and terrain wetness index (e)

The elevation-, slope gradient percentage- and TWI-based soil salinity regression predictive models were used to map the risk of dryland salinity in the Sandspruit catchment using ArcMap 10.0 software (http://www.esri.com). The soil salinity maps based on elevation, slope and TWI predictive models are shown in Figure 7.8. The elevation-based soil salinity map of the Sandspruit catchment indicates that the low lying areas in the catchment are highly prone to

salinity (Figure 7.8a). This is probably due to the fact that groundwater flows from elevated areas to low lying areas. Generally, the results indicate that low lying areas, flat areas and wet areas are highly susceptible to soil salinity. This is consistent with the findings by Barrett-Lennard and Nulsen (1989) and De Clercq *et al*. (2010) who established that salinised land frequently develop in lower valley locations.
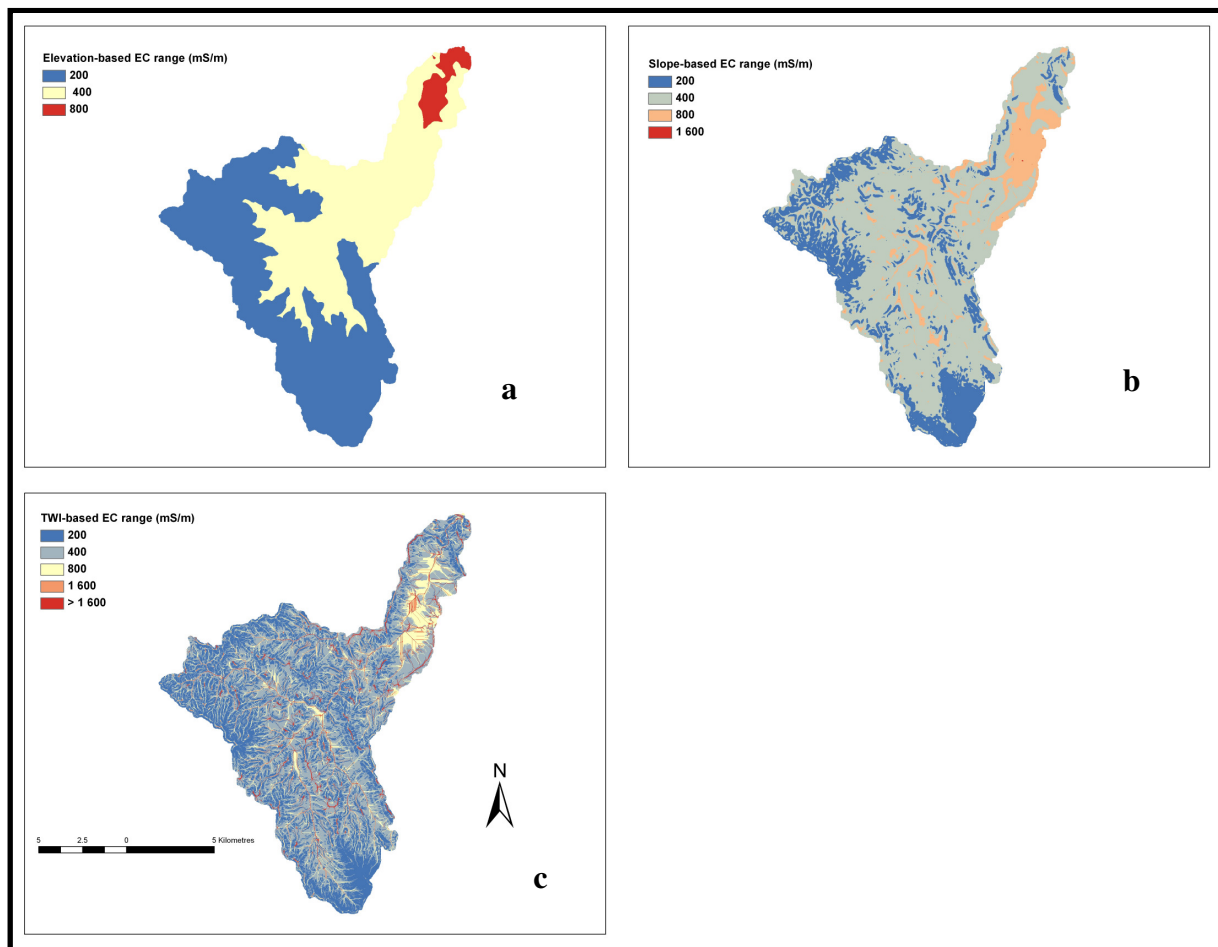
Figure 7.8 Sandspruit catchment potential saline areas derived from elevation (a), slope gradient percentage (b) and terrain wetness index (c)

The percentage of accurately predicted elevation-, slope- and TWI-based soil salinity predictive models is recorded in Table 7.4. The models yielded low to moderate predictions. The TWI-based soil salinity predictive model yielded the highest percentage of accurately predicted soil salinity followed by the slope-based soil salinity predictive model and the elevation-based soil salinity predictive model (Table 7.4). The soil EC-based soil salinity predictive model yielded slightly more accurate predictions than the groundwater EC-based soil salinity predictive model.

This is perhaps due to the resolution of the DEMs. While a very high-resolution (5 m) was used in the case of the soil EC-based model, a 20-m WCDEM was used in the calibration of the groundwater EC-based predictive model.

Overall, the soil salinity predictions from all the models used in this study are somewhat less accurate. These inaccurate predictions are likely due to the data being recorded as EC ranges and that the spread of the groundwater hydro-census samples was course. While the soil EC-based predictive models were based on a very high-resolution DEM, there were not sufficient soil samples to use for calibration. A more detailed sampling and a higher-resolution DEM would likely yield better results. It is clear that there is good potential for DEM-based prediction of soil salinity at local and regional scales.

Table 7.4 Soil salinity predictive models accuracy

Soil salinity regression predictive model	Percentage accurately predicted (%)
Elevation-based soil salinity predictive model	16.36
Slope-based soil salinity predictive model	25.45
TWI-based soil salinity predictive model	34.69

7.4.3 Stepwise multiple regressions

A correlation matrix showing the R^2 values amongst the variables is given in Table 7.5. The EC values were transformed to natural log for this analysis. While the R^2 of the natural EC with the aridity index and annual evapotranspiration was higher, precipitation and elevation yielded a moderate R^2 with EC (Table 7.5). The relationship of EC with elevation and TWI was moderate (R^2 = 0.25 and 0.38 for TWI and slope respectively). A multiple stepwise regression was conducted using Statgraphics software to establish a possible interaction between elevation, slope gradient percentage, TWI, aridity index and annual evapotranspiration. Stepwise multiple linear regressions revealed that the best soil salinity predictive model included elevation, TWI and annual evapotranspiration. The stepwise soil salinity predictive model is given by equation

7.1. The statistics of the model are given Table 7.6. Table 7.7 gives the analysis of variance for the model.

LogEC = -0.00305 (elevation) + 0.031452 (TWI) + 0.00668 + 0.920952 (annual evapotranspiration)   7.1

The R^2 and the R^2 (adj) for the stepwise soil salinity predictive models are comparable. This suggests that the model is somewhat stable in its predictions. The SE and mean absolute error for the stepwise regressions salinity model are low (Table 7.6). This is most likely due to the use of log transformed EC values.

Table 7.5 Matrix of R^2 and p values amongst the soil chemical variables

	Aridity index R^2 (p)	Evapotranspiration R^2 (p)	Precipitation R^2 (p)	Elevation R^2 (p)	Slope R^2 (p)	TWI R^2 (p)	Plan C R^2 (p)	Profile C R^2 (p)	Log EC R^2 (p)
Aridity index	1								
Evapotranspiration	0.69 (0)	1							
Precipitation	-0.75 (0)	-0.95 (0)	1						
Elevation	-0.59 (0)	-0.68 (0)	0.61 (0)	1					
Slope	-0.46 (0)	-0.58 (0)	0.56 (0)	0.56 (0)	1				
TWI	0.33 (0)	0.37 (0)	-0.37 (0)	-0.42 (0)	-0.46 (0)	1			
Plan C	-0.17 (-0.0962)	-0.24 (0.02)	0.25 (0.02)	0.17 (0.09)	0.15 (0.14)	-0.29 (0)	1		
Profile C	-0.033 (0.74)	-0.044 (0.66)	0.09 (0.40)	-0.142 (0.16)	0.05 (0.61)	0.1 (0.31)	-0.28 (0.00)	1	
Log EC	0.57 (0)	0.56 (0)	-0.50 (0)	-0.51 (0)	-0.38 (0)	0.25 (0.01)	-0.14 (0.17)	-0.0049 (0.96)	1

Table 7.6 The statistics of the stepwise multiple regression soil salinity predictive model

Independent variable	coefficient	Std. error	t-value	Significance level
Constant	0.920952	0.329265	2.7970	0.0057
Elevation	-0.000305	0.000124	-2.4639	0.0146
TWI	0.031452	0.00942	3.3389	0.0010
Evapotranspiration	0.000668	0.000141	4.7236	0.0000

R^2 (adj) = 0.53), standard error = 0.31, mean absolute error = 0.24

Table 7.7 Analysis of variance for the stepwise regressions models based on TWI, elevation and evapotranspiration model

	Sum of squares	DF	Mean square	F-Ratio	p-value
Model	14.48	3	4.83	51.65	0.00
Error	18.60	199	0.09		

$R^2 = 0.44$, R^2 (adj.) = 0.43

The soil salinity map of the Sandspruit catchment based on the stepwise multiple linear regressions model is shown in Figure 7.9. The percentage of accurately predicted EC range values was computed using the EC of the groundwater hydro-census samples covering the Sandspruit catchment. The stepwise multiple linear regressions soil salinity model yielded an accuracy of 40.82%. This is higher than the groundwater and soil EC-based soil salinity predictive models. While the soil EC-based slope gradient and elevation soil salinity regression predictive models yielded accuracies of 39.13 and 8.68% respectively, the groundwater EC-based slope gradient percentage, elevation and TWI soil salinity predictive models yielded accuracies of 25.45, 16.36 and 34.69% respectively. The slightly higher prediction capacity of the multiple stepwise linear regressions models is most likely due to the fact that it uses more variables. The prediction of the stepwise multiple linear regressions soil salinity predictive model is consistent with the theory that low lying areas are more prone to salinity than elevated areas in the landscape.

