EVALUATING SOIL AND TERRAIN VARIABLES IN A PRODUCTION ENVIRONMENT: IMPLICATIONS FOR AGRICULTURAL LAND ASSESSMENT

by

Kurt Russell Barichievy

A dissertation presented for the degree of
Doctor of Philosophy
Agricultural Sciences

at

Stellenbosch University
Department of Soil Science, Faculty of AgriSciences

Supervisor: Dr Catherine E. Clarke
Co-Supervisor: Dr Andrei B. Rozanov

December 2022

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 part submitted it for obtaining any qualification.

December 2022

Kurt Barichievy

SUMMARY

Agricultural land in South Africa is under increasing pressure to produce more food from an ever-shrinking land base, as more land is being converted to non-productive uses. Additional to these pressures, is the concept of land reform and strategic land acquisition, aimed at agrarian transform within the rural landscape. It is estimated that less than 15% of South Africa is suitable for dryland cultivation. Consequently, the sustainable utilisation of these scarce resources and preservation of agricultural land is of paramount importance, to ultimately ensure some measure of national food security in the years to come. Agricultural land evaluation is a critical tool that can achieve this goal. Unfortunately, in recent decades the development of revised or novel land evaluation methodologies has stalled for South African farm-level assessments, the scale at which land release decisions are made. Further, the relationship between productivity and individual land assessment attributes has not been adequately quantified or incorporated into contemporary local assessment procedures.

It is envisaged that this study would influence and help guide in-field methodologies, as well as draft legislation and best-practice strategies, with a view of both standardising and improving agricultural land assessment techniques. By emphasising the importance of agricultural land and the accurate assessment thereof, this research also aims to increase our understanding of production-based approaches at an operational scale, though the novel combination of traditional approaches and use of newer technologies. It is anticipated that this improved understanding will be employed to not only protect more agricultural land, which may have been undervalued by historical methods, but also as an intuitive assessment tool to highlight the yield gap between potential and actual production levels.

A review of pertinent literature identified the need for local verification studies to evaluate the performance of land assessment methodologies currently used in industry. To address this, five methods were verified using land assessment polygons in a commercial production environment, in the Province of KwaZulu-Natal, South Africa. The resultant classifications, derived from 225 soil observations, were compared to actual land use and precision yields achieved by dryland maize and soybean, across five growing seasons (2016 - 2020). By comparing land use with broad arability, four of the five land assessment methods were found to adequately classify arable land. Additionally, land evaluation polygons, linked to dryland precision maize and soybean yields can provide a general overview of method performance. However, it was concluded that yield performance and variation, across land evaluation

methods and classes, is only explicit on or near a soil observation point where measurements are taken. Accordingly, seasonal variograms for maize and soybean were developed, to establish a representative yield buffer around individual soil observation points. This, along with yield normalisation strategies were employed, to improve verification procedures across multiple growing seasons.

To determine crop productivity drivers, significant land assessment attributes *inter alia* slope, effective rooting depth, soil texture, soil group and soil wetness limitations were analysed against maize and soybean yields. It was found that the two crops respond differently to individual land assessment attributes and these differences should be taken cognisance of in new, crop-specific land evaluation methodologies and weighted accordingly.

In an attempt to improve productivity-based land classification 78 attributes; derived from land assessment methodologies, digital terrain analysis, the pedological survey and soil colour spectrophotometry were collated. From these attributes, three new approaches, one based on biophysical scoring criteria and two based on machine learning, were developed across two commercial farming operations, in northern KwaZulu-Natal. These new methodologies were then tested on three separate commercial operations, located in different regions of the province.

The biophysical scoring classification generally outperformed machine learning models and was particularly accurate when classifying observations associated with either extremely poor or extremely advantageous soil and terrain attributes. The transferability of the models to other regions, with different resources produced mixed results, highlighting the need for wider calibration in some instances. The study also found that the new productivity-based approaches can have useful applications in commercial farm management, where crop specific classification can identify underperforming areas and yields gaps, which can be ringfenced for appropriate interventions.

The newly developed biophysical scoring classification was used to demonstrate the utility of these approaches in broader agricultural land release applications. The study found the new approaches better reflect production potential and should be used to supplement existing methodologies in land release assessments. Ultimately, the application of these production-based approaches can assist the land assessor to better classify the production potential of the land, as well as the decision-making authority to justify preserving more land for agricultural purposes.

OPSOMMING

Landbougrond in Suid-Afrika is onder toenemende druk om meer voedsel van 'n steeds krimpende grondbasis te produseer, aangesien meer grond na nie-produktiewe gebruike omgeskakel word. Bykomend tot hierdie druk is die konsep van grondhervorming en strategiese grondverkryging, gemik op agrariese transformasie binne die landelike landskap.

Daar word beraam dat minder as 15% van Suid-Afrika geskik is vir droëlandverbouing. Gevolglik is die volhoubare benutting van hierdie skaars hulpbronne en bewaring van landbougrond van kardinale belang, om uiteindelik 'n mate van nasionale voedselsekerheid in die komende jare te verseker. Landbougrond evaluering is 'n kritieke instrument wat hierdie doelwit kan bereik. Ongelukkig het die ontwikkeling van hersiene of nuwe grondevaluering metodologieë in die afgelope dekades vir Suid-Afrikaanse plaasvlak-assesserings, die skaal waarop besluite oor grondvrystelling geneem word, tot stilstand gekom. Verder is die verwantskap tussen produktiwiteit en individuele grondbeoordeling eienskappe nie voldoende gekwantifiseer nie, en ook nie ingesluit in kontemporêre plaaslike assessering prosedures nie.

Daar word in die vooruitsig gestel dat hierdie studie in-veld metodologieë, sowel as konsep-wetgewing en beste-praktyk strategieë sal beïnvloed en help rig, met die oog op beide standaardisering en verbetering van landbougrond assessering tegnieke. Deur die belangrikheid van landbougrond en die akkurate beoordeling daarvan te beklemtoon, poog hierdie navorsing ook om ons begrip van produksie gebaseerde benaderings op 'n operasionele skaal te verhoog, al is die nuwe kombinasie van tradisionele benaderings en die gebruik van nuwer tegnologieë word is missing. Daar word verwag dat hierdie verbeterde begrip aangewend sal word om nie net meer landbougrond, wat moontlik deur historiese metodes onderwaardeer is, te beskerm nie, maar ook as 'n intuïtiewe assessering instrument om die opbrengsgaping tussen potensiële en werklike produksievlakke uit te lig.

'n Oorsig van toepaslike literatuur het die behoefte aan plaaslike verifikasie studies geïdentifiseer om die prestasie van grondbeoordeling metodologieë wat tans in die industrie gebruik word (removed comma) te evalueer. Om dit aan te spreek, is vyf metodes geverifieer deur gebruik te maak van grondevaluering poligone in 'n kommersiële produksie omgewing, in die provinsie KwaZulu-Natal, Suid-Afrika. Die gevolglike klassifikasies, afgelei van 225 grond waarnemings, is vergelyk met werklike grond-gebruik en presisie-opbrengste wat deur droëland-mielies en sojabone behaal is, oor vyf groeiseisoene (2016 - 2020). Deur

grondgebruik met breë bewerkbaarheid te vergelyk, is vier van die vyf grondbeoordeling metodes gevind om bewerkbare grond voldoende te klassifiseer. Boonop kan grondevaluering poligone, gekoppel aan droëland-presisiemielies en sojaboon opbrengste, 'n algemene oorsig van metode prestasie verskaf. Daar is egter tot die gevolgtrekking gekom dat opbrengsprestasie en variasie, oor grondevaluering metodes en -klasse heen, slegs eksplisiet is op of naby 'n grondwaarnemingspunt waar metings geneem word. Gevolglik is seisoenale variogramme vir mielies en sojabone ontwikkel om 'n verteenwoordigende opbrengsbuffer rondom individuele grondwaarnemingspunte te vestig. Dit, tesame met opbrengs normalisering strategieë, is aangewend om verifikasie prosedures oor verskeie groei seisoene te verbeter.

Om oesproduktiwiteit drywers te bepaal, is beduidende grondbeoordeling eienskappe, onder andere helling, effektiewe worteldiepte, grondtekstuur, grondgroep- en grondnat beperkings, ontleed teen mielie- en sojaboon opbrengste. Daar is gevind dat die twee gewasse verskillend reageer op individuele grondbeoordeling eienskappe en hierdie verskille moet in nuwe, gewas-spesifieke grondevaluering metodologieë in ag geneem word en dienooreenkomstig geweeg word.

In 'n poging om produktiwiteit-gebaseerde grondklassifikasie te verbeter 78 eienskappe; afgelei van grondevaluering metodologieë, digitale terrein analise, die pedologiese opname en grondkleur spektrofotometrie is saamgestel. Uit hierdie eienskappe is drie nuwe benaderings, een gebaseer op biofisiese telling kriteria en twee gebaseer op masjienleer, ontwikkel oor twee kommersiële boerdery bedrywighede, in die noorde van KwaZulu-Natal. Hierdie nuwe metodologieë is toe getoets op drie afsonderlike kommersiële bedrywighede, geleë in verskillende streke van die provinsie.

Die biofisiese punte-klassifikasie het oor die algemeen beter as masjienleer-modelle presteer en was besonder akkuraat wanneer waarnemings geassosieer met óf uiters swak óf uiters voordelige grond- en terrein kenmerke geklassifiseer is. Die oordraagbaarheid van die modelle na ander streke, met verskillende hulpbronne, het gemengde resultate opgelewer, wat die behoefte aan wyer kalibrasie in sommige gevalle beklemtoon. Die studie het ook bevind dat die nuwe produktiwiteit gebaseerde benaderings nuttige toepassings in kommersiële plaasbestuur kan hê, waar gewas-spesifieke klassifikasie onderpresterende gebiede en opbrengsgapings kan identifiseer, wat afgesper kan word vir toepaslike ingrypings.

Die nuut ontwikkelde biofisiese punte-klassifikasie is gebruik om die nut van hierdie benaderings in breër landbougrond-vrystelling toepassings te demonstreer. Die studie het bevind die nuwe benaderings weerspieël produksie potensiaal beter en moet gebruik word om bestaande metodologieë in grondvrystelling evaluerings aan te vul. Uiteindelik kan die toepassing van hierdie produksie gebaseerde benaderings die grond beoordelaar help om die produksie potensiaal van die grond beter te klassifiseer, asook die besluitneming gesag om die behoud van meer grond vir landbou doeleindes te regverdig.

I dedicate this thesis to my wife, Michelle and my children, Joshua and Rachel.
You are my centre when I spin away.

BIOGRAPHICAL SKETCH

I, Kurt Barichievy, graduated with a BSc degree in hydrology and environmental science from the University of KwaZulu-Natal in December 2004 with a Deans Commendation and the Roland Schulze Award for the top, final year hydrology student. In December 2005 I completed my BSc (Hons) *cum laude* in Hydrology, which focussed on using wetland water budgeting to quantify typical wetland functioning. In August 2009, under the supervision of Professor Roland Schulze, I completed my MSc in Hydrology, which investigated the higher order impacts of climate change on eco-hydrological indicators and water temperature in South Africa.

Between 2008 and 2014 I was employed at SiVEST Consulting Engineers in Pietermaritzburg, where I focussed on specialist pedological and hydrological assessments. Since 2014 I have worked as a Professional Agricultural Scientist in the Natural Resource Sub-Directorate, at the KwaZulu-Natal Department of Agriculture and Rural Development. My current role focusses on pedological surveys and agricultural land assessment, with the primary aim of increasing agricultural production while preserving the natural resource base. I am a registered Earth Science Professional with the South African Council of Natural and Professional Scientists (SACNASP). I am also a member of the Soil Science Society of South Africa (SSSSA), the South African Soil Surveyors Organisation (SASSO) and hold nominated positions on the South African Soil Classification Working Group and National Sub-working Group on Natural Resources Inventories and Assessments (NRIA).

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- To Dr Cathy Clarke for her tireless encouragement and guidance. Thank you for the lengthy brainstorming sessions and for always supporting my research ideas. For this I am very grateful. I look forward to working together on new projects and challenges!
- To Dr Andrei Rozanov for his co-supervision. Your experience and guidance were most welcomed.
- To my colleagues at the Natural Resources Sub-Directorate, particularly Cobus Botha, Felicity Mitchell, Janet Taylor and Michelle Keith. Thank you for the support and sharing the countless hours classifying, sampling and discussing the 500 or so soil profiles used in this project.
- To Dr Jonathan Atkinson for the encouragement and input during our long telephone conversations between Cedara and Bergville. Never stop having too many irons in the fire!
- To Mr Stephan van der Westhuizen for his general advice but specifically for the development of the machine learning models and assistance with the yield buffer analysis.
- To the KwaZulu-Natal Department of Agriculture and Rural Development for the support of this research and partial bursary funding.
- To the farmers, Lester Vickers, Egon and Tyson Zunckel, Bertjie Pretorius and Koot Smit, thank you for allowing me access to your farms and yield data. I hope to do your soils justice.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABREVIATIONS AND ACRONYMS

ALOS	Advanced Land Observing Satellite
ANOVA	Analysis of Variance
ARC	Agricultural Research Council
AUC	Area under curve
BSC	Biophysical scoring classification
BRG	Bioresource group
BRU	Bioresource unit
CA	Classification accuracy
CARA	Conservation of Agricultural Resources Act of 1983
ERD	Effective rooting depth
ESRI	Environmental Systems Research Institute
DAFF	National Department of Agriculture, Forestry and Fisheries
DAFF LC	National Department of Agriculture, Forestry and Fisheries digital Land Capability
DALRRD	Department of Agriculture, Land Reform and Rural Development
DEM	Digital elevation model
DSM	Digital soil Mapping
FAO	Food and Agricultural Organisation of the United Nations
FCC	Fertility Capability Classification
GIS	Geographical information systems
GPS	Global position systems
KZN	KwaZulu-Natal
KZNDARD	KwaZulu-Natal Department of Agriculture and Rural Development
KZN LC	KwaZulu-Natal Land Capability System
LC	Land Capability
LESA	Land evaluation and site assessment
MCC	Matthews Correlation Coefficient
ML	Machine learning
MR	Misclassification rate
NEMA	National Environmental Management Act of 2008
PA	Precision agriculture
PDALB	Preservation and Development of Agricultural Land Bill
RF	Random Forest classification

ROC	Receiver operating characteristic curve
RSA LC	South African Land Capability System
RSP	Relative slope position
SAGA	System for geoscientific analysis
SAGL	South African grain laboratories
SALA	Subdivision of Agricultural Land Act of 1970
SCGW	Soil classification working group
SD	Standard deviation
SNV	Standardised normal values
SRTM	Shuttle radar topography mission
SPR	Soil potential ratings
SPLUMA	Spatial Planning and Land Use Management Act of 2013
SVM	Support vector machines
TPI	Terrain position index
USDA	United States Department of Agriculture
VSA	Visual soil assessment

1. INTRODUCTON

One of the greatest threats facing food security in South Africa is the loss of agricultural land to non-productive land uses (Newby et al., 2018). A spatial analysis undertaken by the National Department of Agriculture, Forestry and Fisheries (DAFF) in 2011 estimates that 3 million hectares of agricultural land has already been lost to urban and mining developments, with more agricultural land being lost each year (DAFF, 2016a). This problem is further exacerbated by a growing population and a greater demand for foodstuffs, with an estimated 7 million more mouths to feed in South Africa by 2030 (United Nations, 2019). Consequently, commercial farmers will need to sustainably produce more food from an ever-shrinking land base.