Although the models presented here are based on reliable information, it is a question of scale. Information that was sampled at different scales or support cannot necessarily produce reliable secondary information. Comparing information that was sampled at different scales is therefore not advisable as different scales of information support different processes. As a concluding argument for this work, we strongly advise that the resolution or scales of information be comparable when mapping soil salinity.
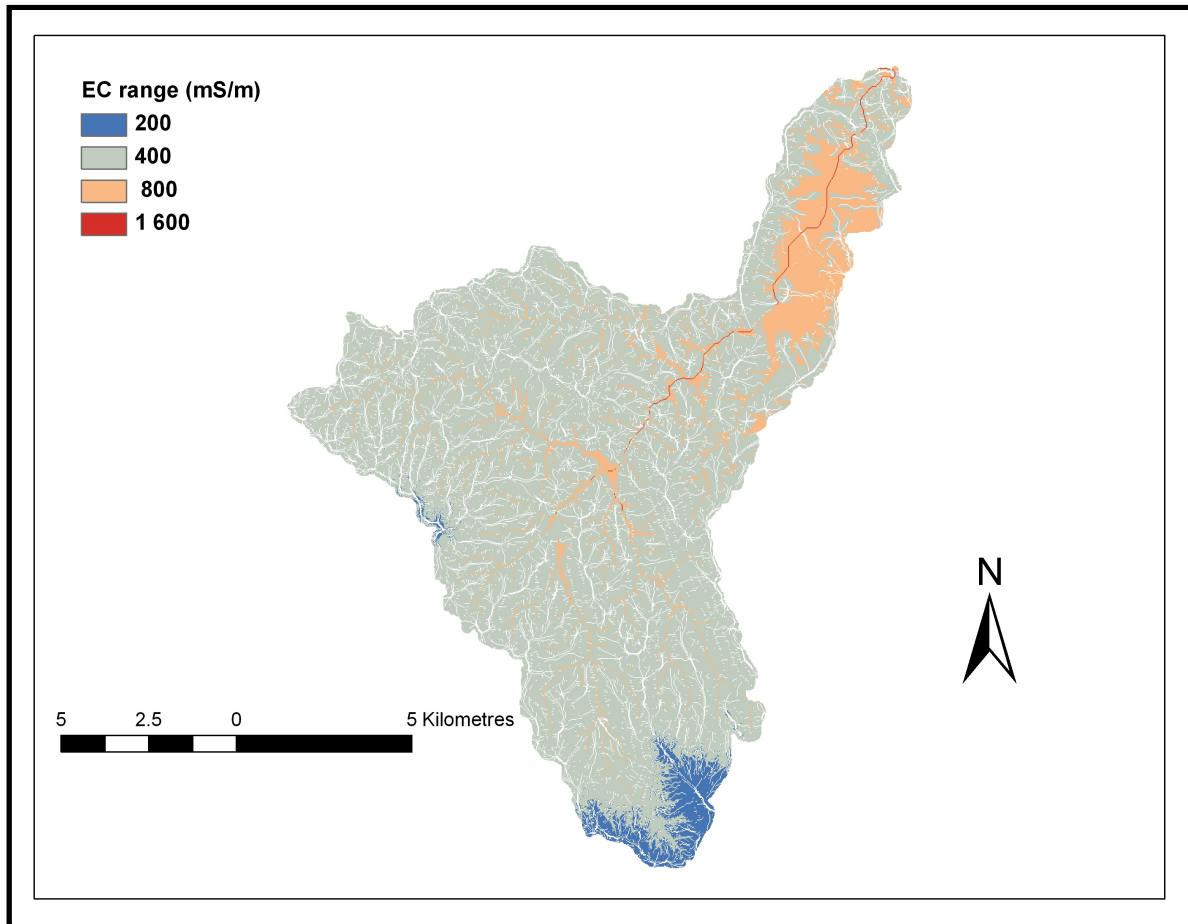
Figure 7.9 Sandspruit soil salinity map derived from stepwise multiple regression based on evapotranspiration, TWI and elevation

7.5 CONCLUSSIONS

This work investigated the potential of terrain attribute-based mapping of soil salinity at catchment scales. Terrain attributes-based soil salinity predictive models were computed using the EC of topsoil samples collected from within the Sandspruit catchment. In addition, terrain attributes-based soil salinity predictive models were developed using the EC of groundwater. Stepwise multiple linear regression soil salinity predictive models were also investigated. The stepwise multiple linear regressions predictive model based on TWI, elevation and evapotranspiration yielded the most accurate predictions of soil salinity. Promising soil salinity maps for the Sandspruit were produced using the slope-, elevation-, evapotranspiration and TWI-based soil salinity predictive models. The study established that wet areas in the landscape are more prone to soil salinity than drier areas. It appears that areas with slope gradient less than 4%

are highly susceptible to soil salinity. Despite less accurate predictions, we conclude that there is good potential for terrain attributes-based soil salinity predictive models. The results of this study lay a foundation for using terrain attributes-based predictive models to map soil salinity. Further investigations using more detailed sampling and higher-resolution DEMs are needed to improve the results. Also, object-based analysis for mapping soil salinity using terrain attributes should be explored.

CHAPTER 8

SYNTHESIS: IMPROVED SOIL SALINITY DETECTION, ANALYSIS AND MONITORING

8.1 BACKGROUND

The emergence of dryland salinity in the BRC will have devastating consequences for wheat production, irrigation agriculture and the supply of clean water in the region. Owing to insufficient funds available for research in South Africa, it will be restrictive in terms of costs to detect and monitor dryland salinity in the BRC using traditional wet chemistry methods. The literature review in this research exposed that remote sensing, NIR spectroscopy, hydrological parameters, land components and topography can be useful for the detection, analysis and mapping of soil salinity. The use of NIR spectroscopy will pave the way for analysing soil salinity cost effectively in the laboratory and *in situ*. NIR spectroscopy will also enhance our understanding of the spectral behaviour of salt-affected soils. Thus, such information is crucial for improving the use of hyperspectral remote sensing to map salt-affected soils remotely from satellite and airborne sensors. Hyperspectral remote sensing has the potential to overcome the limitations of broadband sensors for mapping salt-affected soils. Land components possess great valuable for studying soil properties. Soil properties (including soil salinity) can be mapped at local and regional scales using land components. The accuracy of hydrological parameters is also key to accurate modeling of soil salinity at catchment scales. So, it would be beneficial to evalue the accuracy of DEM-delineated streamlines and catchment boundaries. Finally, terrain attributes possess the ability to map soil salinity at local and regional levels. Thus NIR spectroscopy, hyperspectral remote sensing and DEM derivatives would enhance studying soil-affected soils in the laboratory, *in situ* and at local and regional scales.

Bearing in mind the above, the foundation for this study was to develop accurate, less tedious and cheaper techniques for detection and mapping soil salinity. The resultant techniques will be applied to detect and map soil salinity in the chosen study site (Sandspruit catchment of the BRC). These techniques will form the basis for enhanced monitoring of the soil salinity risk in the BRC and other areas where appropriate.

8.2 REVISTING THE OBJECTIVES OF THE STUDY

This study was aimed to evaluate the value of hyperspectral remote sensing, NIR spectroscopy, DEM-based hydrological parameters and land components, and terrain attributes for quantitative analysis, characterization, mapping and modelling of dryland salinity in the Sandspruit for enhanced monitoring. Six (6) objectives were set to realize the overall goal of this study. The degree to which each of the objectives was achieved is outlined in the following subsections.

8.2.1 Review of soil salinity analysis, detection and mapping techniques

The first objective was to review the literature to reveal the value of remote sensing, NIR spectroscopy, hydrological parameters, land components and terrain attributes for better monitoring of soil salinity. The literature established that although incorrect irrigation accounts for most of the human induced salinization, dryland salinization is increasingly contributing to the rising human caused salinization. It was exposed in the literature that broadband sensors are inadequate to study subtly occurrences of soil salinity. Owing to high spectral resolution, hyperspectral remote sensing is promising to enhance the mapping of soil salinity. In the case of NIR spectroscopy, the literature revealed that while it has potential to be used as a reliable, cheap and less labour intensive technique to quantitatively analyse soil salinity, it has not yet been widely adopted. Composite sampling was also exposed to reduce sampling related bias, thus can improve model calibration. Regarding land components, the literature exposed that they have good potential for mapping soil properties. However, traditional methods of delineating them are subjective, tedious and expensive. Land components can be delineated from DEMs with less costs and effort. But, the accuracy of land components derived from different DEM sources is not yet known. It is crucial to reveal the accuracy of DEM-delineated land components before they can be used for mapping soil properties. Consequently, the accuracy of land components derived from different DEMs has to be investigated first. In the case of hydrological parameters,

it was revealed that they have great value for improving the accuracy of soil salinity modeling from hydrological models. The accuracy of hydrological parameters extracted from DEMs should be investigated. Finally, the literature demonstrated that there is a relationship of soil salinity with topography. It was revealed that saline areas occur at low lying areas in the landscape. The literature also revealed that soil salinity is most likely to occur at wet areas and concave curvature profiles. Additionally, groundwater discharge areas occur at low lying areas and breaks of slope.

8.2.2 Hyperspectral remote sensing of saline soils

The second objective was to investigate the value of hyperspectral remote sensing to enhance the mapping of soil salinity by remote sensing. Individual bands, a NDSI, PLSR and bagging PLSR were investigated. Furthermore, field-based regression predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were developed. It was established that accurate predictions of soil salinity for dry soils can be made using untransformed reflectance individual band (at 2257 nm), a NDSI, PLSR using untransformed reflectance and bagging PLSR using first derivative reflectance predictive models. The predictive models yielded validation R^2 values of 0.85, 0.70, 0.65 and 0.60 based on bagging PLSR using first derivative reflectance, PLSR using untransformed reflectance, a NDSI and untransformed individual band at 2257 nm respectively. The study also concluded that reliable predictions of EC, pH, soluble Ca, Mg, Na, Cl and SO_4 in the field are possible using first derivative reflectance. The R^2 values for field-based predictive models for EC, pH, soluble Ca, Mg, Na, Cl and SO_4 were 0.85, 0.50, 0.65, 0.84, 0.79, 0.81 and 0.58 respectively. These findings are applicable to dry soils and can only be applicable for mapping soil salinity during dry seasons. Thus, this study affirmed that it is possible to improve soil salinity mapping using airborne and satellite hyperspectral data.

8.2.3 NIR spectroscopy of saline soils

The third objective was to investigate the value of NIR spectroscopy as a reliable, less labour intensive and cost effective alternative or complimentary method to analyse soil salinity. The results showed that reliable predictions of EC, pH, Ca, Mg, Na, Cl and SO_4 could be made based on PLSR regression predictive models. Owing to enhanced predictions of some soil chemicals, it is also concluded that the use of composite sampling possibly yielded superior PLSR predictive

models. The study also revealed that although MIR spectroscopy is reported to perform better than NIR spectroscopy , the NIR PLSR predictive models used in this study were more accurate than previously reported MIR results for some soil chemicals investigated. These findings affirm that soil salinity can be analysed accurately, cheaper and faster using NIR spectroscopy. NIR spectroscopy findings also confirmed that the most information on EC and salt minerals are found in the NIR and SWIR regions of the spectrum. Therefore, this information will also be useful for enhancing hyperspectral remote sensing of salt-affected soils.

8.2.4 The accuracy of DEM-delineated streamlines and catchment boundaries

The fourth objective evaluated the value of DEMs to delineate accurate hydrological parameters (streamlines and catchment boundaries). Accurate hydrological parameters will enhance the modeling of salinity at catchment scales using hydrological models. The research affirmed that higher resolution DEMs are required to derive correct hydrological parameters. The ED ratio developed in this study promises to be a valuable technique to compare DEM extracted raster datasets with reference datasets. In general, it appears as if usable hydrological information can be derived from the SRTM DEM and the ASTER GDEM2. This study established that streamlines and catchment boundaries delineated from DEMs generated from high resolution stereo-images can produce results comparable to those of LiDAR DEM's. Accurate streamlines and catchment boundaries are indispensable for improving the modelling of soil salinity using hydrological models. We believe that accurate hydrological parameters can enhance the modelling of soil salinity using hydrological models at local and regional scales.

8.2.5 The accuracy of DEM-delineated land components

The fifth objective evaluated the utility of DEMs to delineate accurate land components. The study concluded that the 5-m GEOEYE DEM was the most successful in producing land component boundaries that coincides with morphological discontinuities. The 5-m SUDEM L2 produced similar land components to the GEOEYE DEM and was almost as successful in maximizing internal (interclass) homogeneity and minimizing external (intraclass) homogeneity. The SRTM DEM appears to be somewhat superior to the ASTER GDEM2 for land component mapping. A novel measure, namely the MSGLV ratio, was developed and tested in this study for evaluating how well land component boundaries coincide with morphological discontinuities.

The MSGLV ratio measures the relationship between internal homogeneity and external heterogeneity of land components. The MSGLV ratio complimented the other validation techniques used in this study. It appears that the MSGLV has good potential to be used for evaluating the efficacy of DEMs to identify morphological discontinuities. Accurate land components delineated in a cost effective and objective manner provide new opportunities for mapping soil salinity at local and regional levels.

8.2.6 The value of terrain attributes to map soil salinity

Finally, the sixth objective investigated the value of mapping saline prone areas using terrain attributes. Terrain attributes-based soil salinity regression predictive models were developed based on the salinity of the soil and groundwater. Additionally, stepwise multiple linear regression soil salinity predictive models were investigated. The stepwise multiple linear regressions soil salinity predictive model based on TWI, elevation and evapotranspiration yielded more accurate soil salinity predictions than the predictive models based on TWI, slope gradient percentage and elevation. Sound relationships of slope, elevation and TWI with soil salinity were observed. Although not very accurate, sound potential salinity maps based on slope, elevation, evapotranspiration and TWI which identified areas at risk of soil salinity in the landscape for the Sandspruit catchment were produced. Although the accuracy is low, the study found that there is potential to map soil salinity at local and regional scales based on slope, elevation, evapotranspiration and TWI.

8.3 NEW DEVELOPMENTS

Traditional methods for analysing, detection and mapping of saline soils by remote sensing have been inadequate. While conventional analysis by wet chemistry methods is tedious and expensive, broadband sensors (for example SPOT and Landsat) are inadequate for mapping soil properties due to limited spectral resolution. This study established that there is potential to map soil salinity using airborne and satellite hyperspectral data. The study found that hyperspectral remote sensing has the possibility to enhance the detection of saline soils using an individual band (at 2257 nm), a novel NDSI, PLSR and bagging PLSR. Bagging PLSR using first derivative reflectance was established to be the most accurate method that has the potential to accurately detect slight to moderately salt-affected soils. With respect to NIR spectroscopy, this

study recognized that there is potential for it to be used as a less tedious and cost saving technique for quantifying soil salinity and soluble ions in the soil. Furthermore, this study established that although useful land components and hydrological parameters can be delineated from freely available high- to medium-resolution DEMs, very high-resolution DEMs are needed to delineate accurate land components and hydrological parameters. Two novel indices were also developed, namely an ED index for comparing raster data and a MSGLV ratio for validating the ability of DEMs to detect terrain morphological discontinuities. Regarding the value of terrain attributes to map soil salinity, it was demonstrated that there is potential to predict the risk of dryland salinity using slope-, elevation- and TWI-based regression predictive models. A stepwise multiple linear regressions soil salinity predictive model based on evapotranspiration, TWI and elevation was found to predict soil salinity more accurately than the slope-, elevation- and TWI-based regression predictive models. Sound dryland salinity risk maps for the Sandspruit catchment were produced using regression predictive models based on elevation, TWI and slope. Overall, this study demonstrated that NIR spectroscopy, hyperspectral remote sensing, land components, hydrological parameters and terrain attributes can enhance the detection of salt-affected soils in the laboratory, *in situ* and at local and regional scales.

8.4 DIRECTIONS FOR FURTHER RESEARCH

The hyperspectral individual bands, NDSI, PLSR and bagging PLSR predictive models developed in this study were tested under controlled conditions. This work should be expanded and evaluated using aerial and satellite data. Also, these predictive models were developed based on soil samples of varying properties covering the whole of South Africa. Further research should be conducted using soil samples at local and regional scales. It was established in this research that accurate land components and hydrological parameters can be extracted from DEMs. Research should be conducted to investigate the value of land components to study soil salinity and their relationship to soil properties. Land components have the potential enhance the detection of soil salinity using vegetation indices, and more research is needed to expose this value. Object-based methods for mapping soil salinity using land components should be explored. The utility of DEM-derived hydrological parameters to enhance soil salinity modeling using hydrological models should also be investigated. The potential for NIRS spectroscopy to quantify soil salinity was demonstrated in this study using a limited number of soil samples,

more research is needed to verify the findings. The data used in this study to investigate the value of terrain attributes to map soil salinity was not adequate. More research should be done using more detailed sampling and high-resolution DEMs to improve the results.

REFERENCES

Abbas A, Khan S, Hussain N, Hanjra MA, Akbar S. In press. Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Physics and Chemistry of the Earth*.

Acosta JA, Faz A, Jansen B, Kalbitz K, Martínez-Martínez S. 2011. Assessment of salinity status in intensively cultivated soils under semiarid climate, Murcia, SE Spain. *Journal of Arid Environments* 75: 1056-1066.

Adediran AO, Parcharidis I, Poscolieri M, Pavlopoulos, K. 2004. Computer-assisted discrimination of morphological units on north-central Crete (Greece) by applying multivariate statistics to local relief gradients. *Geomorphology* 58: 357-370.

Akramkhanov A, Martius C, Park SJ, Hendrickx JMH. 2011. Environmental factors of spatial distribution of soil salinity on flat irrigated terrain. *Geoderma* 163: 55-62.

Ardianyah PO, Yokoyawa R. 2002. DEM generation method from contour lines based on the steepest slope segment chain and monotone interpolation function. ISPRS *Journal of Photogrammetry and Remote Sensing* 57: 86-101.

Al-Khaier F. 2003. Soil salinity detection using satellite remote sensing**.** MSc thesis, International Institute for Geo-information Science and Earth Observation, and Utrecht University, Netherlands.

Allison GB, Cook PG, Barnett SR, Walker GR, Jolly ID, Hughes MW. 1990. Land clearance and river salinization in the western Murray basin. *Australian Journal of Hydrology* 119: 1-20.

Allison LE, Brown JW, Hayward HE, Richards LA, Berstein L, Fireman M, Pearson GA, Wilcox LV, Bower CA, Hatcher JT, Reeve RC. 1969. Diagnosis and improvement of saline and alkali soils, United States Department of Agriculture, Washington.

Andersson JCM, Zehnder AJB, Rockström J, Yang H. 2011 Potential impacts of water harvesting and ecological sanitation on crop yield, evaporation and river flow regimes in the Thukela River basin, South Africa. *Agricultural Water Management* 98: 1113-1124.

Argialas DP. 1995. Towards structured-knowledge models for landform representation. *Zeitschrift für Geomorfologie N.F* 101: 85-108.

ASTER GDEM Validation Team. 2011. Aster global digital elevation model version 2 – Summary of validation results. NASA land Processes Distributed Archive Centre and Joint Japan – US Aster Team.

ASTER GDEM. 2011. [Online]. Available: http://www.gdem.aster.ersdc.or.JP. Date accessed: 20 July 2011.

Barker ME, Weller DE, Jordan TE. 2006. Comparison of automated watershed delineations: Effects on land cover areas, percentages, and relationships to nutrient discharge. *Photogrammetric Engineering and Remote Sensing* 72: 159-168.

Barrett-Lennard EG, Nulsen RA. 1989. Dryland soil salinity—cure, containment or catastrophe. In: *Proceedings of the 5th Australian Agronomic Conference,* 24-29 September, Perth, Australia: The University of Western Australia. pp 304-311.

Barrett-Lennard EG, Nulsen, RA, 1989. Dryland soil salinity cure, containment or catastrophe? In: Proceedings of the 5th Australian Agronomy Conference, 24-29 September, Perth, Australia: The University of Western Australia. pp 212-220.

Beherens T, Zhu A-X, Schmidt K, Scolten T. 2010. Multi-scale digital terrain analysis and feature selection for digital soil mapping. *Geodema* 155: 175-185.