Coupled with the threat of unsustainable land use change, is the concept of land reform and strategic land acquisition, aimed at agrarian transform within the rural landscape. While the difficulties of land reform are well documented (Cloete, 1992; Kotze and Basson, 1993; Beehner, 2006), there is renewed interest in policy transformation and ultimately constitutional amendments, in the hopes of rectifying historical imbalances, while not impacting upon the Nation’s ability to meet its food security requirements.

Land as a resource is frequently undervalued by a wide range of scientists as well as by spatial planners and policy makers (Davidson, 2002). It is estimated that less than 14% of South Africa’s land surface is suitable for dryland cropping and of this area, only 3% is land of high potential (Smith, 2006). The sustainable utilisation of these scarce resources and preservation of agricultural land is thus of paramount importance in South Africa to ensure some measure of national food security in the years to come. Agricultural land evaluation is a critical tool that can achieve this goal by classifying and ultimately recommending an appropriate farming system (Camp et al., 1995).

Land requires evaluation, as it is not equal and varies in physical, social, economic and geographic properties (Rossiter, 1996). Further, land can be evaluated and classified in innumerable ways, based on different factors and for different purposes and objectives (Tesfagiorgis, 2004). When placed within an agricultural context, land evaluation involves undertaking and interpreting natural resource surveys of climate, soil, terrain and vegetation, in terms of requirements and realistic land use options (FAO, 1976). Further, van Niekerk (2010) considers land evaluation an integral part of land use planning to ultimately support sustainable land use management.

In terms of land evaluation methodologies and tools, there are a myriad available to planners, technical specialists and scientific advisors who work in agricultural and natural environments. As the field of land evaluation expands and becomes more interdisciplinary, the development and availability of new tools equally increases and diversifies (van Diepen et al., 1991) Consequently, land evaluation approaches come in many forms, ranging from frameworks (e.g. FAO, 1976) and manuals (Scotney et al., 1991) to computer-based models (Rossiter & Wambeke, 1997) or simply GIS layers and spatial databases (e.g. DAFF, 2018) from which data can be extracted. The Food and Agricultural Organisation (FAO) alone, provides links to over 30 individual tools, which fall under its *Land Resource Planning Toolbox* umbrella (*www.fao.org).* In addition, there are also many country specific and regionalised methods, which aim to provide more relevant results through the development and utilisation of locally adapted models and input datasets.

Over recent decades, rapid anthropogenically driven land use change and urbanisation have triggered a need to critically examine land assessment methodologies (Beek, 1978). However, many local and international agricultural land evaluation methods still being utilised today, were developed decades ago without receiving further revisions or updates. For example, the South African Land Capability System was last revised in 1991, to incorporate changes introduced by the now outdated Taxonomic Soil Classification System (SCWG, 1991). The need for regular revision of land assessment methodologies is widely acknowledged (e.g. FAO, 2007 and Laker, 2004), with even the developers of the South African system stating that "*further research and understanding of the environment will lead to more dynamic and automatic land evaluation systems that will replace the simplistic approach presented*" (Scotney et al., 1991). Unfortunately, no revised nor novel land evaluation methodologies have recently been released for South African farm-level assessments. Ultimately, regular scientific validation, review and advancement is critical, to ensure the methods being utilised in practice are accurate and serve their intended purpose, which to identify and preserve important agricultural resources.

1.1 Rationale

There is a need for accurate and reliable agricultural land assessment methodologies, particularly in South Africa, where competition for land is high (Akinyemi and Mushunje, 2019; Simpson et al., 2019). To substantiate release, a change of land use or subdivision of agricultural land, the National Department of Agriculture, now part of DALRRD: Department of Agriculture, Land Reform and Rural Development, requires that an *Agro-Ecosystems* report be compiled, which includes a polygon based, farm-level survey and agricultural assessment (DAFF, 2018b).

Current methods of evaluation, such as land capability order, class, subclass, or unit rely on predefined soil-physical and terrain related properties to delineate the landscape into relatively homogenous polygons of similar limitation or production potential (Scotney et al., 1991). Although there is a wide range of both international and locally calibrated evaluation methods, available to South African land assessment practitioners, their accuracy and validity has not been sufficiently tested in a local production environment. Further, the relationship between productivity and individual land assessment attributes, has not been quantified nor incorporated into contemporary assessment procedures.

1.2 Aims and Objectives

It is envisaged that this study would influence and help guide in-field methodologies, as well as draft legislation and best-practice strategies, with a view of both standardising and improving agricultural land assessment techniques. By emphasising the importance of agricultural land and the accurate assessment thereof, this research aims increase to our understanding of production-based approaches at an operational scale, though the novel combination of traditional approaches and use of newer technologies. It is anticipated that this improved understanding will be employed to not only protect more agricultural land, which may have been undervalued by historical methods, but also as an intuitive assessment tool for farmers to highlight the yield gap between actual and potential production levels. To realise these primary aims the following research objectives were developed:

1. Explore pertinent literature and legislation surrounding agricultural land assessment and where applicable, highlight challenges and the need for review.
2. Assess if soil and land assessment approaches, currently being practised in industry, reflect actual land utilisation and production levels.
3. Investigate and quantify the relationship between individual land assessment attributes and productivity.
4. Develop novel, locally calibrated procedures for use in a specific commercial production environment.
5. Test the utility and robustness of these approaches in different locales and for different applications.

1.3 Thesis Outline

This thesis consists of seven chapters. This first introductory chapter is followed by a literature review, which presents an overview of the key concepts surrounding agricultural land assessment and evaluation. The importance of proper land use planning, as well as a summary of pertinent land policies in South Africa provides an initial framework for the study. A review of prominent international evaluation frameworks and assessment methods is included, along with a review of regionalised methods designed for South Africa. Finally, the need for revised and more dynamic systems is presented.

Chapter 3 deals with the verification of land assessment polygons in a commercial production environment. The resultant classifications, derived from five different land assessment methodologies, were verified using actual land use and precision yields achieved by dryland maize and soybean across five growing seasons (2016 - 2020).

Chapter 4 builds on the previous verification exercise by focussing on land assessment classification at a point scale. Seasonal variograms for maize and soybean were developed, in order to establish a representative yield buffer around individual soil observation points. This, along with yield normalisation strategies were employed, to improve verification techniques. Key land assessment attributes were identified and compared to maize and soybean performance, with the view of developing new production-based methodologies.

Chapter 5 begins with the incorporation of new land assessment attributes; derived from digital terrain analysis, the pedological survey and soil colour spectrophotometry. Three new approaches, two based on machine learning and one using biophysical scoring criteria, were developed using soil and terrain attributes and their relationship to crop productivity. These new methodologies were then tested on three separate commercial operations, located in different regions of the KwaZulu-Natal Province. The utility of these approaches in broader agricultural land release applications and commercial farm management completes the chapter.

Chapter 6 provides an integrated discussion, where the major findings of this research are contextualised within broader international literature and research.

The thesis concludes with Chapter 7, which incorporates the general conclusions and recommendations from the study, as well as the possibilities for future work.

2. LITERATURE REVIEW

2.1 Introduction

To effectively explore the concepts surrounding agricultural land evaluation and assessment a subset of literature has been selected, based on its relevance to the following questions:

1. What is land and land use?
2. What is agricultural land evaluation and why is it important?
3. What subject-specific terminology is commonly used within the field of land evaluation?
4. What are the Governmental legislative policies, which guide agricultural land management, land assessment and planning in South Africa?
5. What are the primary methods used to assess agricultural land internationally?
6. What are the primary methods used to assess agricultural land in South Africa?
7. Do South African land assessment methods need to be reviewed and updated?

2.2 What is Land and Land Use?

Land is a finite and precious resource, whose use and function is multi-faceted; providing food, shelter and a basis for development (Verheye, 2009). Due to these complexities, land should not be viewed simply as "soil" and should rather include the overarching physical environment (FAO, 1976).

More recently the Food and Agriculture Organisation (FAO) accepts the more holistic definition of "land", as provided by The United Nations (1994), "as a delineable area of the earth's terrestrial surface, encompassing all attributes of the biosphere immediately above or below this surface including those of the near-surface climate, the soil and terrain forms, the surface hydrology (including shallow lakes, rivers, marshes and swamps), the near-surface sedimentary layers and associated groundwater reserve, the plant and animal populations, the human settlement pattern and physical results of past and present human activity (terracing, water storage or drainage structures, infrastructure, buildings, etc.)."

This broad definition places more emphasis on the environmental aspects of land and importantly takes into account anthropogenic influences (FAO, 2007). These anthropogenic influences manifest themselves as a particular *land use* which is the application of human control on natural systems, in order derive a benefit from it (Sys, 1985). Examples of major, single land uses include rainfed and irrigated agriculture, grassland, forestry and recreation (FAO, 1976).

The term “land use” is synonymous with anthropogenic activities, which are directly connected to land, utilising its natural resources, or having an impact on it (DAFF, 2016). The increasing demand for food (Schiefer et al., 2016) has led to the agricultural sector being a major driver of land use change, with an estimated 30% increase in arable land expansion in Sub-Saharan Africa between 1961 and 2005 (Nkonya et al., 2012). Ultimately a change in land use will often require a land assessment evaluation, in order substantiate such change (FAO, 2007).

2.3 What is Agricultural Land Evaluation and why is it Important?

Stability in agricultural production and the conservation of limited resources, particularly in South Africa, can be achieved through the appreciation of the natural factors governing production and the implementation of sustainable land use systems (Smith, 2006). Agricultural land evaluation is a critical tool, which can assess these natural factors and recommend appropriate land uses. Specifically, Verheye (2009) defines land evaluation as a tool or technique that “assesses the performance of land based on a more or less systematic analysis of the physical land conditions and on the impact these have on present and alternative land use systems.” The land evaluation process is purpose driven and can be carried out at various scales; ranging from global and national level assessments through to detailed farm evaluations, all with different levels of quantification (Eliasson, 2007). Further, the evaluation process is an estimation or predication of potential use (Dent and Young, 1981; Rossiter, 1996) that should provide reliable scientific data and viable land use options, but does not in itself, determine the changes to be effected (FAO, 1976; Verheye, 2009). This process usually includes a form of soil-based survey and interpretation thereof, with aim of improving land use planning and decision making (Sonneveld et al., 2010; Manikandan et al., 2013).

Land evaluation has principally developed from soil survey interpretation and land classification (Beek, 1978). Furthermore, from both a historic perspective, as outlined in van Diepen et al. (1991), and within the context of this study, it is this soil and land resource survey, which draws the greatest attention. However, Sys (1985) importantly notes that soil and land resource surveys form only one aspect in the overarching field of land evaluation, which includes *inter alia* socio-economic, developmental and human resource factors.

Ultimately, when implemented within a holistic decision-making framework, agricultural land evaluation can optimise land use, reduce environmental degradation and improve productivity, all of which contributes towards long-term and sustainable food security (Smith, 2006; Bryant, 2017).

2.4 What Terminology is Commonly used within the Field of Land Evaluation?

Within the field of agricultural land evaluation, a number of subject-specific terms are commonly used and these terms should be pre-defined, in order to reduce misinterpretation.

2.4.1 Land capability

Land capability is an interpretive method of land classification and evaluation, which is determined by the combined effects of soil, terrain and climate (Scotney et al., 1991). It is a hierarchical classification that was originally developed to assist farmers with planning and conservation practices, where hazard of use is highlighted and the control of soil erosion prioritised (van Diepen et al., 1991). According to Smith (2006) land capability is concerned with the wise use of land in order for it to produce both economically and sustainability under explicit uses and treatments. The primary objective of this classification is the systematic arrangement of land to indicate its most intensive long-term use as well as its associated permanent hazards (Scotney et al., 1991).

The eight-class land capability classification method, developed by Klingebiel and Montgomery (1961) of the USDA, is the most well-known system and has served as a basis for many other attempts (Scotney, 1971; van Diepen et al., 1991; Schoeman et al., 2002;). Wherein a land capability class, groups land units with similar potentials and limitations (Schoeman et al., 2002).

The USDA method is viewed as a classical land capability approach and generally does not provide crop specific ratings, nor does it consider the dynamic nature of soil fertility in its classification (Klingebiel and Montgomery, 1961; Scotney et al., 1991). Adaptation of the USDA method, for use in other countries are numerous and include England, Wales and Scotland (Bibby & Mackney, 1969), New Zealand (New Zealand Ministry of Works, 1969) and India (Dahake, 1971). Other regionalised adaptions include Land Capability Classification for Tasmania (Hawkins, 1989), West Australia (Wells & King, 1989) and KwaZulu-Natal, South Africa (Camp et al., 1998).

From a National perspective, *A System of Soil and Land Capability Classification for Agriculture in South Africa,* was compiled by a task team appointed by the multilateral technical committee for Agriculture and Environment Affairs (Scotney et al., 1987) (Scotney *et al.* 1987). The system was later revised in 1991 to incorporate changes introduced by Taxonomic Soil Classification System, released in the same year (Scotney et al., 1991). A more detailed overview of both USDA Land Capability as well as South Africa's regionalised method is provided in Sections 2.6.2 and 2.7.2.

2.4.2 Land suitability

The term land suitability is most associated with *the FAO Framework for Land Evaluation* (1976), (cf Section 2.6.1). By definition, "land suitability is the fitness of given piece of land for a defined land use, with the degree of suitability being determined by the relationship between benefits and required inputs associated with that use" (FAO, 1976; Scotney et al., 1991). For example a tract of land may be highly suited for grain production but not suitable for the production of vegetables, due to the inputs required to obtain a beneficial yield (DAFF, 2018a). The land suitability criteria, as used in the example above, depends on the criteria for optimal use, with the most utilised criteria being maximum benefit and minimum losses (Beek, 1978).

A synthesised definition provided by (Rossiter, 1996), outlines the quantification and expression of land suitability in order to indicate land use fitness. In this he provides two options in terms of suitability expression, the first is on a continuous scale of "goodness" (e.g. 0 to 100) and the other a set of discrete classes, ranging from "completely suited" to "completely unsuited". In land suitability evaluation there is no "good" or "bad" land but only appropriate or inappropriate land uses (Eliasson, 2007), this is a major departure from land capability-based approaches.

2.4.3 Agricultural potential

Agricultural production potential is generally considered to be determined by physical land factors *inter alia* the quality of the soil, availability of water and the prevailing climate (Haverkort, 1988). Smith (2006) defines agricultural potential as "a measure of possible productivity per unit area and unit time, achieved through specific management inputs at farm level and is largely determined by the interaction of climate, soil and terrain." This potential can be linked to a range of beneficial land uses for given crop, for example maize, or veld type for primary production (DAFF, 2016).

2.5 What are the Governmental Legislative Policies which Guide Agricultural Land Management, Land Assessment and Planning in South Africa?

This section summarises pertinent land policies in South Africa and provides an initial legislative framework for the study.

2.5.1 The Subdivision of Agricultural Land Act

The Subdivision of Agricultural Land Act (SALA), commonly referred to as *Act 70 of 1970*, is enacted legislation to control the subdivision, and connection therewith, the use of agricultural land in South Africa (Republic of South Africa, 1970). The primary purpose of the Act is to prevent the fragmentation of agricultural land into non-viable economic units. The Act is also used to protect agricultural land from non-productive land uses through unsustainable land use change, such as the expansion of residential developments onto farmland. SALA essentially provides a measure of legislative control in order protect agricultural land and the production thereof.

One of the core issues in the Act and subsequent amendments is the actual definition of “Agricultural Land”. In terms of the Act (Republic of South Africa, 1970:1) all land in South Africa is agricultural land except:

- “land situated under the jurisdiction of a local council, for example a municipality, or local board such as village management board, or health committee, land included in Ordinances and specified land excluded by the Minister in the *Government Gazette*.; or
- land which forms part of any area subdivided in terms of the Agricultural Holdings (Transvaal) Registration Act, 1919 (Act No. 22 of 1919); or
- land which is a township as defined in section 102(1) of the Deeds Registries Act, 1937 (Act No. 47 of 1937), but excluding a private township; or
- any state-owned land or any land held in trust by the State or a Minister for any other person; or
- any land that the Minister has excluded from the Act by notice in the *Government Gazette*.”

It is also important to note that the exemptions listed above also include land held by Traditional Councils, such as the Ingonyama Trust Board, which has jurisdiction over nearly a third of rural land in the Province of KwaZulu-Natal (CLS, 2015). Servitudes and land owned by parastatals such as Eskom and Transnet are excluded from provisions outlined in SALA.

Critically SALA prohibits a number of related actions on agricultural land on which it applies. This includes the "subdivision of agricultural land, the transfer of agricultural land into undivided shares, leasing of agricultural land for longer than ten years, the establishment or extension of a development area or area under local jurisdiction and the development of a land use scheme on agricultural land" (Republic of South Africa, 1970:3). SALA, where applicable, is also used to specifically control the release of farmland to a non-agricultural use. The potential release is done on application-by-application basis, for each individual farm portion, where the impact of losing that particular farm portion is ultimately assessed by the National Department of Agriculture with advisory recommendations emanating from their Provincial counterparts.

SALA is a polarising piece of legislation and has received notable criticism from a legislative, applicability and planning perspective (e.g. Brink, 2015). Many quarters see SALA as a hindrance to rural development and land reform, a vital issue in South Africa's new social and political landscape (Ramothar et al., 2021). An application for sub-division or a change of land use is often associated with long application lead times, which slows development and delays land reform projects. SALA also lacks a comprehensive set of norms and standards, which can lead to inconsistent decisions and lengthy appeals. Furthermore, SALA is often in conflict with more contemporary planning legislation (e.g. Formal Township Establishment Act of 1991, Physical Planning Act of 1991, Development Facilitation Act of 1995 and finally the encompassing Spatial Planning and Land Use Management Act of 2013).

These problems and aforementioned issues regarding the broad definition of agricultural land culminated in SALA being repealed by the Subdivision of Agricultural Land Repeal Bill of 1997, which was assented into a law in September of 1998. Critically the repeal comes with the provision and that it will only come into effect on a date fixed by the President by proclamation in the Gazette (Republic of South Africa, 1997). This proclamation has yet to occur, resulting in SALA still currently being utilised by the national Department of Agriculture (Collett, 2009). The major reason for the delay is that there is currently no suitable legislation to replace SALA, although the Preservation and Development of Agricultural Land Bill (cf Chapter 2.5.5) is envisioned to ultimately repeal SALA.

2.5.2 The Conservation of Agricultural Land Act

The Conservation of Agricultural Land Act of 1983 (CARA) was assented to law in April 1983 and came into effect on the 1st of June the following year (Republic of South Africa, 1983). This Act repealed a number of Acts and associated amendments including the Weeds Act (Act 42 of 1937) and the Soil Conservation Act (Act 76 of 1969). From an agricultural perspective CARA is a critical piece of legislation as it directly addresses the sustainable use and protection of agricultural resources. CARA aims to conserve the natural resources of the South Africa, particularly within the agricultural, non-urban, landscape. The objectives of the Act are specifically to provide "*for the conservation of the natural agricultural resources by the maintenance of the production potential of land, by the combating and prevention of erosion and destruction of the water sources, and by the protection of the vegetation and the combating of weeds and invader plants"* (Republic of South Africa, 1983).

At its heart, the Act provides control measures, backed by a detailed suite of regulations (Republic of South Africa, 2001) for a wide range natural resource issues *inter alia* veld management, weed and erosion control, sustainable grazing, fire use, water resource protection and rehabilitation. From a land perspective CARA provides specific controls for the cultivation of virgin soil, the utilization and cultivation of land and the irrigation of land.

As per the Act, virgin soil is defined as soil that has not been previously cultivated or has not been actively cultivated for ten consecutive years, land meeting these criteria cannot be broken without official approval. Approval of cultivation of virgin land is granted after review process, which could include the digging of soil observation pits to assess the lands inherent potential to cultivated agriculture. The Act also limits the maximum permissible slope which can be cultivated. This ranges between 12% and 20%, depending on the locality and inherent soil properties (Republic of South Africa, 1983).

The Act also empowers the overseeing Authority to issue directives to non-compliant land owners in the form of fines and/or imprisonment. On face value the Act and control regulations provide significant protection to agricultural resources however the key to all legislation is its implementation and enforcement. Unfortunately, the lack of functioning Conservation Committees and shortage of Resource Auditors means that non-compliance is rarely punished and ultimately it is the resource base that suffers through degradation (Collett, 2009).

There is however, some overlap between the CARA and the regulations National Environmental Management Act, 107 of 1998, including the transformation of virgin land, which includes cultivation. Collett (2009) notes that this duplication of responsibilities needs to be addressed and this remains the case.

2.5.3 The National Environmental Management Act

The National Environmental Management Act (NEMA), Act 107 of 1998, and various amendments provides for co-operative environmental governance in South Africa (Republic of South Africa, 1998, 2008).. The Act offers a holistic framework, detailed guidelines, procedures and enforceable regulations in order to allow for sustainable environmental management.

According to the Act, "*everyone has the right to have the environment protected through reasonable legislative measures that prevent pollution and degradation, promote conservation and secure ecological sustainable development"* (Republic of South Africa, 1998, 2008). Although not strictly agricultural legislative, NEMA, due to its holistic nature and overarching jurisdiction does provide a certain level of protection to important agricultural resources such as soils, indigenous vegetation, agricultural land, wetlands and peat soils. The Act is backed by three Listing Schedules for various activities. Activities listed in the Schedules requires Environmental Authorisation prior to the commencement of the activity. Activities that pertain to agricultural resources include soil excavation, clearance of indigenous species, transformation of land, alteration of virgin soil and land development, where the land was previously used for agriculture. NEMA acknowledges CARA and when properly enforced can provide a suitable level of protection to soil and vegetative resources from degradation.

2.5.4 The Spatial Planning and Land Use Management Act

The Spatial Planning and Land Use Management Act (Act 16 of 2013), SPLUMA, provides a framework for all spatial and land use management in South Africa. SPLUMA aims to reduce the historical uncertainty and ambiguity that surrounds land use planning in South Africa's complex urban and rural landscape (Republic of South Africa, 2013). SPLUMA promotes both integrated planning at all levels of Government, while also achieving constitutional imperatives, such as environmental protection, protection of property rights and the right to sufficient food and water. From an agricultural perspective SPLUMA does provide some protection to agricultural land, including land under the control of Traditional Authorities, where Act 70 of 1970 does not have jurisdiction. Further, SPLUMA is founded on a number of sound planning principles including spatial sustainability, where special consideration is given to the protection of prime and unique agricultural land as well as its sustainable use (Republic of South Africa, 2013).

When SPLUMA was enacted it repealed a number of parallel planning legislations and included the repealing of the Removal of Restrictions Act (84 of 1967), Physical Planning Act (88 of 1967), Less Formal Township Establishment Act (113 of 1991), Physical Planning Act (125 of 1991) and the Development Facilitation Act (Act 67 of 1995). However, and importantly from an agricultural perspective, SPLUMA did not repeal Act 70 of 1970. Both pieces of legislation are thus in effect and are often in direct conflict with each other, when being applied to a common application. An example of this is that SPLUMA requires Municipalities to create "wall to wall" land use schemes, which includes land regulated by Act 70 of 1970. Act 70 in-turn, does not allow land use schemes to be placed on agricultural land, thus creating a paralysed feedback loop. These planning conflicts have placed ever increasing pressure on the National Department of Agriculture to replace the problematic Act 70 of 1970 with legislation that is more compatible with contemporary planning legislation. This has culminated in the development of the Preservation of Agricultural Land Bill.

2.5.5 Preservation and Development of Agricultural Land Bill

The Draft Preservation of Agriculture and Development of Agricultural Land Bill (PDALB) was released for comment in 2016 and once enacted will fully repeal SALA, (Act 70 of 1970). The primary aim of the Bill is to provide for the protection of agricultural land. The Bill and underlying regulations are far more detailed than its precursor and not only covers agricultural subdivision and change of land use but also provides additional rules, norms and standards, institutional frameworks, dispute mechanisms and coordinated planning guidelines. The Bill also attempts to reduce conflicts with existing planning legislation while increasing the scope of its applicability, especially in areas where SALA was not applicable. Further the Bill states that if there is conflict the PDALB will prevail if the conflict directly concerns the management and development of agricultural land (DAFF, 2016). Critically the Bill recognises that it is in the National interest to preserve and, promote the sustainable use and development of agricultural land. Further, it recognises that high value agricultural land is rare, is under developmental pressure and that it is in the best interest to have agricultural land protected (DAFF, 2016).

PDALB importantly provides detailed definitions, pertaining to critical terms used in agricultural planning, which have in the past caused significant confusion and conflicting legal viewpoints. This includes definitions pertaining to, *inter alia*; "agricultural land", "agricultural potential", “high value agricultural land” and “unique agricultural land”. Clarity and consistency of use of these terms, which are often used interchangeably, allows for improved classification and regulation development (DAFF, 2016).

PDALB also provides guidance in terms of agricultural land classification “where the Minister of Agricultural may establish a system of land capability classification, within an appropriate land evaluation framework, for determining the physical capability of land at national, regional and local scale”. This evaluation is aimed at classifying land in order to determine its most intensive long-term use. Along with physical capability the Minister may determine agricultural land zones according to its suitability for a variety of agricultural activities (Section 8.1, Para 1-3, DAFF, 2016).

The Draft Bill also stipulates that high level Agricultural Plans be created at Municipal level to ensure important agricultural areas are ring fenced and protected from non-productive land uses. This protection can culminate with a particular area being deemed a *Protected Agricultural Area*, which will make release of agricultural land, to other land uses, within this protective buffer extremely difficult. The proclamation of a Protected Agricultural Area will have the same binding effect as they do with Ecological Conservation Areas, where protection is actually reflected on the land’s title deed.

In terms of an application for farm subdivision or change of land use the Bill stipulates that an agro-ecosystem report, which complies with certain standard should be compiled. This report will include a land evaluation in terms of land capability and suitability, the impact of the development application on surrounding farmland as well impact on landscape character. All with the aim to retain productive land, reduce fragmentation and safeguard food security. The finalisation of the land evaluation framework, methodology for farm level capability and suitability assessment as well the agro-ecosystem report will fall to the Minister of Agriculture when the Bill is enacted. This is a critical point as a suite of robust and scientifically based methodologies for farm evaluation will need to be available for consideration. To date there is still no approved National Framework or norms and standards for the submission, consideration and approval/rejection of application for subdivision and/or change in land use of agricultural land. This leads to an uncoordinated and inconsistent approach to decision making across various governmental departments (DAFF, 2016).

Ultimately, current governmental and legislative failures are leading to the loss of both productive and potentially productive agricultural land, decreasing overarching agricultural sustainability and the degrading its resource base.

2.5.6 Status Quo: Applicable Legislation

At this point in time although repealed, SALA is still in effect as there is no ratified piece of legislation to replace it. Despite the contestation surrounding the definition and demarcation of “Agricultural Land” the Constitutional Court ruled that national control over agricultural land remains in place despite the creation of transitional councils (Constitutional Court of South Africa 2008; Collett, 2008) with this mandate falling to Department of Agriculture. SALA is thus still utilised to control the subdivision and use of agricultural land in South Africa. CARA and NEMA are both in effect and apply beyond the bounds of what is deemed “Agricultural Land”. CARA and NEMA both aim to preserve environmental integrity and its natural resources though regulation of activities. SPLUMA is also active legislation, which acknowledges the need to protect prime and unique agricultural resources. SPLUMA cannot be completely adhered to due to its non-compatibility with SALA, often creating a paralysed feedback loop (cf Section 2.3.4). There is thus a need for the development of new legislation to replace SALA, with PDALB being the forerunner to do this. PDALB, when enacted, will completely repeal SALA with the aim to reduce legislative conflicts and also provide more holistic protection for agricultural resources in South Africa. Recent legislative policy, proposed through the PDALB (2016), aims to establish a broad framework to classify rainfed agricultural land according to the most intensive long-term use thereof determined by the interaction of climate, soil and terrain. This proposed framework intends to include methods for “*determining physical capability at national, regional and local scales*” PDALB (2016).