Bellon-Maurel V, Fernandez-Ahumada E, Palagos B, Roger, JM, MacBratney A. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the

prediction of soil attributes by NIR spectroscopy. *Trends in Analytical Chemistry* 29: 1073-1081.

Bellon-Maurel V, MacBratney A. 2011. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils e Critical review and research perspectives. *Soil Biology and Biochemistry* 43: 1398-1410.

Ben-Dor E, Banin A. 1994. Visible and near-infrared (0.4-1.1 µm) analysis of arid and semiarid soils. *Remote Sensing of Environment* 48: 261-274.

Ben-Dor E, Chabrilat S, Dematte JAM, Taylor GR, Hill J, Whiting ML, Sommer S. 2009. Using imaging spectroscopy to study soil properties. *Remote Sensing of Environment* 113: 538-555.

Ben-Dor E, Patkin K, Banin A, Karnieli A. 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data – a case study over clayey soils in Israel. *International Journal of Remote Sensing* 23: 1043-1062.

Ben-Dor E, Taylor RG, Hill J, Dematte JAM, Whiting ML, Chabrillat S, Sommer S. 2008. Imaging spectroscopy for soil applications. *Advances in Agronomy* 97: 321-392.

Bennet SJ, Barret-Lennard EG, Colmer TD. 2009. Salinity and waterlogging as constraints to saltland and pasture production: A review. *Agricultural Ecosystems and Environment* 129: 349-360.

Bertel L, Deronde B, Fernandez M, Kempeneers P, Knaeps E, Meuleman K, Reusen I, Ruddick K, Sterckx S, Trefois, P, Mol V. 2006. Hyperteach: Training in imaging spectroscopy. AFRICAN SUN MeDia, Stellenbosch.

Bhaskaran S, Paramananda S, Ramnarayan M. 2010. Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data. *Applied Geography* 30: 650-665.

Bikindou FDA, Gomat HY, Deleporte P, Bouillet J-P, Mouikini R, Mbedi Y, Ngouaka E, Brunet D, Sita S, Diazenza J-B, Vouidibio J, Mareschal L, Ranger J, Saint-Andre L. 2012. Are

NIR spectra useful for predicting site indices in sandy soils under Eucalyptus stands in Republic of Congo? *Forest Ecology and Management* 266: 126-137.

Bilgili AV, Van Es HM, Akbas F, Durak A, Hively WD. 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments* 74: 229-238.

Bilgili AV, van Es HM, Akbas F, Durak A, Hively WD. 2010. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments* 74: 229-238.

Blaschke T, Stobl J. 2003. Defining landscape units through integrated morphometric characteristics. In: Buhmann E, Ervin S (eds.), *Landscape Modeling: Digital Techniques for Landscape Architecture*. Heidelberg: Wichmann Verlag. pp 104-113.

Blascke T. 2010. Object-based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing* 65: 2-16.

Bock M, Xofis P, Mitchley J, Rossner G, Wissen M. 2005. Object-oriented methods for habitat mapping at multiple scales – Case studies from Northern Germany and Wye Downs, UK. *Journal for Nature Conservation* 13: 75-89.

Bouaziz M, Matshullat J, Gloaguem R. 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *Comptes Rendus Geoscience* 343: 795-803.

Brown DJ, Shephered KD, Walsh MG, Mays MD, Reinsch TG. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 132: 273-290.

Bruker Optics. 2011. [Online]. Available: http://www.brukeroptics.com/ft. Date accessed: 05 July 2011.

Brus DJ, De Gruijter JJ. 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil. *Geoderma* 80: 1-44.

Callow JN, Van Niel KP, Boggs GS. 2007. How does modifying a DEM to reflect known hydrology affect subsequent terrain analysis? *Journal of Hydrology* 332: 30-39.

Campbell JB. 2007. *Introduction to remote sensing (4th edn)*. New York: The Guilford Press.

Cecillon L, Certini G, Lange H, Forte C, Strand LT. 2012. Spectral fingerprinting of soil organic matter composition. *Organic Chemistry* 46: 127-136.

CGIAR-CSI. 2011. [Online]. Available: http://srtm.csi.cgiar.org. Date accessed: 20 July 2011.

Chang CI. 2003. *Hyperspectral imaging: techniques for spectral detection and classification*. New York: Kluwer Academic/Plenum Publishers.

Chang C-W, Laird DA, Mausbach MJ, Hurburgh Jr. CR. 2001. Near-infrared reflectance spectroscopy—principal components regression analysis of soil properties. *Soil Science Society of America Journal* 65: 480-490.

Charrier R, Li Y. 2012. Assessing resolution and source effects of digital elevation models on automated floodplain delineation: A case study from the Camp Creek Watershed, Missouri. *Applied Geography* 34: 38-46.

Chaubey I, Cotter AS, Costello TA, Saorens TA. 2005. Effect of DEM data resolution on SWAT output uncertainty. *Hydrological Processes* 19: 621-625.

Cho MA, Skidmore A, Corsi F, Van Wieren SE, Sobhan I. 2007. Estimating of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation* 9: 414-424.

Christofides TC. 2003. Randomized response in stratified sampling. *Journal of Statistical Planning and Inference* 128: 303-310.

Clarke CJ, George RJ, Bell RW, Hobbs RJ. 1998. Major faults and the development of dryland salinity in the Western Wheatbelt of Western Australia. *Hydrological Earth Systems Science* 2: 77-91.

Collell C, Gou P, Arnau J, Comaposada J, 2011. Non-destructive estimation of moisture, water activity and NaCl at ham surface during resting and drying using NIR spectroscopy. *Food Chemistry* 129: 601-607.

Cox JW, McFarlane DJ. 1995. The causes of waterlogging in shallow soils and their drainage in southwestern Australia. *Journal of Hydrology* 167: 175-194.

Cozzolino D, Moron A. 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. *Journal of Agricultural Sciences* 140: 65-71.

Cozzolino D, Cynkar WU, Shah N, Smith P. 2011. Multivariate data analysis applied to spectroscopy: Potential applications to juice and fruit quality. *Food Research International* 44: 1888-1896.

Csillag F, Pasztor L, Biehl L. 1993. Spectral Band selection for the characterization of salinity status of soils. *Remote sensing of environment* 43: 231-242.

Daniel KW, Tripathi NK, Honda K. 2003. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Australian Journal of Soil Research* 41: 47-59.

de Clercq WP, Fey MV, Jovanovic NZ. 2010. Land use impacts on salinity in the Berg River catchment. Report No. 1503/1/10. Pretoria: Water Research Commission.

de Clercq WP, Van Meirvenne M, Fey MV. 2009. Prediction of the soil-depth salinity-trend in a vineyard after sustained irrigation with saline water. *Agricultural Water Management* 96: 395-404.

Dehaan RT, Taylor GR. 2002. Field-derived spectra of salinized soils and vegetation indicators of irrigation induced soil salinization. *Remote Sensing of Environment* 80: 406-417.

Dehaan RT, Taylor GR. 2003. Image-derived endmembers as indicators of salinization. *International Journal of Remote Sensing* 24: 775-794.

Delbart N, Kergoat L, Toan TL, Lhernitte J, Picard G. 2005. Determination of phenological dates in boreal regions using normalized difference water index. *Remote Sensing of Environment* 97: 26-38.

Dematte JAM, Campos RC, Alves MC, Firio PR, Nanni MR. 2004. Visible–NIR reflectance: a new approach on soil evaluation. *Geoderma* 121: 95-112.

Deng Y. 2007. New trends in digital terrain analysis: landform definition, representation, and classification. *Progress in Physical Geography* 31: 405-419.

Dent D. 2007. Environmental geophysics mapping salinity and water resources. *International Journal of Applied Earth Observation and Geoinformation* 9: 130-136.

Ding J, Wu M, Tashpolat T. 2011. Study on soil salinization information in arid region using remote sensing technique. *Agricultural Sciences in China* 10: 404-411.

Dixon B, Earls J. 2009. Resample or not? Effects of resolution of DEMs in watershed modeling. *Hydrological Processes* 23: 1714-1724.

Dong Y-W, Yang S-Q, Xu C-Y, Li Y-Z, Bai W, Fan Z-N, Wang Y-N, Li, Q-Z. 2011. Determination of soil parameters in apple-growing regions by Near- and Mid-Infrared Spectroscopy. *Pedosphere* 21: 591-602.

Douaoui AEK, Nicolas H, Walter C. 2006. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* 134: 217-230.

Dowling T, Summerell GK, Walker J. 2003. Soil wetness as an indicator of stream salinity: a landscape position index approach. *Environmental Modelling and Software* 18: 587-593.

Drăgut L, Tiede D, Levick SR. 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24: 859-871.

Drăgut L, Blaschke T. 2006. Automated classification of landform elements using object-based image analysis. *Geomorphology* 81: 330-344.

Drăgut L, Eisank C, Strasser T. 2011. Local variance for multi-scale analysis in geomorphometry. *Geomorphology* 130: 162-172.

Drăgut L, Shauppenlehne T, Muhar, T, Strobl J, Blaschke T. 2009. Optimization of scale and parameterization for terrain segmentation: An application to soil-landscape modeling. *Computers and Geosciences* 35: 1865-1883.

Drăgut L, Tiede D, Levick SR. 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science* 24: 859-871.

Drăgut L, Eisank C. 2011. Object representations at multiple scales from digital elevation models. *Geomorphology* 129: 183-189.

Drake NA. 1995. Reflectance spectra of evaporite minerals (400-500 nm): Applications for remote sensing. *International Journal of Remote Sensing* 16: 2555-2571.

Draper NR, Smith H. 1966. *Applied Regression Analysis*. New York: Wiley.

Dutkiewicz A. 2006. Evaluating hyperspectral imagery for mapping the surface symptoms of dryland salinity. PhD thesis, University of Adelaide, Australia.

Dwivedi RS. 1992. Monitoring and the study of the effects of image scale on delineation of salt-affected soils in the Indo-Gangetic plains. *International Journal of Remote Sensing* 13 (8): 1527-1536.

Esbensen KH, Paasch-Mortensen P. 2010. Process sampling: theory of sampling – the missing link in process analytical technologies (PAT). In: Bakeev K (ed.) *Process sampling: theory of sampling – the missing link in process analytical technologies (PAT)*. John Wiley and Sons Ltd. pp 37.