Ultimately, contemporary and reliable land evaluation methodologies are critical to support land and agricultural related legislation, whose primary aim should be to accurately classify, delineate and protect valuable and unique resources for sustainable food production and security.

2.6 What are the Primary Methods used to Assess Agricultural Land Internationally?

There is no single or universal method, which is suitable for all land assessment applications (Beek, 1978; Rossiter, 1994). Thus, the selection of the most appropriate tool, or suite of tools, is therefore paramount to obtain the most accurate and relevant results. There are, however, selection criteria which can narrow ones focus and assist to provide a smaller set of suitable methods to choose from. The evaluation objective, the scale of applicability, and the field of study can all be used to eliminate tools, which are not suited for a particular application (Rossiter, 1994; Eliasson, 2007).

For the purposes of this review it is impractical to evaluate all available land assessment methods and tools. Therefore a total of six tools were selected based, on the following criteria, which chiefly correspond to the objectives of this research:

- *The tool must be a non-regionalised model or assessment framework, which aims to provide a predominately soil based evaluation of non-irrigated cropland with an applicability scale which includes site, farm or landscape level assessments.*

This criteria eliminates tools such land degradation methods, tools for purely irrigated or non-cropping land uses such as forestry and rangelands. Regionalised methods which are non-transferable or methods that are only applicable to a scales ranging from Global to District level were also excluded. Many of the tools reviewed in Sections 2.6.1-2.6.6 were originally developed decades ago, however they provide a sound basis for agricultural land evaluation.

2.6.1 FAO Framework for Land Evaluation and associated land use guidelines

In the early 1970's, a need for some form of global standardisation and transferability, in the field of land evaluation, had been emphasised (Beek, 1978). This standardisation came in the form of a *Framework for Land Evaluation,* which was developed through a consultative approach, by the FAO and published in FOA Soils Bulletin 32 (FAO, 1976). The Framework presented in this Bulletin was not an evaluation system *per se*, but rather a set of principles and concepts, which can form the basis of an area-specific evaluation methodology, with global applicability (FAO, 1976). These broad principles were further advanced in more detailed land evaluation publications and guidelines for specific land uses including Rainfed Agriculture (FAO, 1983). These land use specific Guidelines sit midway between the broad Framework for Land Evaluation and the more detailed region specific manuals and include 25 land qualities which should be considered when evaluating land for rainfed agriculture (FAO, 1983).

The Framework for Land Evaluation is based on the concept of land suitability, where land should be assessed and classified based on specific kinds of sustainable use. A multidisciplinary approach, which includes contextualising economic and social issues, is also demanded, as well as comparisons of more than one kind of use (FAO, 1976). The Framework provides two analytical approaches to land classification, namely the parallel and two-stage approach. The parallel approach analyses both the physical and socio-economic aspects concurrently. While, the more pragmatic two-stage approach, evaluates the lands physical potential in the first stage, followed by an economic and social analysis, in the second (FAO, 1976). Verheye (2009) suggests that although both approaches have their own unique advantages, in reality the two-stage approach has been applied more often, as the need for physical potential evaluation is inevitable and is ultimately more quantifiable, than socio-economic factors. Furthermore, he concludes that often land assessment is finalised at the physical evaluation stage, without conducing the secondary social or economic analysis.

After initial surveys and consultations, land is classified via the suitability of various land uses, by matching and comparing these uses, with their differing requirements and limitations, to land mapping units and their associated land qualities. The level of detail associated with the resulting suitability classification is flexible and is dependant on the objectives and scale of the land evaluation process. The Framework provides four levels of detail, via a hierarchical classification, ranging from broad Suitability Orders through to the detailed Suitability Units (Rossiter, 1994). Suitability Order for example, simply has two classes namely *S* and *N*, S denotes that specific land portion is suitable for a particular land use, while N denotes that it is not suitable.

Although, the Framework has not gained much traction in South Africa it has been successfully applied in various countries and regions across the globe, and for a multitude of land uses. Early applications of the Framework include evaluations in annual and perennial crop production in Brazil, oil palm production in Surinam and land utilisation type in Kenya as presented in FAO (1976) and Beek (1978), more recent national applications and assessment include Zambia (Chinene, 2007) and Iran (Bagheri Bodaghabadi et al., 2015). Technical advances in computing, simulation modelling, remote sensing and GIS have all contributed towards automating and advancing the original Framework. Examples include the Automated Land Evaluation System: ALES (Rossiter, 1988 and 1990) GIS based crop-land suitability analysis using neural networks (Bagherzadeh et al., 2016) and soil site suitability analysis through geo-statistics (Mandal et al., 2020).

While, the Framework for Land Evaluation is recognised as one of most important and widely used FAO methodologies, in the field of land resource management (FAO, 2007; Rossiter, 2009), it is not without its limitations. A detailed review by Van Diepen et al. (1991) highlights a number of operational and philosophical constraints surrounding the original Framework. Perceived Framework deficiencies and more recent paradigm shifts in land evaluation, culminated in the writing of a discussion paper entitled “Land Evaluation: Towards A Revised Framework” (FAO, 2007), which details various shortcomings and potential changes, which could assist in updating the Framework in order to include more ecological based concerns and technological advancements. At this time no updated FAO Framework or guidelines have been published but founding principles and associated guidelines should be considered an important reference document for all land evaluation practitioners.

2.6.2 Agricultural Land Evaluation and Site Assessment

Agricultural Land Evaluation and Site Assessment (LESA), originally released in 1981, is an analytical framework developed by the United States Department of Agriculture, Natural Resources Conservation Service (USDA: NRCS) to determine the quality of agricultural land for productive land uses as well as their agricultural economic viability (USDA, 2011). The system is designed to be transparent, defensible, repeatable and adaptable enough to accommodate different environments and local situations (Qian et al., 2021). Further, it was developed to assist decision makers when comparing two sites based on their agricultural value (Pease and Coughlin, 1996). LESA has been applied in a number of States, regions and counties across the United States of America (Steiner et al., 1987), examples of state-wide available LESA systems are California, Delaware, Hawaii and Connecticut.

LESA consists of two separate components, namely Land Evaluation (LE) and Site Assessment (SA), providing systematic and objective methods to quantify sites by agricultural importance (FAO, 2021). This dual rating approach is common to all LESA models, however the individual factors that are selected can vary considerably, to meet spatially specific needs and conditions (California Department of Conservation, 1997). The land evaluation (LE) process within LESA revolves around soil survey outcomes, where soils are given area-specific ratings and grouped from best to worst suited, for a stated agricultural use. The site assessment component (SA) alternatively identifies non-soil related factors, that influence the quality of the land for agricultural use. These selected factors are also subsequently graded to meet the needs and objectives of the local assessment method (USDA, 2011). Site assessment factors include limitations to agricultural productivity, development pressures, and factors measuring other public values (Southeastern Wisconsin Regional Planning Commission, 2007). The ratings for each component are then aggregated to provide an overall agricultural value of the site for the specified agricultural use, such as cropland, forest land or rangeland. Both the LESA handbook (USDA, 2011) and guidebook (Pease and Coughlin, 1996), were developed to assist in creating area specific LESA systems. Both documents stress the need for local participation from planners, extension officers, agricultural scientists and farmers, in order to create a local LESA committee and ensure local input and buy-in.

From a cropland specific perspective LESA stresses the need for an integrated approach to the land evaluation component, where a combination of accepted methodologies is ideally applied, rather than a relying on an individual method. Accepted methodologies for land evaluation as provided in LESA handbook (USDA, 2011) and guidebook (Pease and Coughlin, 1996) include:

- Land Capability Classification, based on the classical eight class USDA approach, identifies the degree of agricultural limitation inherent in soils of the study area (cf Chapter 2.4.1, 2.7.2 and 2.7.3). Additional soil-based limitations are usually designated through the use of capability sub-classes, for example Class II e, would be defined Land Capability Class II, due to soil erosion limitations
- Soil Productivity Ratings, considers soil productivity in terms of predicted yield for a specified indicator crop(s), which is commonly grown on the study area, e.g. maize. This rating is often used as a dual indicator, for not only soil productivity but potential economic returns. Soil productivity ratings are similar to the ultimate aim of the KwaZulu-Natal ecotope concept which links soil properties to crop specific yields (cf Chapter 2.7.4).
- Soil Potential Ratings, indicate the relative quality of a given soil based on the previously determined soil productivity ratings, as described above. The locally established standard yield for a specified indicator crop(s) is compared with the costs associated with overcoming inherent soil limitations as well as any continuing limitations, such as soil fertility corrections (cf Chapter 2.6.5).
- Important Farmland Classification, standardises national criteria and definitions, which in turn allows local planners to consider overarching national efforts to protect prime and unique resources. This classification provides a consistent basis for comparing land in different areas. The minimum soil-based criteria to achieve the designation of “prime farmland” is very specific and includes limits such as minimum soil depths, permissible soil acidity, soil permeability ranges and even permissible soil temperatures.

Once the most appropriate land evaluation method or suite of methods are selected and applied for the study area, they are subsequently rated, along with the selected site assessment factors using a numerical scoring system. For example, if land capability classification was selected then Class I land, with no significant limitations, could equate to a score of 100 for this specific land evaluation factor. These ratings are then assigned a relative weight to recognise differing factor importance. Finally, these weighted scores for all selected land evaluation and site assessment factors, are tallied to obtain an overall LESA score, to which thresholds are assigned to assist land use decisions (Southeastern Wisconsin Regional

Planning Commission, 2007; FAO, 2021). A site with a high LESA score indicates that is important from an agricultural perspective and a proposed change in land use would not likely be granted from planning authorities.

In 2006, the American Farmland Trust recognised the LESA system as the best agricultural land suitability tool available in the United States, which allows users to evaluate agricultural importance and encouraged dialogue between a diverse groups of stakeholders. The Trust did however acknowledge some drawbacks linked to the system, which include the extended time taken to develop and field-test a usable system, scoring inconsistencies and the need for regular re-evaluation.

From a South African perspective, where natural resource diversity is high, the flexibility of the LESA system is attractive as well as its integration of multiple land evaluation methodologies. This integration reduces the risk of an over-dependence on single land evaluation method, which may not be applicable in all situations. However, the lack of detailed soil information at local level in South Africa may be problematic as outlined by van Zijl and Botha (2016), particularly when developing soil ratings across large areas and for multiple crops.

2.6.3 Fertility Capability Classification

The Fertility Capability System (Buol, 1972; Sanchez et al., 1982, 2003), commonly abbreviated to FCC, is a technical land soil evaluation system which aims to group soils based on the physical and chemical limitations they present for agronomic production and management. According to the developers of the system, FCC attempts to "*bridge the gap between soil classification and soil fertility*".

FCC focuses on quantifiable plant growth parameters in both the top- and subsoil and consists of three categorical levels: type (topsoil texture), substrata type (subsoil texture) and modifiers (15 defined parameters) (FAO, 2020). Type and substrata type are relatively self-explanatory and are limited by either plough layer depth in the topsoil or by specified depths. The third categorical level is defined by 15 unique abbreviated modifiers, which include both physical and chemical criteria. Physical modifiers include, for example, the presence of gleyed (wetness) or vertic (structural) characteristics. While chemical modifiers include, criteria relating to low cation exchange capacity, salinity and aluminium toxicity. The parameters provided in FCC are strictly defined and correspond with quantitative limits provided by either the USDA Soil Taxonomy classification system (USDA, 1975) or the FAO-Unesco Legend of the Soil Map of the World (FAO-Unesco, 1974)

Soils are ultimately assigned an FCC code, where type and substrata type are presented in capital letters while modifiers are listed in lower-case, with the most prominent modifier appearing first. For example many well-drained soils (oxisols) would be assigned an FCC code of ***C a e I k*** which is defined as a clayey soil associated with aluminium toxicity, low cation exchange capacity, high Phosphorous fixation and low Potassium Reserves (Sanchez et al., 1982). Later publications, relating to FCC, also include an interpretation of the various classification codes which aids to managing the various limitation presented (Buol, 1986).

The FCC, although developed in the tropics (Buol, 1972) has been widely used and evaluated in other climatic areas (Sanchez et al., 1982, 2003). These evaluations showed that the FCC system is a useful tool for relating physical and fertility limitations to crop yield responses in a wide variety of soils and crops. One of the limitations of this method that it does not take terrain variability into account.

From a South African perspective this classification has some potential in terms of local application. Primarily as the classification can be specifically adapted, where existing modifiers and their associated limits are calibrated to local conditions or new local quantitative modifiers are introduced to the system. However, the concept of continually changing fertility levels in a typical crop production environment needs to be developed further, to make this a reliable method in land use change or release applications.

2.6.4 The Storie Index and Revised Storie Index Soil Rating

The Storie Index, developed by Professor Earl Storie at the University of California, is an adaption of his previous work which developed *An Index for the Rating of Agricultural Value of Soils* in the 1930's. The Index is a semi-quantitative soil rating method and evaluates a land's potential utilisation and productive capacity (Storie, 1978). The last version of the original Storie Index was published in 1978 and is widely used, particularly in the State of California in the United States of America (USA), where it was developed.

The Index uses four characteristics to which percentage values are assigned, with 100% being the highest possible score. The four characteristics used in the Index are: Factor A, the rating of the soil profile characteristics; Factor B, the texture of the soil surface; Factor C, the slope of the land and finally Factor X, the rating of other soil and terrain features including drainage, alkalinity, fertility levels, acidity, erosion and microrelief. The four factors are then multiplied together, with an equal weighting, to calculate the Storie Index. In California the method is supported by further sub-divisions in terms of soil profile groups and soil gradings, with the intention to improve method accuracy and its interpretation (Storie, 1978).

The original Storie Index, is a comparatively simplistic land evaluation method, so far as that its factors are relatively easy to determine. The Index does, however, have a number of fundamental issues. For example if one wants apply the method outside California, then new rating tables are required, with local calibration for soil type factors and soil grading groupings. Further drawbacks of the method include that it is predominately used in irrigated land use scenarios and that the ratings are highly subjective, depending on evaluator's background and experience (FAO, 2020). These problems reduce the overall reliability and repeatability of the Index which has given rise to the development of the Revised Storie Index (O'Geen & Southard, 2005; O'Geen et al., 2008).