Evans R, Brown C, Kellet J, 1990. In: Mackay N, Eastburn D (eds.). *Geology and groundwater*. Canberra: The Murray–Darling Basin Commission. pp 77–93.

Farifteh J, Farshad A, George RJ. 2006. Assessing salt-affected soils using remote sensing, solute modelling, and geophysics. *Geoderma* 130: 191-206.

Farifteh J, Van der Meer F, Atzberger C, Carranza EJM. 2007. Quantitative analysis of salt-affected soil reflectance spectra: a comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment* 110: 59-78.

Farifteh J, Van der Meer F, Van der Meijde M, Atzberger C. 2008. Spectral characteristics of salt-affected soils: A laboratory experiment. *Geoderma* 145: 196-206.

Farifteh J. 2007. Imaging spectroscopy of salt-affected soils: Model-based integrated method. PhD thesis, International Institute for Geo-information Science and Earth Observation, and Utrecht University, Netherlands.

Farr T, Kobrick M, 2001. The shuttle radar topography mission. *American Geophysical Union EOS* 81: 583-585.

Farr TG. 2000. The shuttle radar topography mission. *IEEE Aerospace Conference Proceedings* 63.

Fernandez-Buces N, Siebe C, Cram S, Palacio JL. 2006. Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. *Journal of Arid Environments* 65: 644-667.

Fetter CW, 2001. *Applied Hydrogeology (4th edn)*. Upper Saddle River: Prentice-Hall.

Fey MV, de Clercq WP. 2004. A pilot study investigating the role of dryland salinity and the quality of the water of Berg River. Report No. 1342/1/04 (IASB No. 1770053059). Pretoria: Water Research Commission.

Fey MV, de Clercq WP. 2004. Dryland salinity impacts on Western Cape rivers. Report No. 1342/1/04. Pretoria: Water Research Commission.

Fitzpatric E.A. 1993. *Soils: Their Formation and Distribution*. London: Longman.

Fitzpatric EA. 1980. Soils: Their formation and distribution, Longman, London.

Flügel WA. 1995. River salinity due to dryland agriculture in the Western Cape Province, Republic of South Africa. *Environmental International* 21: 679-686.

Freeze RA, Cherry JA, 1979. *Groundwater*. Englewood Cliffs: Prentice-Hall.

Frey H, Paul F. In press. On the suitability of the SRTM DEM and ASTER GDEM for the compilation of topographic parameters in glacier inventories. *International Journal of Applied Earth Observation and Geoinformation*.

Fuentes M, Hidalgo C, Gonzalez-Martin I, Hernandez-Hierro JM, Govaerts B, Sayre KD, Etchevers J. 2012. NIR spectroscopy: An alternative for soil analysis. *Communication in Soil Science and Plant Analysis* 43: 346-356.

Gao B. 1996. NDWI – A normalized difference water index for remote sensing of vegetation from space. *Remote Sensing of Environment* 58: 257-266.

Gao F, Huang Q, Sun X, Yan Z. 2011. Study of dynamic changes of the soil salinization in the upper stream of the Tarrim River based on RS and GIS. *Procedia Environmental Sciences* 11: 1135-1141.

Ge Y, Morgan CLS, Grunwald S, Brown DJ, Sarkhot D. 2011. Comparison of soil reflectance spectra and calibration models obtained using multiple spectrometers. *Geoderma* 161: 202-211.

Geological Survey of South Africa. 1984. Geological map of the Republics of South Africa, Transkei, Bophuthatswana, Venda and Ciskei and the Kingdoms of Lesotho and Swaziland. Pretoria: Council for Geoscience.

George C, Rowland C, Gerard F, Baltter H. 2006. Retrospective mapping of burnt areas in central Siberia using a modification of the normalized difference water index. *Remote Sensing of Environment* 104: 346-359.

Gichamo TZ, Popescu I, Jonoski A, Solomatine D. 2012. River cross-section extraction from the ASTER global DEM for flood modelling. *Environmental Modelling and Software* 31: 37-46.

Gomez C, Viscarra Rossel RA, McBratney AB. 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma* 146: 403-411.

Graff LH, Usery EL. 1993. Automated classification of generic terrain features in digital elevation models. *Photogrammetric Engineering and Remote Sensing* 59: 1409-1417.

Greiner R. 1998. Catchment management for salinity control: Model analysis for the Liverpool Plains in New South Wales. *Agricultural Systems* 56: 225-251.

Gutierrez M, Johnson E. 2010. Temporal variations of natural soil salinity in an arid environment using satellite images. *Journal of South American Earth Sciences* 30: 46-57.

Gyasi-Agyei Y, Willgoose G, Detroch FP. 1995. Effects of vertical resolution and map scale of digital elevation models on geomorphological parameters used in hydrology. *Hydrological Processes* 9: 363-382.

Hansen G, Schjoerring JK. 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sensing of Environment* 86: 542-553.

He Y, Huang M, Garcia A, Hernandez A, Song H. 2007. Prediction of soil macronutrients content using near-infrared spectroscopy. *Computers and Electronics in Agriculture* 58: 144-153.

Herold M, Roberts DA, Gardner ME, Dennison PE. 2004. Spectrometry for urban area remote sensing - Development of a spectral library from 350 to 2400 nm. *Remote Sensing of Environment* 91: 304-319.

Hughes JD, Crosbie RS, Van de Ven RJ. 2008. Salt mobilisation process from salinized catchment featuring a perennial stream. *Journal of Hydrology* 362: 308-319.

Inoue Y, Penuelas J, Miyata A, Mano M. 2007. Normalized difference spectral indices for estimating photosynthetic efficiency and capacity at a canopy scale derived from hyperspectral and CO_2 flux measurements in rice. *Remote Sensing of Environment* 112: 156-172.

Islam K, Singh B, McBratney AB. 2003. Simultaneous estimation of various soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* 41: 1101-1114.

Iwai CB, Oo AN, Topark-ngarm B. 2012. Soil property and microbial activity in natural salt affected soils in an alternating wet–dry tropical climate. *Geoderma* 189-190: 144-152.

Janik LJ, Forrester ST, Rawson A. 2009. The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial-least squares regression and neural

networks (PLS-NN) analysis. *Chemometrics and Intelligent Laboratory Systems* 97: 179-188.

Janik LJ, Merry RH, Skjemstad JO. 1998. Can mid infra-red diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture* 38: 681-696.

Jenny H. 1941. *Factors of soil formation*. New York: McGraw-Hill.

Jian-li D, Man-chun W, Tiyip T. 2011. Study on soil salinization information in arid region using remote sensing technique. *Agricultural Sciences in China* 10: 404-411.

Jins S, Sader SA. 2004. Comparison of time series tasselled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment* 94: 364-372.

Jones KH. 1998. A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers and Geosciences* 24: 315-323.

Katerji N, Van Hoorn JW, Hamdy A, Mastrorilli M. 1998. Response of tomatoes, in a crop of indeterminate growth, to soil salinity. *Agricultural Water Management* 38: 59-68.

Khan NM, Rastoskuev VV, Sato Y, Shiozawa S. 2005. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agricultural Water Management* 77: 96-109.

Kienzle S. 2004. The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS* 8: 83-111.

Kim J, Sungur EA, Heo T. 2006. Calibration approach in stratified random sampling. *Statistics and Probability Letters* 77: 99-103.

Kingswell R, John M. 2007. The influence of farm landscape shape on the impact and management of dryland salinity. *Agricultural Water Management* 89: 29-38.

Kirby SD, 1996. Integrating a GIS with an expert system to identify and manage dryland salinity. *Applied Geography* 16: 289-303.

Kitamura Y, Yano T, Honna T, Yamamoto S, Inosako K. 2006. Causes of farmland salinization and remedial measures in the Aral Sea basin—Research on water management to prevent secondary salinization in rice-based cropping system in arid land. *Agricultural Water Management* 85: 1-14.

Klug W, Graziani G, Grippa G, Pierce D, Tassone C (eds). 1992. *Evaluation of long range atmospheric transport models using environmental radioactivity data from the Chernobyl accident: the ATMES Report*. London: Elsevier.

Laliberte AS, Rango A, Havstad KM, Paris JF, Beck RF, Mcneely R, Gonzalez AL. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* 93: 198-210.

Land Type Survey Staff. 1972-2003. Land Types of South Africa. Pretoria: ARC-Institute for Soil Climate and Water.

Legesse D, Vallet-Coulomb C, Gasse F. 2003. Hydrological response of a catchment to climate and land use changes in Tropical Africa: case study South Central Ethiopia. *Journal of Hydrology* 275: 67-85.

Lenny MP, Woodstock CE, Collins JB, Hamdy H. 1996. The status of agricultural lands in Egypt: The use of multispectral NDVI features derived from Landsat TM. *Remote Sensing of Environment* 56: 8-20.

Li J, Wong DWS. 2010. Effects of DEM sources on hydrologic applications. *Computers, Environment and Urban Systems* 34: 251-261.

Liang X, Guo J, Leung R. 2004. Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology* 298**:** 287-310.

Lillesand TM, Keifer RW, Chipman JW. 2004. *Remote sensing and image interpretation (5th edn).* New York: John Wiley and Sons, Inc.

MacMillan RA, Jones RK, McNabb DH. 2004. Defining a hierarchy of spatial entities for environmental analysis and modeling using digital elevation models (DEMs). *Computers, Environment and Urban Systems* 28: 175-200.

Maianu A. 1984. Twenty years of research on reclamation of salt-affected soils in Romanian rice fields. *Agricultural Water Management* 9: 245-256.

Malins G, Metternicht G, 2006. Assessing the spatial extent of dryland salinity through fuzzy modeling. *Ecological Modelling* 193: 387-411.

Mancas M, Gosselin B, Macq B. 2005. Segmentation using a region-growing thresholding. In: Proceedings of the Electronic Imaging Conference of the International Society for Optical Imaging (SPIE/EI 2005), San Jose, USA.

Martens HA, Dardenne P, 1998. Validation and verification of regression in small data sets. *Chemometrics and Intelligent Laboratory Systems* 44: 99-121.

Martínez-Sánchez MJ, Pérez-Sirvent C, Molina-Ruiz J, Tudela ML, García-Lorenzo ML. 2011. Monitoring salinization processes in soils by using a chemical degradation indicator. *Journal of Geochemical Exploration* 109: 1-7.