The Revised Storie Index attempts to reduce individual user bias by developing an overarching model to calculate the ratings digitally. The model is based on combinations of discrete and fuzzy logic functions to obtain more reliable scores for the factors associated with the Storie Index (O'Geen & Southard, 2005). The Revised Storie Index uses the National Information Soil System (NASIS), an American computerised database of soil information, as its primary input. The NASIS-derived Index is seen as a rapid and non-bias method of converting soil survey data to The Storie Index Ratings (O'geen et al., 2008). Although an improvement, the lack of transferability of the revised method, outside the NASIS coverage, remains a major drawback for potential international users and applications.

2.6.5 Soil Potential Ratings

Soil Potential Ratings (SPR) is a land evaluation approach, which has been refined for use by the Natural Resource Conservation Service (NRCS) within the USDA. The procedural method is detailed in Part 621 of its National Soil Survey Handbook, Title 430, with the latest amendments to the method being published in 2018 (USDA, 2019). SPR is essentially a class-based rating system that provides an indication of the relative quality of a soil, for a defined use, as compared with other soils within a specified area. Five generic classes are provided for comparative rating of soil potential: very high, high, medium, low and very low. However, the final number of classes used depends on the range of potentials in an area, with more homogenous areas requiring fewer classes (USDA, 2019). SPR can be developed for any geographic are, regardless of scale of unit of mapping. SPR has been developed to evaluate not only agricultural uses, such as cultivation, but also non-productive land uses such as dwellings and septic tanks. However for the purposes of this review its focus, in terms of SPR, will be its application in an agronomic environment. As per the NRCS Handbook SPR considers the following when assigning a rating:

(1) Locally established yield or level of performance;
(2) The relative cost of applying modern technology to minimize the effects of any soil restrictions;
(3) The adverse effects of continuing limitations, if any, on social, economic, or environmental values (USDA, 2019, 621-A.1).

The aforementioned rating classes are based on a numerical Soil Production Index (SPI) for each soil for a specified crop. The SPI is expressed by the equation:

$$SPI = P - (CM + CL) \qquad (\textit{Eq. 2-1})$$

Where:
P = Index of performance or yield as a locally established standard
CM = Index of costs of corrective measures to minimize the effects of soil limitations
CL = Index of costs resulting from continuing limitations (USDA, 2020, 621-A.5).

The index for each soil is normalised, generally between 0 and 100, against the average yield from the most productive soil in the study area. SPR differs from other traditional evaluation procedures as it uses observed yields, in combination with relative production costs, as indicators for soil productivity (van Diepen et al., 1991). This is an important factor as even though two soils may have the same yield, their SPR may differ due to differences in establishment or maintenance costs. CM and CL, the cost indices, are general in nature and can be based on percentage of the cost, where highly detailed economic analyses are not required (USDA, 2019). Corrective measures can also be once off or continuous such as the need for fertilizer application over and above the normal rate.

Critically ratings for a particular use are established for a specific area and thus ratings cannot be reliably transferrable to other areas, where criteria may differ. The ratings should be used principally for planning purposes as well as to provide an indication of relative soil suitability rather than a recommendation for soil or land use. In this regard, they can assist land use planners to prioritise areas, which need to remain under agriculture (Rossiter, 1994).

The development of a reliable set SPR area requires significant data and including the collection of soil data, long term yield data, performance levels as well as costs associated with establishment and typical corrective measures. The SPR method should be seen more as a guiding framework and is not a tool that can be used “out-the-box” or in a previously unstudied area without significant input from specialists and land users. Its development and continued application is also an iterative process where the performance and cost indices can be adjusted over time.

The procedures used for rating soil potentials allow for maximum flexibility however, the benefit of flexibility comes with the cost subjectivity. All parts of the SPR system are locally derived, thus its actual ratings can be subjectively adjusted and even the soil rankings can be influenced by the weighting given to the costing indices versus the performance index (Rossiter, 1994). This subjectivity can lead to inconsistent rankings across different areas. Ultimately, SPR are also not a standalone product, but should be used to supplement other classifications and interpretations (USDA, 2019).

2.6.6 Visual Soil Assessment

Visual Soil Assessment (VSA) is a scoring approach based on the visual evaluation of key soil and plant performance indicators of soil quality (FOA, 2020). The supporting VSA documentation, presented in a 504-page field guide published by the FAO, is an adaption from the original methodology developed by Shepherd et al. (2008). The rationale for a visual approach is that many soil properties, particularly physical soil properties present themselves through visual indicators. These indicators can be reliably identified infield with little training and equipment. Further advantages of VSA is that it is both a rapid and economical approach to evaluate soil quality, especially when compared to more quantitative approaches (e.g. laboratory analysis). Research has shown that many visual indicators are closely correlated to quantitative indicators of soil quality (FOA, 2020).

Unique VSA guides (FOA, 2008) have been developed for both broad land cover categories (annual crops, pastures and orchards) as well as some individual crops (maize, olive orchards, vineyards and wheat). Indicator scoring and their weighting varies depending on the land cover category or specific crop being assessed. The visual indicators of soil quality used in the VSA include soil texture, soil structure, soil colour, soil smell, degree mottling, presence of earthworms, potential rooting depth, surface ponding, surface crusting, surface relief and signs of soil erosion. These indicators are scored, weighted, ranked and summed to provide an overarching soil quality index. This index is used to class the soil as either poor, moderate or good in terms of soil quality for that particular land cover or crop. Soils with good VSA scores will generally give the best yields with the lowest establishment and operational costs (FOA, 2008).

With the exception of soil texture, the aforementioned soil indicators are dynamic, i.e. capable of changing under different management strategies and land uses. Regular assessments are required to gain insight into how these dynamic soil quality indicators are changing over time (FOA, 2008)

Within a South African context, the relatively simple, economical and rapid nature of the VSA methodology is a seen as a major positive. However, the lack of local verification studies, which compare VSA scoring to soil properties and actual productivity has curtailed its local application.

2.7 What are the primary methods used to assess agricultural land in South Africa?

2.7.1 Historical background

The 1923 Drought Investigation Report authored by Heinrich du Toit found that inadequate land assessment was one of the primary reasons for the notable degradation of the natural resources in South Africa (du Toit, 1923). Over the years the need for proper assessment of agricultural land has seen development of various systems.

Schoeman et al. (2002) states the E.C.M Code, first document by Loxton (1962), is the earliest formal system developed for South Africa, from a land capability perspective. The E.C.M code refers to erosion hazard, soil climate and mechanical limitations and was used in the author's implied soil-survey procedure for farming planning. The concept of agricultural land evaluation, through the use of land capability, was further developed and tested by Ludorf (1970) and (Scotney, 1971). At that time no single method was being used, but rather individual, regionalised methods, leading to confusion between users (Laker, 2004). A review by (Van Niekerk, 1983) found that there was no formal or coordinated national effort to develop a standardised approach to land capability evaluation in South Africa as cited in Laker (2004). Ultimately, a standardised approach to land capability in South Africa was required.

2.7.2 South African land capability

To meet the need for a universal capability approach to land evaluation, *A System of Soil and Land Capability Classification for Agriculture in South Africa* was compiled by a task team appointed by the multilateral technical committee for Agriculture and Environment Affairs (Scotney et al., 1987). The system was later revised in 1991 (Scotney et al., 1991) to incorporate changes introduced by Taxonomic Soil Classification System, released in the same year (SCWG, 1991).

Like many other regionalised adaptions, the South African system is based on the USDA Land Capability Classification System developed by Klingebiel and Montgomery (1961). It is a good example of a maximum-limitation classification system, as defined by de la Rosa and van Diepen (2002) as it "*combines soil, terrain and climate factors to classify land in terms of its intrinsic hazards and best sustainable, long-term use*" (van Zijl et al., 2020). It recognises eight land capability classes, with Class I being the most favourable for arable land uses, while Class VIII has the greatest limitations, which preclude its use from commercial agricultural production (Schoeman & Scotney, 1987). These eight classes can also be grouped according to broad land utilisation, where classes I-IV denotes arable land, IV-VII grazing land and class VIII being reserved for wildlife and conservation. It assumed that soils suited for intensive uses such as crop production, Classes I-IV, would also be suitable for less intensive land uses such as grazing.

The system is broad and can be used throughout South Africa but does contain a number of important assumptions, which limit its usefulness and applicability. The South African system was specifically developed for rainfed agriculture applications and classifies land on its present limitations and is not crops specific. Good land management and the requisite soil conservation measures is also assumed, while fertility status and economic factors do not form part of the assessment process (Schoeman & Scotney, 1987).

The South African system uses two terrain factors, five soil factors and one climatic factor, with the least favourable factor ultimately determining the Land Capability Class for a given piece of land. This is classification is achieved using an elimination key (Table 2-1).

Table 2-1 Elimination key to Soil and Land Capability Classes (replicated from Scotney et al., 1991)

<table>
<tr><th rowspan="2">SOIL CAPABILITY CLASS</th><th colspan="2">TERRAIN FACTORS</th><th colspan="5">SOIL FACTORS</th><th rowspan="2">CLIMATIC FACTORS</th><th rowspan="2">LAND CAPABILITY CLASS</th></tr>
<tr><th>Erosion Hazard</th><th>Flood Hazard</th><th>Effective Depth</th><th>Texture</th><th>Internal Drainage</th><th>Mechanical Limitations</th><th>Other</th></tr>
<tr><td>1</td><td>E 1</td><td>F 1</td><td>D 1</td><td>T 1</td><td>W 1</td><td>M 1</td><td>O 1</td><td>C 1</td><td>I</td></tr>
<tr><td>2</td><td>E 2</td><td>F 2</td><td>D 2</td><td>T 2</td><td>W 2</td><td>M 2</td><td>O 2</td><td>C 2</td><td>II</td></tr>
<tr><td>3</td><td>E 3</td><td>F 3</td><td>D3</td><td rowspan="6">T 3</td><td>W 3</td><td>M 3</td><td>O 3</td><td>C 3</td><td>III</td></tr>
<tr><td>4</td><td>E 4</td><td>F 4</td><td>D 4</td><td>W 4</td><td>M 4</td><td>O 4</td><td>C 4</td><td>IV</td></tr>
<tr><td>5</td><td>E 5</td><td rowspan="4">F5</td><td rowspan="4">D 5</td><td rowspan="4">W 5</td><td>M 5</td><td rowspan="4">O 5</td><td>C 5</td><td>V</td></tr>
<tr><td>6</td><td>E 6</td><td>M 6</td><td rowspan="3">C 6</td><td>VI</td></tr>
<tr><td>7</td><td>E 7</td><td>M 7</td><td>VII</td></tr>
<tr><td>8</td><td>E 8</td><td>M 8</td><td>VIII</td></tr>
</table>

Erosion and flood hazards are used to characterise terrain hazards and incorporate slope, soil and wind erodibility, flood frequency and flood duration. Soil factors include depth, textural and wetness limitations, hazards linked to stone and bedrock exposures, the presence of erosion, slope limitations to mechanical implements as well tillage and salinity limitations.

Climatic criteria, within the growing season, incorporate rainfall, evaporation, minimum and maximum temperatures as well as limitations linked to frost, wind and hail hazards. When assessing climatic factors, the system provides a dual approach, whereby permanent terrain and soil factors can be used to define *soil capability*, while with the addition of climatic factors can then be used to determine *land capability* (Scotney et al., 1991). However, the authors recommend the use of the overarching land capability classification in rainfed production environments, while soil capability can be used to comparatively rate soil capability of different areas.

The major benefit emanating from this system was that its development allowed a locally calibrated and standardised approach to land capability assessments in South Africa. Further, positives related to the method's relatively simple evaluation procedure as well as its nationwide applicability. These positives are also at the heart of the system's deficiencies. A review by Laker (2004), found the system was too general and does not provide guidelines for different land utilisation types. Further, Schoeman et al. (2002), states that the South African System was never tested or refined and serious inconsistencies remained uncorrected.

2.7.3 KwaZulu-Natal's land capability system

KwaZulu-Natal (KZN) is a diverse Province in terms of soil, terrain and climate (Camp et al., 1995). Consequently, many broad nationalised land classification systems, such as the South African Land Capability, do not always provide consistent results along steep environmental gradients as often encountered in KZN. In order to improve its usefulness and reliability, a more regionalised land capability method was developed and refined for KZN by the Natural Resources Section at Cedara (Camp et al., 1995, 1998; Smith, 2006). Similarly, to the USDA and South African Land Capability Systems, the KZN Land Capability recognises eight land capability classes.

The KZN method stresses the need to acknowledge the role of overriding climatic conditions when assessing the significance of soil characteristics (Smith, 2006). For this reason three separate flow charts or decision trees are presented (an example is provided in Figure 2-1), which cater for high-, medium- and low-rainfall Bioresource Groups within KZN (Camp et al., 1998; Smith, 2006). Class breaks for the various land capability criteria, such as slope and effective depth, differ depending which flow chart is applicable. In lower rainfall areas such as the Dry Zululand Thornveld or Sandy Bushveld natural vegetation cover or crop cover is reduced and therefor criteria for the classification of slope, soil texture and soil depth are more stringently applied (Smith, 2006). The flow charts also have the advantage of being more user friendly and importantly illustrate how important class breaks are in the finally land capability classification. For example, with all other criteria remaining static, a slope change of a single percent, from 2 to 3%, in drier regions will result in a lower land capability class.

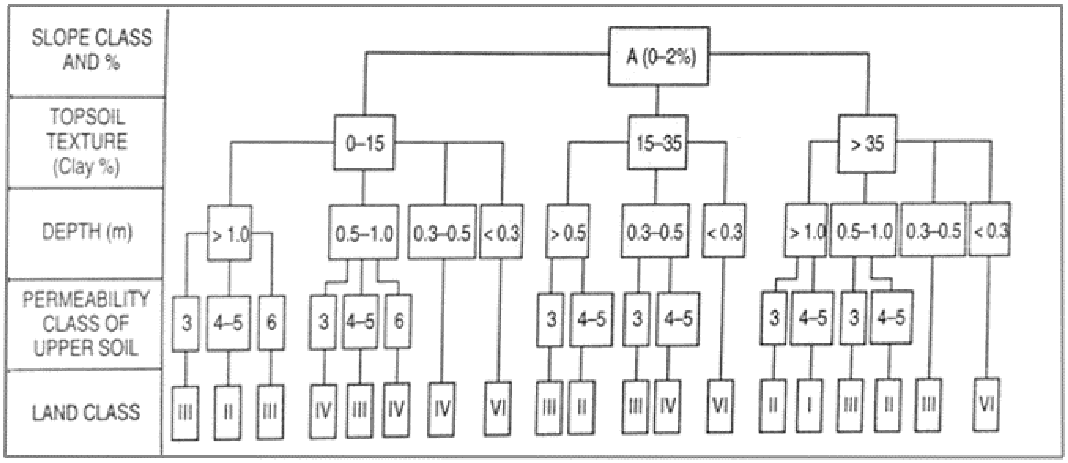

Figure 2-1 Capability class determination guidelines for low rainfall BioResource Groups in KwaZulu-Natal (replicated from Smith, 2006 as adapted from Camp et al. 1998)

The completion of the correct flow chart provides an initial land class, which then needs to be modified based on wetness, permeability, rockiness and soil surface crusting limitations to provide the final land capability class (Camp et al., 1998). Although flow chart selection is based on broad climatic conditions, climate capability downgrades are only considered as part of the land potential classification, this is a significant divergence between the KZN method and the South African Land Capability Classification.