Martz L, De Jong E. 1998. CATCH: A FORTRAN program for measuring catchment area from digital elevation models. *Computers and GeoSciences* 14: 627-640.

Mashimbye ZE, Cho MA, Nell JP, De Clercq WP, Van Niekerk A, Turner DP. 2012. Model-based integrated methods for quantitative estimation of soil salinity from hyperspectral remote sensing data: A case study of selected South African soils. *Pedosphere* 22: 640-649.

Mashimbye ZE. 2005. Remote sensing-based identification and mapping of salinised irrigated land between Upington and Keimoes along the lower Orange River, South Africa. MSc thesis, Stellenbosch University, South Africa.

Mathieu R, Freeman C, Aryal J. 2007. Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. *Landscape and Urban Planning* 81: 179-192.

McBratney AB, Mendonca Santos ML, Minasny B. 2003. On digital soil mapping. *Geodema* 117: 3-52.

Mckenzie NJ, Ryan PJ. 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89: 67-94.

Melendez-Pastor I, Pedreno JN, Koch M, Gomez I. 2010. Applying imaging spectroscopy techniques to map saline soils with ASTER images. *Geoderma* 158: 55-65.

Metternicht G. 2001. Assessing temporal and spatial changes of salinity using fuzzy logic, remote sensing and GIS: Foundations of an expert system. *Ecological Modelling* 144: 163-179.

Metternicht GI, Zinck JA. 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sensing of Environment* 85: 1-20.

Metternicht GI. 2003. Categorical fuzziness: a comparison between crisp and fuzzy class boundary modelling for mapping salt-affected soils using Landsat TM data and a classification based on anion ratios. *Ecological Modelling* 168: 371-389.

Minar J, Evans IS. 2008. Elementary forms for land surface segmentation: the theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology* 95: 236-359.

Mirlas V. 2012. Assessing soil salinity hazard in cultivated areas using MODFLOW model and GIS tools: A case study from Jezre'el Valley, Israel. *Agricultural Water Management* 109: 144-154.

Moller M, Volk M, Friederich K, Lymburner L. 2008. Placing soil genesis and transport processes into landscape context: a multiple terrain analysis approach. *Journal of Plant and Soil Science* 171: 419-430.

Molnár DK, Julien PY. 2000. Grid-size effects on surface runoff modeling. *Journal of Hydrologic Engineering* 5: 8-16.

Mukherjee S, Joshi PK, Mukherjee S, Ghosh A, Garg RD, Mukhopadhyyaya A. 2013. Evaluation of vertical accuracy of open source digital elevation model (DEM). *International Journal of Applied earth Observation and Geoinformation* 21: 205-217.

Mulder VL, De Bruin S, Schaepman ME, Mayr TR. 2011. The use of remote sensing in soil and terrain mapping - A review. *Geoderma* 162: 1-19.

Nell JP. 2009. The primary salinity, sodicity, and alkalinity status of South African soils. PhD thesis, University of the Free State, South Africa.

O'Callaghan JF, Mark DM. 1984. The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics and Image Processing* 28: 323-344.

Pannell DJ, Ewing MA. 2005. Managing secondary dryland salinity: Options and challenges. *Agricultural Water Management* 80: 41-56.

Patel AD, Bhensdadia H, Pandey AN. 2009. Effect of salinization of soil on growth, water status and general nutrient accumulation in seedlings of *Delonis regia* (Fabaceae). *Acta Eclogica Sinica* 29: 109-115.

Perica S, Foufoula-Georgiou E. 1996. Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *Journal of Geophysical Research* 101: 26347-26361.

Pontius RG, Boersma W, Castella J-C, Clarke K, De Nijs T, Dietzel C, Duan Z, Fotsing E, Goldstein N, Kok K, Koomen E, Lippitt CD, McConell W, Sood AM, Pijanowski B, Pithadia S, Sweeny S, Trung TN, Veldkamp AM, Verburg PH. 2008. Comparing the input, output, and validation maps for several models of land change. *Annals of Regional Science* 42: 11-37.

Prima ODA, Echigo A, Yokoyama R, Yoshida T. 2006. Supervised landform classification of Northeast Honshu from DEM-derived thematic maps. *Geomorphology* 78: 373-386.

Quinn P, Beven K, Chevallier P, Planchon O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes* 5: 59-79.

Rao BRM, Dwivedi RS, Venkataratnam L, Ravishakar T, Thammappa SS. 1991. Mapping the magnitude of sodicity in part of the Indo-Gangetic plains of Uttar Pradesh, northern India using Landsat-TM data. *International Journal of Remote Sensing* 12: 419-425.

Reeves III JB, Mcarty GW, Meisnger JJ. 1999. Near infrared reflectance spectroscopy for the analysis of agricultural soils. *Journal of Near Infrared Spectroscopy* 7: 179-193.

Reeves III JB, Mcarty GW. 2001. Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and fibre-optic probe. *Journal of Near Infrared Spectroscopy* 9: 25-34.

Renssen H, Knoop JM. 2000. A global river routing network for use in hydrological modeling. *Journal of Hydrology* 230: 230-243.

Rietz DN, Haynes RJ. 2003. Effects of irrigation-induced salinity and sodicity on soil microbial activity. *Soil Biology and Biochemistry* 35: 845-854.

Robertson MJ, Measham TG, Batchelor G, George R, Kingwell R, Hosking K. 2009. Effectiveness of a publicly-funded demonstration program to promote management of dryland salinity. *Journal of Environmental Management* 90: 3023-3030.

Rodriguez E, Morris CS, Belz JE, Chaplin EC, Martin JM, Daffer W, Hensley S. 2005. *An assessment of the SRTM topographic products*. Pasadema: Jet Propulsion Laboratory.

Rongjiang Y and Jingsong Y. 2010. Quantitative evaluation of soil salinity and its spatial distribution using electromagnetic induction. *Agricultural Water Management* 97: 1961-1970.

Rowel DL 1994. *Soils Science: Methods and applications*. London: Prince Hall.

Saha K, Well NA, Munro-Stasiuk M. In press. An object-oriented approach to automated landform mapping: A case study of drumlins. *Computers and Geosciences*.

Salama RB, Farrington P, Bartle GA, Watson GD, 1993. The role of geological structures and relic channels in the development of dryland salinity in the wheatbelt of Western Australia. *Australian Journal of Earth Sciences* 40: 45-56.

Sankey JB, Brown DJ, Benard ML, Lawrence RL. 2010. Comparing local vs. global visible and near-infrared (VIS-NIR) diffuse reflectance spectroscopy (DRS) calibrations for the prediction of soil clay, organic C and inorganic C. *Geoderma* 148: 149-158.

Sarkhot DV, Grunwald S, Ge Y, Morgan CLS. 2011. Comparison and detection of total and available soil carbon fractions using visible/near infrared diffuse reflectance spectroscopy. *Geoderma* 164: 22-32.

Seyler F, Muller F, Cochonneau G, Guimareas L, Guyot JL. 2009. Watershed delineation for the Amazon sub-basin system using GTOPO30 DEM and a drainage network extracted from JERS SAR images. *Journal of Hydrology* 23**:** 3173-3185.

Shafique M, Van der Meijde M, Kerle N, Van der Meer F. 2011. Impact of DEM source and resolution on topographic seismic amplification. *International Journal of Applied Erath Observation and Geoinformation* 13: 420-427.

Sharma RC, Bhargarva GP. 1988. Landsat imagery for mapping saline soils and wetlands in north-west India. *International Journal of Remote Sensing* 9: 39-44.

Shepherd KD, Walsh MG. 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal* 66: 988-998.

Shepherd KD, Walsh MG. 2002. Infrared spectroscopy- enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *Journal of Near Infrared Spectroscopy* 15: 1-19.

Shibusawa S, Imade SW, Amon SS, Sato SS, Hirako S. 2001. Soil mapping using real-time spectrometer. In: Greiner G, Blackmore S (eds.), Third European Conference on Precision Agriculture volume 1. Agro Montpellier. pp 497-508.

Siart C, Bubenzer O, Eitel B. 2009. Combining digital elevation data (SRTM/ASTER), high resolution satellite imagery (Quickbird) and GIS for geomorphological mapping: A multi-component case study on Mediterranean karst in Central Crete. *Geomorphology* 112: 106-121.

Simin C, Rongqun Z, Liming L, De Z. 2010. A method of salt-affected soil information extraction based on a support vector machine with texture features. *Mathematical and Computer Modelling* 51: 1319-1325.

Slavich PG, Petterson GH, Grieffin D. 2002. Effects of irrigation water salinity and sodicity on infiltration and Lucerne growth over a shallow watertable. *Australian Journal of Experimental Agriculture* 42: 281-290.

Smith G, Beare M, Boyd M, Downs T, Gregory M, Morton D, Brown N, Thompson A. 2007. Land cover map production through generalization of OS mastermap. *The Cartographic Journal* 44: 276-283.

Speight JG. 1977. Landform pattern description from aerial photographs. *Photogrammetria* 32: 161-182.

Stellenbosch University Institute for Wine Biotechnology. 2012. [Online]. Available: http://academic.sun.ac.za/wine_biotechnology. Date accessed: 15 March 2012.

Sumner ME. 2000. *Handbook of Soil Science*. Boca Raton: CRC Press.

Szabolcs I. 1994. Soils and salinisation. In*:* Pessarakli M. (ed.), *Handbook of plant and crop stress*. New York: Marcel Dekker Inc. pp 1-11.

Talsma T. 1981. Transport of salts in catchments and soils. *Agricultural Water Management* 4: 103-113.

Tamas J, Lenart C. 2006. Analysis of a small agricultural watershed using remote sensing techniques. *International Journal of Remote Sensing* 27: 3727-3738.

Tarboton DG. 1997. A new method for the determination of flow directions and contributing areas in grid digital elevation models. *Water Resources Research* 33: 309-319.

Tarekegn TH, Haile AT, Reintjies T, Reggiani P, Alkema D. 2010. Assessment of an ASTER-generated DEM for 2D hydrodynamic flood modeling. *International Journal of Applied Earth Observation and Geoinformation* 12: 457-465.

Taud H, Parrot JF, Alvarez R. 1999. DEM generation by contour line dilation. *Computers and Geosciences* 25: 775-783.

Thakur A, Anand RS. 2005. The local statistics based region growing segmentation method for ultrasound medical images. *International Journal of Signal Processing* 1: 141-146.

Thompson JL, Bell JC, Butler CA. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modelling. *Geoderma* 100: 67-89.

Trimble. 2011. *eCcognition Developer 8.64.1 reference book*. Munich: Trimble Germany GmbH.

Turcotte R, Fortin JP, Rousseau AN, Massicotte S, Villeneuwe JP. 2001. Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network. *Journal of Hydrology* 240: 225-242.

Tweed SO, Leblanck M, Webb JA, Lubczynski MW. 2007. Remote sensing and GIS for mapping groundwater recharge and discharge areas in salinity prone catchments, southeastern Australia. *Hydrogeology Journal* 15: 75-96.

Utset A, Borroto M. 2001. A modeling-GIS approach for assessing irrigation effects on soil salinisation under global warming conditions. *Agricultural Water Management* 50: 53-63.

Vagen T-G, Shepherd KD, Walsh MG. 2006. Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using VIS-NIR spectroscopy. *Geoderma* 133: 281-294.

van Asselen S, Seijmonsberen AC. 2006. Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM. *Geomorphology* 78: 309-320.

van Niekerk A. 2001. Western Cape Digital Elevation Model product description. Stellenbosch: Centre for Geographical Analysis, Stellenbosch University.

van Niekerk A. 2008. CLUES: A web-based land expert system for the Western Cape. PhD thesis, Stellenbosch University, South Africa.

van Niekerk A. 2010. A comparison of land unit delineation techniques for land evaluation in the Western Cape, South Africa. *Land Use Policy* 27: 937-945.

van Niekerk A. 2011. Stellenbosch University Digital Elevation Model product description. 2011 Edition. Stellenbosch: Centre for Geographical Analysis, Stellenbosch University.

Vaze J, Teng J, Spencer G. 2010. Impact of DEM accuracy and resolution on topographic indices. *Environmental Modelling and Software* 25: 1086-1098.

Verma KS, Saxena RK, Barthwal AK, Deshmukh SN. 1994. Remote sensing technique for mapping salt affected soils. *International Journal of Remote Sensing* 15: 1901-1914.

Viscarra Rossel RA. 2008. ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometric Inteligent Laboratory Systems* 90: 72-83.

Viscarra Rossel RA, Behrens T. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* 158: 46-54.

Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 131: 59-75.

Viscarra Rossel RA. 2007. Robust modelling of soil diffuse reflectance spectra by "bagging-partial least squares regression". *Journal of Near Infrared Spectroscopy* 15: 39-47.

Vogt JV, Colombo R, Bertolo F. 2003. Deriving drainage networks and catchment boundaries: a new methodology combining digital elevation data and environmental characteristics. *Geomorphology* 53: 281-298.

Walker JP, Willgoose GR. 1999. On the effect of digital elevation model accuracy on hydrology and geomorphology. *Water Resources Research* 35: 2259–2268.

Wang Q, Li P, Chen X. 2012. Modeling salinity effects on soil reflectance under various moisture conditions and its inverse application: A laboratory experiment. *Geoderma* 170: 103-111.

Wang W, Yang X, Yao T. 2011. In press. Evaluation of ASTER GDEM and SRTM and their suitability in hydraulic modelling of a glacial lake outburst flood in southeast Tibet. *Hydrological Processes*.

Wang X, Yin Z-Y. 1998. A comparison of drainage networks derived from digital elevation models at two scales. *Journal of Hydrology* 210: 221-241.

Weepener HL, Van den Berg HM, Metz M, Hamandawana H. 2012. The development of a hydrologically improved digital elevation model and derived products for South Africa based on the SRTM DEM. Report number 1908/1/11. Pretoria: Water Research Commission

Weng YL, Gong P, Zhu ZL. 2010. A spectral index for estimating soil salinity in the Yellow River Delta Region of China using EO-1 Hyperion data. *Pedosphere* 20: 378-388.

Williamson DR, 1998.Waterlogging and salinisation. In: Williams J, Hook RA, Gascoigne HL. (eds.), *Farming Action Catchment Reaction*. Collingwood: CSIRO Publishing. pp 162-187.

Wise SM. 2007. Effect of differing DEM creation methods on the results from a hydrological model. *Computers and Geosciences* 33: 1351-1365.

Wuest W, Zhang Y. 2009. Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison. *ISPRS Journal of Photogrammetry and Remote Sensing 64:* 55-64.

Wulder MA, White JC, Hay GJ, Castilla G. 2008. Pixels to objects to information: Spatial context to aid in forest characterization with remote sensing. In: Blaschke T, Lang S, Hay GJ (eds.), *Object-Based Image Analysis: Spatial concepts for knowledge-driven remote sensing applications*. Berlin: Springer. pp 345-363.

Xie K, Wu Y, Ma X, Liu Y, Liu B, Hessel R. 2003. Using contour lines to generate digital elevation models for steep slope areas: A case study of the Loess Plateau in North China. *Catena* 54: 161-171.

Xu P, Shao Y. 2002. A salt-transport model within a land-surface scheme for studies of salinisation in irrigated areas. *Environmental Modelling and Software* 17: 39-49.

Yao RJ, Yang JS. 2010. Quantitative evaluation of soil salinity and its spatial distribution using electromagnetic induction method. *Agricultural Water Management* 97: 1961-

Yu R, Liu T, Xu Y, Zhu C, Zhang Q, Qu Z, Liu X, Li C. 2010. Analysis of salinization dynamics by remote sensing in Hetao Irrigation District of North China. *Agricultural Water Management* 97: 1952-1960.

Zeilhofer P, Schwenk LM, Onga N. 2011. A GIS-approach for determining permanent riparian protection areas in Mato Grosso, central Brazil. *Applied Geography* 31: 990-997.

Zhang T-T, Zeng S-L, Gao Y, Ouyang Z-T, Li B, Fang C-M, Zhao B. 2011. Using hyperspectral vegetation indices as a proxy to monitor soil salinity. *Ecological Indicators* 11: 1552-1562.

Zhang W, Montgomery. 1994. Digital elevation model grid size, landscape representation and hydrologic simulations. *Water Resources Research* 30: 1019-1028.

APPENDIX A

SOIL SAMPLE DATA

Table A1 Sandspruit catchment soil samples

Samples	EC (mS m^{-1})	pH	Ca(mg/l)	Mg (mg/l)	Na (mg/l)	Cl (mg/l)	SO$_4$ (mg/l)
Sand_09	20.00	6.66	20.80	12.14	20.20	18.24	4.17
Sand_101	8.00	5.92	4.60	2.01	11.07	7.21	7.18
Sand_103	33.00	6.46	29.74	8.32	117.58	81.23	34.15
Sand_104	1353.00	7.80	728.06	442.23	3984.73	9386.96	2436.60
Sand_108	4.00	6.50	3.13	1.71	6.46	4.31	1.42
Sand_111	11.00	6.76	14.89	7.53	10.78	8.70	4.74
Sand_114	728.00	6.85	344.00	223.58	1290.73	3728.91	356.76
Sand_115	1520.00	8.06	313.81	349.82	4263.26	10255.90	843.28
Sand_116	1130.00	7.25	129.21	236.72	3233.63	7484.11	359.93
Sand_15	16.00	6.75	27.83	11.18	3.95	9.15	
Sand_20	36.00	6.65	54.19	7.40	8.68	28.37	8.10
Sand_21	20.00	6.79	36.52	4.25	5.51	11.86	5.16
Sand_23	83.00	6.75	68.09	16.67	10.08	45.34	11.32

Sand_28	82.00	6.06	107.84	8.98	28.47	45.39	287.54
Sand_32	143.00	6.60	89.66	23.37	150.90	265.00	242.07
Sand_34	230.00	6.69	103.96	43.99	257.70	707.00	119.56
Sand_39	77.00	6.88	41.28	14.74	73.06	93.35	48.20
Sand_61	92.00	6.74	52.51	17.96	12.70	34.33	16.88
Sand_71	154.00	7.28	62.50	27.81	104.13	211.89	63.58
Sand_72	79.00	7.01	52.02	8.76	23.27	30.23	12.43
Sand_89	120.00	6.37	59.98	20.42	109.98	374.43	37.16
Sand_90	912.00	6.58	164.11	293.99	1486.62	3644.34	754.25
Sand_98	7.00	5.70	4.00	1.21	3.46	4.22	5.36

Table A2 Agricultural Research Council ad hoc samples

Samples	EC (mS m^{-1})	pH	Ca (mg/l)	Mg (mg/l)	Na (mg/l)	Cl (mg/l)	SO$_4$ (mg/l)
JHB_2229	121.00	7.44	60.10	20.89	166.59	57.92	457.16
JHB_2230	145.00	7.29	122.65	38.93	130.14	221.11	485.30
JHB_2231	88.00	7.09	51.22	16.95	90.55	62.75	225.66
JHB_2232	121.00	6.94	105.28	37.40	86.23	115.25	408.37
JHB_2233	221.00	6.87	201.75	64.18	189.07	172.42	858.19
JHB_2234	186.00	5.70	149.91	67.64	146.25	338.13	590.31
JHB_2235	215.00	7.56	142.53	49.21	244.51	195.19	851.29
JHB_2236	216.00	7.33	153.41	49.59	213.93	377.85	659.06
JHB_2237	162.00	7.65	83.78	32.00	193.79	109.38	644.67
JHB_2238	257.00	6.95	207.18	69.41	229.94	321.05	991.53
JHB_2239	429.00	7.29	528.59	132.53	339.43	238.68	2598.31
JHB_2240	220.00	7.07	184.78	99.46	78.30	166.66	779.21
JHB_2241	265.00	7.65	214.13	64.45	261.10	225.56	1047.74
JHB_2242	229.00	7.52	256.10	59.44	190.55	285.62	821.82
JHB_2243	161.00	7.78	118.48	37.37	156.04	130.24	603.69
JHB_2244	119.00	7.74	104.97	27.78	100.39	80.09	248.27
JHB_2245	60.00	7.81	36.58	13.16	56.95	63.93	198.26
JHB_2246	50.00	7.79	26.73	9.02	49.66	32.98	138.38
JHB_2247	39.00	7.98	33.33	11.31	19.96	19.59	30.16
JHB_2248	117.00	7.93	88.94	28.22	96.03	57.86	411.79
JHB_2249	37.00	8.01	25.44	8.90	26.30	14.71	60.83
JHB_2250	43.00	7.82	19.79	7.33	45.39	24.87	71.03
JHB_2251	78.00	7.83	34.83	12.36	87.52	28.23	263.49