If so required, the land capability class can then be subsequently matched with the overriding climate capability of the study site in order to achieve the land potential classification. The KZN system uses a modified version of the climate capability classes as presented by Scotney et al. (1991), cf Chapter 2.7.2.

Although this regionalised approach shows refinement, compared to the South African System, many faults still exist. The first is the issue of class breaks and classification downgrading severity, which have serious implications to the final classification. For example, an agricultural field in a high rainfall area, with a 4 % slope, can no longer be considered land capability class I. A lack of detailed verification and re-calibration studies, in different parts of the KZN Province, is also considered a fault associated with this method.

2.7.4 Kwazulu-Natal Ecotope Classification

The term ecotope has its roots in landscape ecology and is defined as the smallest ecologically-distinct unit in a landscape, where further subdivision will have no significance (Smith, 2006; Ellis, 2011). In terms of South African agricultural land assessment the ecotope concept has been used in KwaZulu-Natal agricultural planning for over four decades, first being described by Schoeman & Macvicar (1978). Within this agricultural context, an ecotope is described as unit of land, defined in terms of soil functional group, texture, depth, wetness, slope and soil surface characteristics (Camp, 1999). An ecotope is thus associated with low environmental and spatial variation, such that relative uniformity exists in terms of land use options, agricultural yield and production techniques (Smith, 2006).

The ecotope boundaries are identified by using soil, terrain and climatic characteristics and can be refined using aerial imagery (Camp et al., 1998). An *Ecotope Norming Exercise* undertaken by the Natal Region of the Department of Agriculture and Fisheries in (1982). states that an ecotope is a resource classification unit, intended to assist in agricultural decision making including:

- which agricultural enterprises can or should be practised;
- how these enterprises are best and sustainably undertaken; and
- what yields are likely under defined management conditions.

The ecotope planning unit also lies at the heart of The Bioresource Programme of KwaZulu-Natal (BRUP), which is a computerised natural resource inventory tool for the Province (Camp, 1999). The BRUP utilises a code to define soil and crop ecotopes:

Soil Functional Group. Topsoil Clay (%). Effective Soil Depth (mm). Slope (%). Rockiness

A soil ecotope is defined by all 5 characters (e.g. B.1.2.a.0), while a crop ecotope, used to estimate suitability and yields, only requires the first 3 characters (e.g. B.1.2), as shown in Table 2-2. Soil forms as defined in both Binomial and Taxonomic Classification Systems have been converted to soil functional group. Soils within a particular group share similar potential and functionality from an agronomic perspective. For example *Group B: Well and Moderately*

Drained Soils (Ferrasols) consists of *inter alia* Hutton, Clovelly, and Griffin soil forms, which under the same management will produce similar crop yields (Smith, 2006).

Table 2-2 Defining Properties Ecotope Classification (adapted from Camp et al., 1998; Smith, 2006)

Ecotope Classification			
Soil Functional Group	**Code**	**Topsoil Texture**	**Code**
Deep humic soils	A	> 35 % clay	1
Well and Moderately drained Soils	B	15-35 % clay	2
Unconsolidated sediments	C	< 15 % clay	3
Mottled and moderately drained soils	D	**Effective Depth**	
Mottled and poorly drained soils	E	> 800 mm	1
Black (Margalitic) soils	F	500 – 800 mm	2
Black (Margalitic) poorly drained soils	G	300 – 500 mm	3
Young soils	H	200 – 300 mm	4
Gleyed soils	I	< 200 mm	5
Duplex soils	J	**Slope**	
Soft and/or hard carbonates	K	0 – 3	a
Dorbank	L	4 – 12	b
Man-made soils	M	13 – 15	c
Organic soils	O	16 – 20	d
Podzols	Z	21 – 40	e
Rockiness		> 40	x
No mechanical limitations	0		
Many stones but ploughable	1		
Large stones and boulders, unploughable	2		
Very shallow soil on rock or lack of soil	3		

Soil ecotopes are unique so far as they can describe a land unit's status in terms of both capability and suitability. Figure 2-2, schematically shows how a soil ecotope fits within the greater land capability categories. It illustrates that the ecotope is an example of a land capability unit, the finest scale available for land capability evaluations. Ecotope classification is not simply a land capability unit but can also be combined with crop requirements and thus can also be used to define land suitability and potential yields. This duality is a great advantage of this classification method.

This duality is further utilised within the BRUP by combining crop ecotopes, climatic inventories and local crop models, developed by Smith (2006; 1997) to estimate typical yields for a variety of field crops, cultivated pastures and commercial timber plantations. Laker (2004) is of the opinion that although some ecotope studies have been undertaken the concept probably received much less attention than it should have. He further advocates the use of the ecotope to facilitate knowledge transfer at a farm level. To this day farm assessments and associated production recommendations, conducted by the KwaZulu-Natal Department Agricultural and

Rural Development, continue to use soil and crop ecotopes as basis for resource mapping and planning.

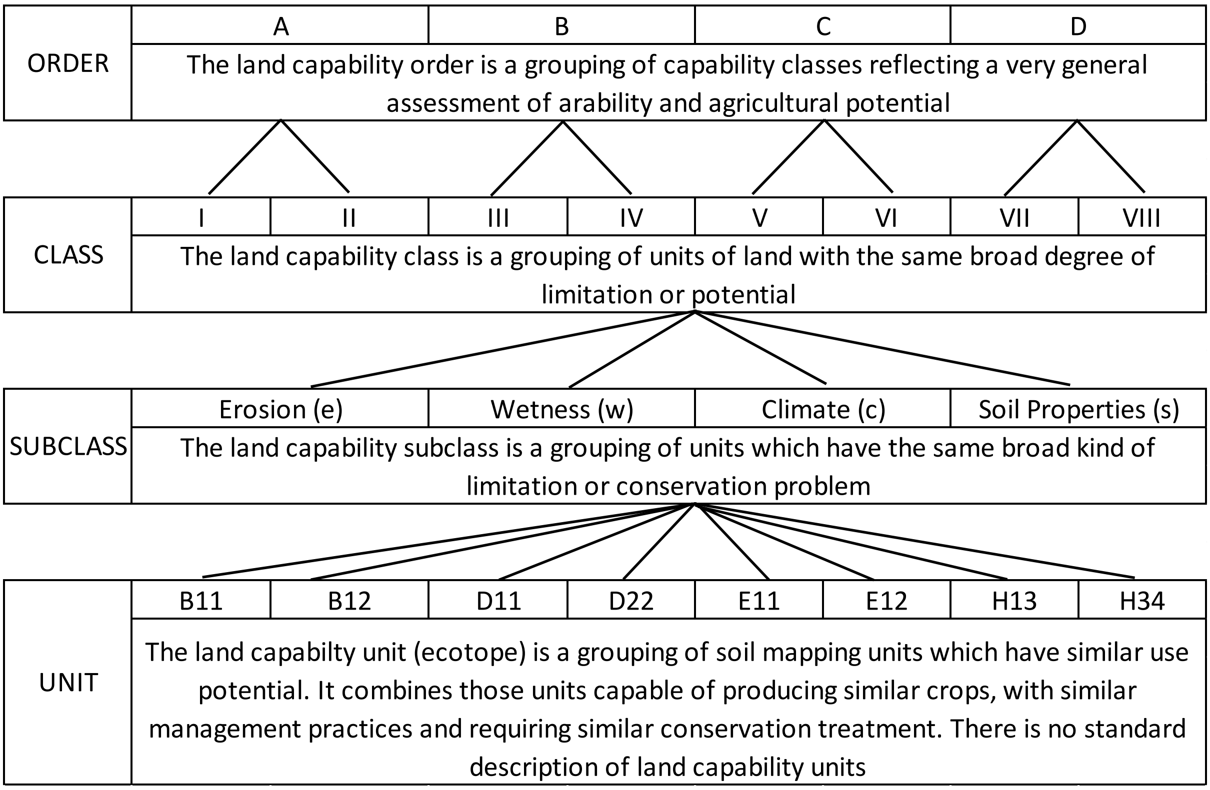

Figure 2-2 A schematic representation of land capability categories (after New Zealand Ministry of Works, 1969; Scotney et al., 1991)

2.7.5 Spatial land capability approaches in South Africa

Advances in Geographic Information Systems (GIS) has correspondingly led to in an increase in spatially-related applications and assessments across all scientific fields, including agriculture and land management (Andreev, 2020). Land evaluation is naturally linked to some form of spatial analysis and thus GIS should play an integral role in methodological development, data management and information processing (Kollias et al., 1999). In South Africa two major projects have been completed with the aim of developing a national spatial land capability product.

The first, undertaken by Schoeman et al. (2002) of the Agricultural Research Council: Institute for Soil, Climate and Water (ARC: ISCW), and commissioned by the National Department of Agriculture, aimed to develop a workable land capability system for South Africa. The primary objectives of the project were to develop, test and document mental models for deriving and presenting land capability and translate these mental models into algorithms (Schoeman et al., 2002). This work culminated in a report and seamless digital land capability spatial layer for South Africa. The final product is a classical eight-class land capability evaluation system that is conservation-orientated and utilizes existing broad national soil, terrain and climate databases (Laker, 2004). According to DAFF (2018) the products generated by the Schoeman et al. (2002) can only be used at scale of 1:250 000 and is not deemed suitable for detailed delineation nor for assisting in decision making relating to capability or potential of agricultural resources. Laker (2004) had earlier come to a similar conclusion, that uncertainties relating to the input data and class determination restricts the ultimate use of the product to very general, land capability assessments at regional or national scale only.

The limitations associated with the work of Schoeman et al. (2002) led the National Department of Agriculture’s Directorate of Land Use and Soil Management (DLUSM), to embark on a new process to refine this initial spatial land capability layer for South Africa. The *Land Capability Evaluation and Classification for South Africa Project* was initiated in 2014, with the initial spatial products being released, two years later, in 2016. This refinement aimed to allow for improved decision-making pertaining to the identification and preservation of high potential agricultural land (DAFF, 2018). Its spatial methodology made use of a three-tier data architecture to produce an agriculturally driven land capability evaluation model comprising of 15 land capability classes at usable scale of between 1:50 000 and 1:100 000 (Collett, 2019). This 15-class system is a departure from the classical eight land capability classes (Klingebiel & Montgomery, 1961) and has an inverted scoring system where class 1 is the lowest and poorest capability class and class 15 is the highest or best possible value. The system generally uses land capability class 7 as the break between arable and non-arable land uses but this can differ, based on the user’s local knowledge (Collett, 2021). According to the user documents provided DAFF (2018) the refined spatial product utilises a complex weighting system of soil, terrain and climatic variables (Figure 2-3). Soil related variables contribute 30% towards the final land capability classification and were mostly extracted from modal profiles collected during the land-type survey. Climate capability factors contribute 40% of the final weighting and are based on data extracted from the South African Atlas of Climatology and Agrohydrology (Schulze, 1997). The terrain capability contributed 30% of the final weighting and terrain related attributes were generated from the 90 m SRTM (Shuttle Radar Topography Mission) digital elevation model.

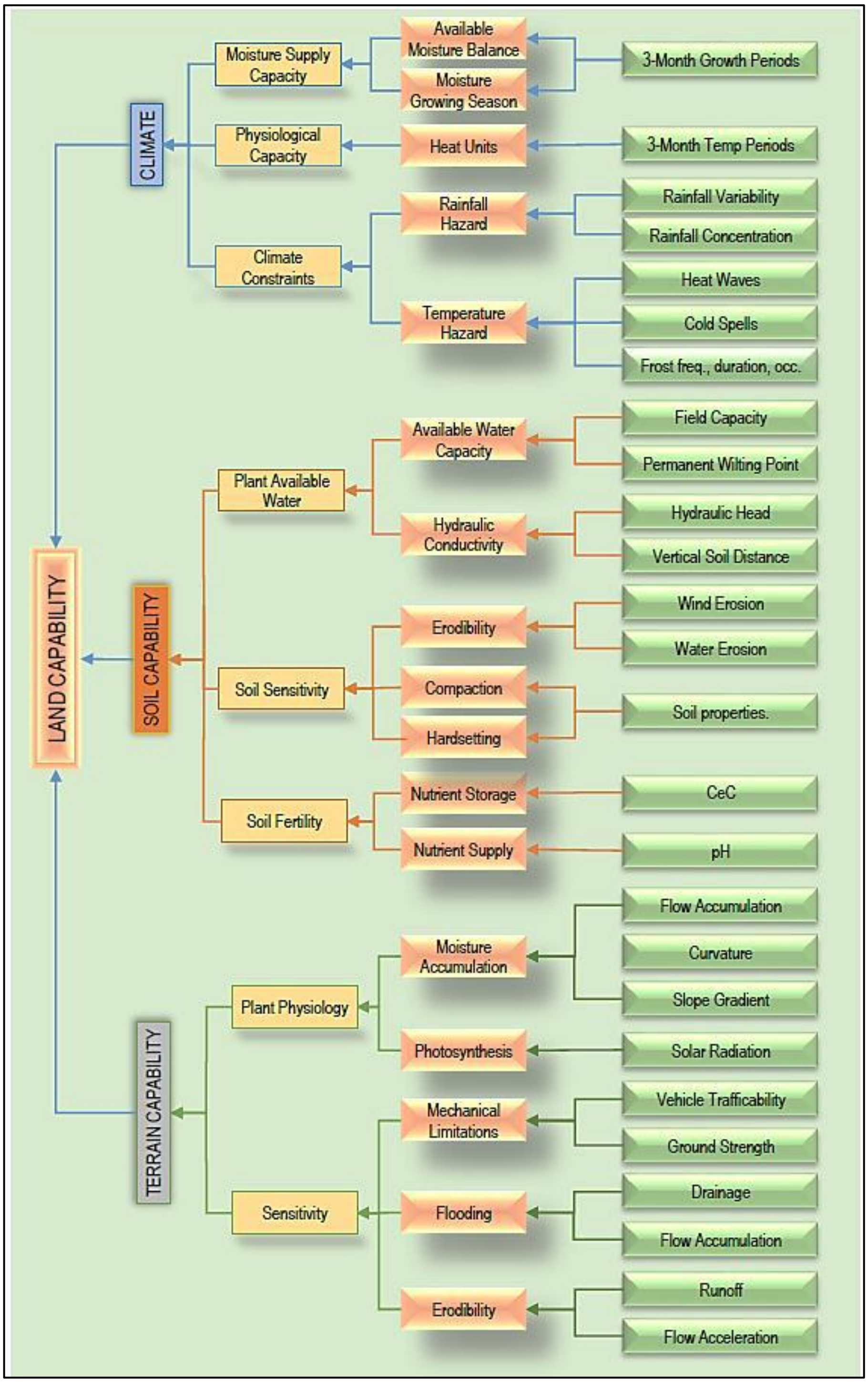

Figure 2-3 Land Capability Evaluation Schematic Model (DAFF, 2018)

The resulting spatial product is user-friendly and does provide national, agriculturally related data in a single repository. However, there are limitations associated with the weighted methodology utilised in this system. The country of South Africa is both complex and diverse in terms natural resource potential (DAFF, 2018) and thus a weighting system, which attempts to encapsulate this diversity in its entirety, often leads to important or “unique” agricultural land, on the fringe, of the weighting spectrum being misclassified. This is highlighted in wheat growing areas of the Western Cape, where actively cultivated land is often classified as non-arable due to the winter rainfall regime. Similar disjuncts were found in Northern KwaZulu-Natal where active cropland was classified in very low, non-arable, capability classes. A regionalised weighting-system could potentially overcome these limitations.