JHB_2252	198.00	7.39	172.02	59.76	154.17	267.20	760.46
JHB_2253	1234.00	7.77	482.69	534.96	1776.85	2776.51	4251.89
JHB_2254	418.00	7.36	167.18	147.68	508.70	811.63	1464.83

Table A3 Land Type Database soil samples

LabNoType	LabNo	SoilClass	X-Coord	Y-Coord	Texture	Org_C	Extr_Exch	Conductivity
C	1759	Ar20	31.96500015	-27.5083	ClLm	2.799999952	Exchangeable	90
C	1997	We12	27.02444458	-29.2325	Cl	0.100000001	Exchangeable	120
C	2799	Va21	26.32999992	-28.5633	Cl	0.400000006	Exchangeable	60
C	2891	Oa46	26.11277771	-29.7806	Lm	0.200000003	Exchangeable	90
C	4556	Av36	26.26000023	-27.3767	SaClLm	0.150000006	Exchangeable	46
C	4560	We13	27.10666656	-27.3567	SaCl	0.289999992	Exchangeable	56
C	4647	Hu46	30.06388855	-24.8856	Lm	0.100000001	Exchangeable	64
C	5107	Bo30	31.52833366	-27.5917	SaLm	0.200000003	Exchangeable	80
C	7264	Oa46	20.2816658	-31.3825	SiLm	0.100000001	Exchangeable	124
D	244	Es15	29.27083397	-23.825	LmSa	0.430000007	Exchangeable	70
C	1903	Kd13	25.48111153	-33.9761	SaClLm	0.200000003	Exchangeable	308
C	2569	Es33	28.57777786	-26.2806	SaCl	0.200000003	Exchangeable	240
C	2698	Oa47	26.09777832	-30.1453	SiCl	0.899999976	Exchangeable	260
C	2708	Ar20	26.38194466	-29.0667	Cl	0.5	Exchangeable	203
C	4728	Hu43	17.22333336	-28.87	SaLm	0.300000012	Exchangeable	300
C	5464	Sw20	22.40166664	-32.37	SaClLm	0.300000012	Exchangeable	234
C	5563	Ss26	25.51388931	-32.3767	Cl	0.200000003	Exchangeable	295
C	5584	Gs23	24.28388977	-32.9625	SaCl	0.300000012	Exchangeable	351
C	6384	Hu34	17.88166618	-30.7697	LmSa	0.600000024	Exchangeable	370
C	6388	Cv21	17.79055595	-30.9731	Sa	0.360000014	Exchangeable	267
C	8643	Ss13	26.74749947	-30.6306	SiClLm	0.059999999	Exchangeable	340
C	8936	Hu44	22.82638931	-30.2228	SaLm	0.180000007	Exchangeable	348
D	306	Oa16	19.52750015	-31.3889	SaLm	0.379999995	Exchangeable	214
C	2298	Rg20	31.7788887	-25.8572	ClLm	0.300000012	Exchangeable	465
C	3448	Kd11	18.09222221	-32.8339	Sa	0.100000001	Exchangeable	460
C	3827	Es14	23.33166695	-33.9508	ClLm	0.100000001	Exchangeable	514
C	4228	Bo41	32.05833435	-28.3833	Cl	0.300000012	Exchangeable	445
C	4281	Ss16	32.24000168	-27.695	Cl	0.100000001	Exchangeable	624
C	4407	Kd16	19.52444458	-33.6828	SaClLm	0.100000001	Exchangeable	650
C	4713	Cv41	16.68333244	-28.8633	Sa	0.100000001	Exchangeable	412
C	4916	Hu32	17.45833397	-29.2208	SaLm	0.100000001	Exchangeable	440
C	5337	Va41	31.26666641	-24.35	SaCl	0.200000003	Exchangeable	690
C	5771	Oa47	20.52000046	-30.4767	SaClLm	0.200000003	Exchangeable	775
C	6511	Hu46	25.89888954	-33.2256	SaClLm	0.200000003	Exchangeable	480
C	6514	Oa46	25.97249985	-33.0444	SaClLm	0.400000006	Exchangeable	400
C	7250	Va20	20.09555626	-31.1361	Lm	0.189999998	Exchangeable	459
C	8312		19.2088089	-31.3861	ClLm	0.730000019	Exchangeable	758
C	1555	Ar20	32.13499832	-27.4983	Cl	0.200000003	Exchangeable	1490
C	1875	Hu46	26.41666603	-33.2	SaClLm		Exchangeable	1172
C	3449	Kd11	18.09222221	-32.8339	LmSa	0.200000003	Exchangeable	980
C	3779	Ms20	20.41388893	-34.4972	LmSa	0.800000012	Extractable	
C	4305	Bo40	32.10833359	-28.3433	SaClLm	0.300000012	Exchangeable	1450
C	4403	Ss26	19.47277832	-33.7839	SaCl	0.300000012	Exchangeable	1158
C	4719	Hu42	17.71333313	-28.9133	Sa	0.100000001	Exchangeable	1407
C	4948	Hu32	17.19000053	-29.1667	Sa	0.100000001	Exchangeable	1370
C	4951	Hu42	17.39722252	-29.3431	LmSa	0.200000003	Exchangeable	1500
C	5146	Hu34	22.37916756	-33.5639	LmSa	0.100000001	Exchangeable	1590
C	5147	Hu34	22.37916756	-33.5639	Sa	0.100000001	Exchangeable	1000
C	5377	Oa44	21.92277718	-33.1136	LmSa	0.100000001	Exchangeable	1200
C	5515	Oa46	22.07250023	-33.1972	SaLm	0.100000001	Exchangeable	1070
C	6244	Hu33	18.66333389	-31.8667	Sa	0.100000001	Exchangeable	1207
C	6419	Cv31	17.66777802	-30.9778	SaLm	0.200000003	Exchangeable	900
C	6420	Cv31	17.66777802	-30.9778	SaLm	0.100000001	Exchangeable	1449
C	8945	Sd30	23.69611168	-30.4869	Lm	0.159999996	Exchangeable	821
C	9485	Va20	29.11666679	-22.6833	SaLm	0.200000003	Exchangeable	978
C	1554	Ar20	32.13499832	-27.4983	Cl	0.800000012	Exchangeable	1860
C	2408	Oa36	19.36555481	-33.6242	SaClLm	0.200000003	Exchangeable	2680
C	4720	Hu42	17.71333313	-28.9133	LmSa	0.100000001	Exchangeable	2345
C	4864	Es30	31.01166725	-28.0217	SaLm	0.300000012	Exchangeable	1644
C	4964	Hu35	18.42333412	-30.3133	Sa	0.230000004	Exchangeable	320
C	5470	Oa47	21.10333252	-30.0014	Cl	0.300000012	Exchangeable	1776
C	5484	Oa47	19.56666756	-30.36	SiCl	0.300000012	Exchangeable	3980
C	5485	Oa47	19.56666756	-30.36	SiLm	0.200000003	Exchangeable	3440
C	5543	Va22	25.11249924	-32.4356	Cl	0.600000024	Exchangeable	1960
C	5978	Oa47	24.08444405	-30.1853	Lm	0.100000001	Exchangeable	2272
C	6270	Va21	24.38999939	-31.7144	Cl	0.200000003	Exchangeable	1634
C	6390	Ss23	17.88916588	-31.0172	SaCl	0.400000006	Exchangeable	1974
C	8987	Oa46	22.93333244	-30.8694	ClLm	0.170000002	Exchangeable	2000

APPENDIX B

ARTICLES SUBMITTED AND PREPARED

de Clercq WP, Javanovic N, Bugan R, Mashimbye E, Du Toit T, Van Niekerk A, Ellis F, Wasserfall N, Botha P, Steudels T, Helschrot J, Flügel WA. 2013. Management of human-induced salinization in the Berg River catchment and development of criteria for regulating agricultural land use in terms of salt generating capacity. Final Report to the Water Research Commission, Report No. 1849/01/2013.

Mashimbye ZE, Cho MA, Nell JP, de Clercq JP, van Niekerk A, Turner DP. 2012. Model-based integrated methods for quantitative estimation of soil salinity from hyperspectral remote sensing data: a case study of selected South African soils. *Pedosphere* 22: 640-649.

Mashimbye ZE, de Clercq WP, van Niekerk A. Under review. An evaluation of digital elevation models for delineating land components. Geoderma.

Mashimbye ZE, de Clercq WP, van Niekerk A. Under review. In Preparation. Assessing the influence of DEM source on derived streamline and catchment boundary accuracy. Water SA.

Mashimbye ZE, de Clercq WP, van Niekerk A. In Preparation. Near infrared spectroscopy and chemometric modeling of salt-affected soils.

Mashimbye ZE, de Clercq WP, van Niekerk A. In preparation. The value of terrain attributes to map soil salinity

CONFERENCE PROCEEDINGS

Mashimbye ZE, Cho MA, Nell JP, de Clercq JP, van Niekerk A, Turner D.P. 2011. Hyperspectral remote sensing of salt-affected soils. Combined Congress, 17 – 21 January 2012, Pretoria, South Africa.

Mashimbye ZE, Cho MA, Nell JP, de Clercq JP, van Niekerk A, Turner DP. 2012. Model-based hyperspectral quantitative analysis of soil salinity: a South African case study. 2nd South African Chemometrix Society Conference, 07 – 09 May 2012, Irene, South Africa.

Mashimbye ZE, de Clercq WP, van Niekerk A, Nieuwoudt HH. 2012. Near infrared spectroscopy and chemometric modeling of salt-affected soils. 2nd South African Chemometrix Society Conference, 07 – 09 May 2012, Irene, South Africa.

Mashimbye ZE, de Clercq WP, van Niekerk A. 2011. Deriving hydrological parameters from DEMs: A case study of the Sanspruit catchment in South Africa. Geography Doctoral and Post-Doctoral Conference, 11th of November 2011, Stellenbosch, South Africa.

Mashimbye ZE, de Clercq WP, van Niekerk A. 2012. Digital terrain mapping of the Sandpsruit catchment of the Western Cape Province in South Africa. Combined Congress, 16 – 19 January 2012, Potchefstroom, South Africa.