The premise behind the 15-class system was to provide more detail for planners and allow for improved decision-making. However, even post-hoc analysis performed by the National Department, aggregates the 15 classes into a more manageable 9-class system (e.g. (Collett, 2019), this secondary aggregation is indicative of redundant classes within the original dataset.

The summarised results provided in the DAFF Report (2018) indicates that of the 69.89% of the KwaZulu-Natal Province falls in Land Capability Class 7 or above, essentially this model predicates that 70% of the Province is potentially arable. This is considerably higher than 41.80% estimated by Schoeman et al. (2002) as well as other regional estimates, where only 16% of land is considered suitable for annual cultivation and an additional 8% for permanent crops (KZNDARD, 2018). The broad scale of the input data, particularly in terms the soil and terrain co-variants could also be seen as a potential source of arability overestimation, particularly at local level. Further, due to the intricate methodology employed, the ultimate spatial product and its associated weighting system cannot to be adapted or replicated by other developers or users, when more detailed or new soil information becomes available.

2.8 Do South African Land Assessment Methods Need to be Reviewed and Updated?

In recent decades, scientific innovation and research have increased our understanding of agricultural systems, their components and their management (Jung et al., 2021). Coupled with this increased understanding is the great advancement in technology used in agricultural and natural resource management such as precision agriculture (e.g. Cisternas et al., 2020), remote sensing (e.g. Huang et al., 2018), GIS (e.g. Ustaoglu et al., 2021) as well as computer aided modelling and machine learning (e.g. Jagtap et al., 2021). These advances should be incorporated into methodologies to ultimately improve land classification and evaluation.

Section 2.7.5 highlights the South African spatial products, which have recently been developed using newer technologies, but this has not translated into revising or developing new in-field land evaluation methodologies, particularly at farm level. This is critical as only farm level, land evaluation assessments are used by the National Department of Agriculture to determine if land should remain under the auspices of agriculture or be released to a non-productive land use (DAFF, 2018).

The recent release of the third South African soil classification system in 2018 (SCWG, 2018), was seen as an ideal opportunity to update antiquated land assessment methodologies, as many use soil classification as a raw input into their models (e.g. land capability and ecotope classification). A recent review by van Zijl et al. (2020), which this study contributed towards, investigated the implications of the new soil classification system on users and related methodologies. This research found that existing agricultural land classification methods will require significant revision to appropriately align with the new soil classification framework, which has now been endorsed as the new, official classification system for South Africa.

Along with the aforementioned technological improvements and revisions to the soil classification system, proposed legislation is also contributing to the need to update land evaluation assessment methodologies and frameworks. PDALB, (cf Section 2.5.5), explicitly refers to the establishment of a system of land capability classification, within an appropriate land evaluation framework, for determining the physical capability of land at national, regional and local scale. Current methodologies would need to be revised and updated in order to meet these legislative requirements.

2.9 Conclusions

In this Chapter, the primary concepts relating to agricultural land assessment and evaluation were introduced. To effectively explore these concepts seven key questions were developed and the following conclusions drawn:

Agricultural land evaluation is a critical process in land use management and when implemented effectively can improve decision-making, optimise land use, reduce environmental degradation and improve productivity.

South Africa has a number of sound legislative policies and Acts, which aim to promote the sustainable use of agricultural resources. However, significant legislative overlap and paralysis exists, which has triggered an uncoordinated and inconsistent approach to decision making across various governmental departments. These administrative and legislative failures are leading to a loss of critical agricultural land and the degradation of the resource base.

South Africa is characterised by natural resource diversity. Consequently, no single or universal method should be relied upon to evaluate all possible scenarios emanating from agriculturally based assessment and land use planning. Although there is a wide range of both international and locally calibrated evaluation methods available to land assessment practitioners, there appears to be an overreliance on land capability-based methodologies in South Africa.

There is a need for local verification studies, to analyse the performance of land assessment methodologies currently been practiced in industry. Additionally, local assessment methodologies, particularly at farm level, are in need of revision to incorporate recent pedological revisions, legislative requirements and address the current challenges facing both land use planners and agricultural scientists. These methodologies should, where applicable, incorporate newer technologies such GIS, precision agriculture, terrain analysis and machine learning.

3. VERIFICATION OF LAND ASSESSMENT POLYGONS IN A PRODUCTION ENVIRONMENT

3.1 Introduction

Recent legislative policy, proposed through the Preservation and Development of Agricultural Land Bill (PDALB), aims to establish a broad framework to classify rainfed agricultural land according to the most intensive long-term use thereof, determined by the interaction of climate, soil and terrain (DAFF, 2016). This proposed framework intends to include methods for determining physical capability at national, regional and local scales. Presently the South African Land Capability System (Scotney et al., 1991), produced to standardise the approach to evaluating agricultural land in South Africa, remains the default methodological approach. However, its limitations in terms of applicability, methodological inaccuracies and lack of in-field testing have attracted criticisms (e.g. Schoeman et al., 2002; Laker, 2004).

More recently alternative land assessment methods developed, both locally and abroad, are gaining greater attention and implementation (e.g. Shepherd et al., 2008 DAFF, 2018a). There is, however, an absence of local, in-field verification studies, for these assessment methods in order to determine if they are performing adequately, particularly in a production environment. Indeed, one of the major shortcomings of local land assessment methods, as identified by both Laker (2004) and Schoeman et al. (2002), is the lack of verification studies.

The aims of this Chapter are therefore to determine if soil and land assessment techniques, currently being practised in industry, reflect actual land utilisation and production levels. To achieve this, five land assessment methods were selected:

1. A System of Soil and Land Capability Classification for Agriculture in South Africa (Scotney et al., 1991), hereafter referred to as “RSA LC”;
2. The KwaZulu-Natal Land Capability System (Camp et al., 1998; Smith, 2006),”, hereafter referred to as “KZN LC”;
3. KwaZulu-Natal Ecotope Classification (Schoeman & Macvicar, 1978; Camp et al., 1998) hereafter referred to as “Ecotope”;
4. Soil Visual Assessment (VSA) for Maize (Shepherd, 2010) and Annual Field Crops (Shepherd et al., 2008) ; and
5. National Department of Agriculture Forestry and Fisheries National Capability Digital Product (DAFF, 2018a), hereafter referred to “DAFF LC”.

Importantly, all the selected methods can be applied "*out-the-box*" without the need for local calibration or modification. Due to their distribution, production outputs and key contributions to food security (Statistics South Africa, 2020), Maize (*Zea mays*) and Soybean (*Glycine max*) were selected as the focus crops of this study. Further, only dryland fields formed part of this analysis, irrigated lands were excluded as they have their own system of assessment and evaluation. Finally, the survey methodology, observation density and data collected follow the prescribed standards as outlined by the National Department of Agriculture (DAFF, 2018b).

The performance of each method is assessed through two procedures. First, the arability results, from the five land assessment methods, are compared to the distribution of actual land use. Second, the final land assessment classification polygons are compared to actual productivity, using precision yields across five growing seasons, between 2015 and 2020.

3.2 Materials and Methods

3.2.1 Study area

The study area is located near the town of Bergville in the northern region of the Tugela Catchment, KwaZulu-Natal (Figure 3-1 a-c). The study area is positioned between Woodstock Dam, forming its western boundary, and the R74 road to the east.

The area extends from 28^{o} 38' 5.8" S; 29^{o} 07' 45.2" E to 28^{o} 41' 14.7" S; 29^{o} 12' 56.5" E and covers some 1 956 ha. The surveyed farming enterprise, known as FCL Farming, is a typical mixed commercial farming enterprise for the region. The study area encompasses land from 16 farm portions and in terms of land use combines dryland and irrigated cultivation of maize and soybean, on a three-year rotation, as well as grazing of livestock on both natural veld and improved pastures (Figure 3-2).

Approximately 75% of the survey area is under currently or previously cultivated land or pastures, the balance is predominately natural veld and wetland areas. The natural vegetation pattern is classified as Moist Transitional Tall Grassveld (Camp, 1999).

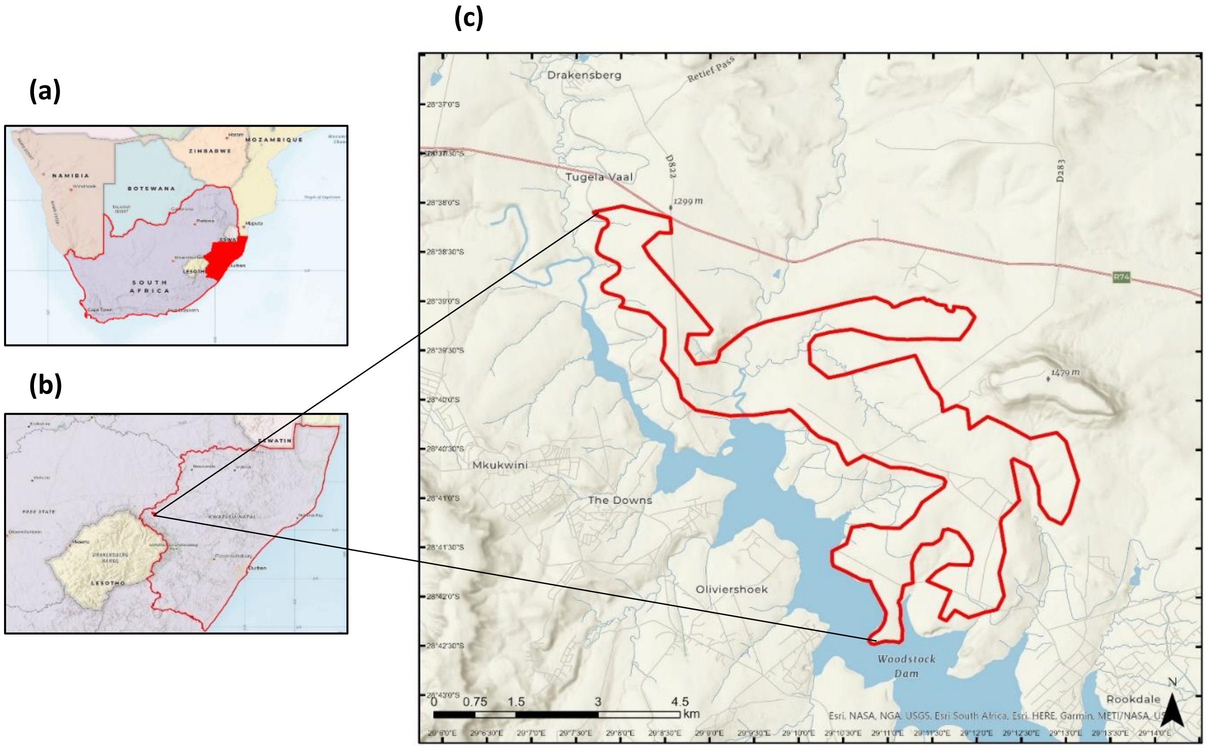

Figure 3-1 Location of the Study area – (a) Location within southern Africa (b) Location within the Province of KwaZulu-Natal (c) Regional Locality Map (Background Layers provided by ESRI, 2021)

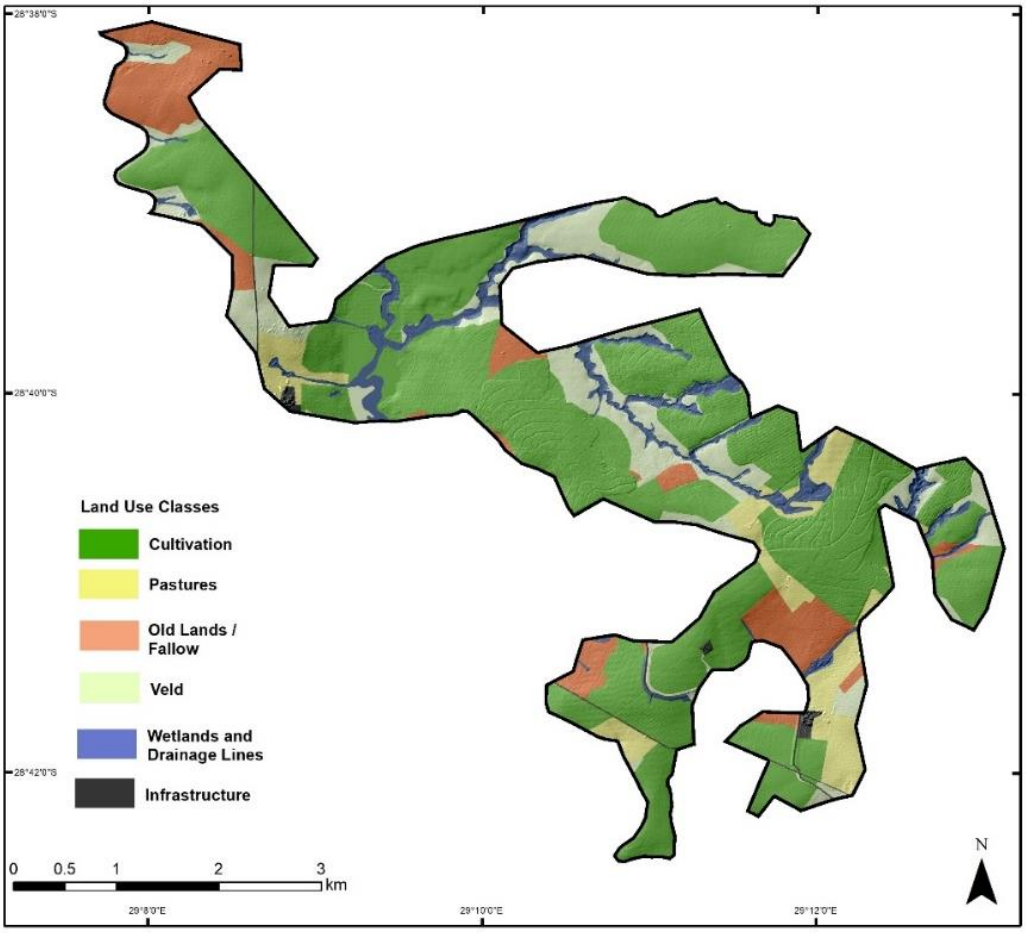

Figure 3-2 Land Use Map

The survey area falls entirely within BioResource Unit (BRU) WXc5 (Woodstock) and includes land types Ab208 and Db261 (Camp, 1999; Land Type Survey Staff, 1976-2006). “WX” in the Bioresource code denotes a mean annual rainfall of between 801 and 900 mm, while the “c” denotes an upland area with an altitude of between 901 and 1400 m. In terms of climate, Camp (1999) classifies this area as having a *Class C5* within the Province of KwaZulu-Natal. This Provincial classification is based on the ratio of annual precipitation to annual A-pan evaporation in combination with mean September, June and annual temperatures (Smith, 2006). Within the provincial context, a C5 climate rating, is described as climate that has a restricted growing season due to low temperatures, frost and/or moisture stress (Camp, 1999). The area has a summer rainfall regime (Table 3-1) with 83% of Mean Annual Precipitation falling between October and March. Suitable, adapted winter crops can only be grown through the application of supplementary irrigation.

Mean daily temperatures range from 23 °C in summer to 12 °C in winter (Figure 3-3). On average, 15 days of heavy frost is expected during early autumn and late winter (Camp, 1999). More spatially relevant rainfall was obtained from on-site rain gauges in order to link actual seasonal rainfall variation to crop productivity (cf Chapter 3.5).

Table 3-1 Climate Summary for the Study Area (Camp, 1995 and Schulze, 1997)

	Unit	Annual	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Median Rainfall	mm	696	131	123	100	36	11	3	2	11	23	61	84	111
Mean Rainfall	mm	824	162	139	98	40	13	9	7	23	49	68	92	124
Mean Average Daily Temp	°C	18.4	22.7	23.1	21.5	18.3	14.8	12.3	12.5	14.6	17.8	19.6	21.3	22.5
Mean Minimum Daily Temp	°C	9.4	15.1	15.4	13.7	9.7	4.9	1.6	1.9	4.6	8.1	10.9	12.9	14.3
Mean Maximum Daily Temp	°C	27.4	30.3	30.9	29.4	27	24.7	23	23.1	24.6	27.5	28.3	29.8	30.7
Relative Daily Humidity	%	56	65	61	66	62	59	41	44	47	49	57	57	62
A-Pan Evaporation	mm	1900	216	179	163	125	105	92	103	138	171	190	196	222

The study area is underlain by a mix of geological materials of the Karoo Super Group. The parent material is mainly shale, siltstone and sandstone of the Estcourt Formation (Beaufort Group); sandstone; maroon, green and blue mudstone of the Tarkastad Formation, alluvium and small areas of dolerite (ENPAT, 2002).

In terms of broad soil patterns, the area can be characterised by having both upland and lowland duplex soils e.g. Valsrivier and Sepane Forms (Solonetz and Luvisols) within a plinthic catena ranging from freely drained red and yellow apedal soils (Acrisols and Ferralsols) in the mid-slopes grading into grey plinthic soils (Plinthosols) on the foot and toe slopes.

The elevation ranges from 1175 to 1300 m above sea level with rolling terrain (Figure 3-3 a). The major topographic feature in the region is the large mesa landform, located to the northeast of the study area, with its southerly slopes falling within the study area (Figure 3-3 b). The study area generally slopes in a south-westerly direction towards Woodstock Dam and is characterised by incised drainage channels and channelled valley-bottom wetland systems, which feed directly into Woodstock Dam.

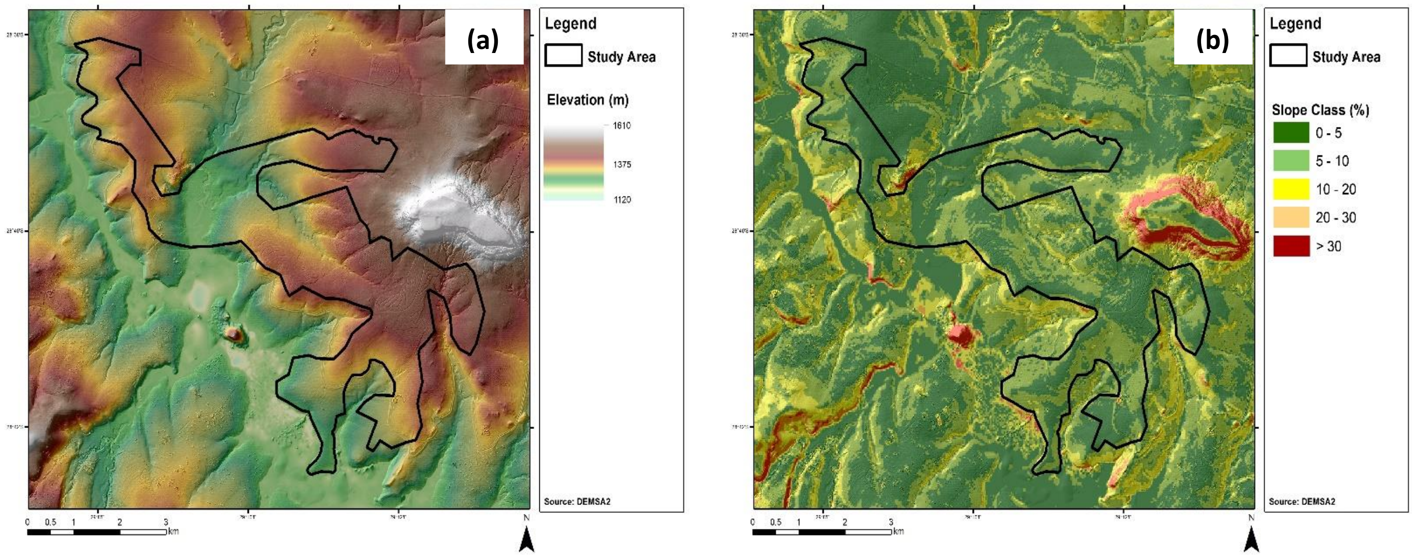

Figure 3-3 Topographic Features – (a) Elevation Map (b) Slope Class Map (Developed from DEMSA2*)*

3.2.3 Soil and land assessment surveys

Land assessment surveys and associated soil sampling took place predominantly over the winter of 2016. Smaller, supplementary surveys were conducted between 2017 and 2019. During these surveys, a total 225 soil observation points were collected, equating to a sample density of one observation per 8.6 ha (Figure 3-4). Most observations were conducted using soil pits dug to at least 1.5 meters or refusal, while small proportion were collected by confirmatory auger holes. All observation points were classified using the *Soil Classification, A Taxonomic System for South Africa* (SCWG, 1991), however all horizons were fully described to convert this Taxonomic Classification to the new South African Classification (SCWG, 2018), which was not formally recognised at the initiation of this project.

Sampling positions and densities were based on norms and standards produced by both National and Provincial Agricultural Departments for agricultural land assessment (DAFF, 2018b). It is a purposive sampling approach using expert knowledge, current land use, slope and topographic positions. At each sampling point the following information was collected:

- **<u>General</u>:** Spatial position, land use, crop type, geology, soil sample reference numbers and general comments;
- **<u>Terrain:</u>** Terrain position, unit and slope class (via Abney level);
- **<u>Soil Classification:</u>** Horizon name, horizon thickness and colour, soil form and family; and
- **<u>Land Assessment Attribute Data:</u>** Total soil depth, effective depth, topsoil clay content, permeability of B1 Horizon, soil structure type and grade, wetness classification, soil crusting, rockiness, soil erosion type and severity as well as method specific land assessment attributes.

This information was collected using handheld *Trimble GeoXT* GPS Units with on-board Terrasync Software (*www.trimble.com*), to record positions and the associated attribute data using customised data dictionaries. All recorded positions were downloaded, differentially corrected, and exported as shapefiles using Trimble GPS Pathfinder Office software.

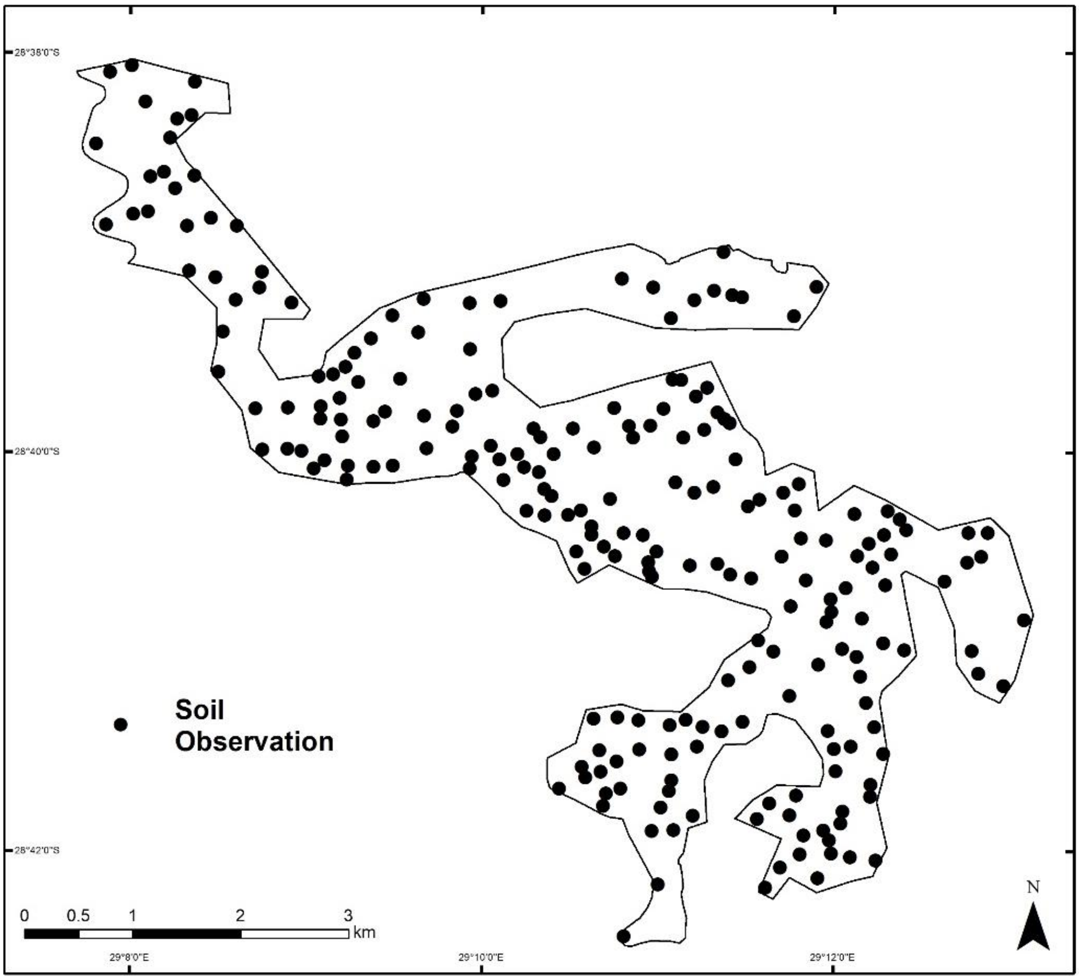

Figure 3-4 Soil observation points across the study area.

3.2.3 Land assessment methods and polygon generation

This polygon-based study analyses the final classification of each land assessment method, for example land capability class for RSA LC or aggregated soil quality index in the case of VSA. For the RSA LC, KZN LC, Ecotope Classification and both VSA methods, all soil observation points were reclassified, based on each method's unique land assessment criteria. An example of this secondary classification process for KZN LC and Ecotope Crop Classification is provided in Figure 3-5.

RSA LC classes were determined using the standardised methodology, as outlined in the systems manual (Scotney et al., 1991) where land is grouped into one of eight classes with the same broad degree of limitation or potential. The determined overarching National climate capability, Class III, was included in order to reach a final land capability classification for this National method, taking into consideration climate is relevant in terms of LC classification in this case, as it assesses rainfed crop production where the climatic potential is restrictive, as recommended by Scotney et al. (1991).

The KZN LC classes were determined using the flow chart provided in Smith (2006) as adapted from Camp et al. (1998). The study area falls within a "moist" climatic Bioresource Group and thus this specific flow chart was used to determine both land class and land capability. In the KZN LC method only broad climatic conditions are considered in terms of land capability classification (cf Chapter 2.7.3). Initial land classes were adjusted based on permeability, wetness, rockiness and surface crusting to achieve the final KZN LC classification. KZN crop ecotopes were determined using soil functional group, topsoil texture and effective rooting depth as outlined in Camp et al. (1998) and Smith (2006).

Individual VSA scores and overarching soil quality index were determined for maize production using the maize specific scoring and weighting system, as outlined in Shepherd (2010). The more generic annual field crops scoring and weighting system was used in respect to soybean production (Shepherd et al., 2008). Individual visual soil quality indicators were scored, ranked and aggregated to determine the soil quality index and assessment class for both maize and soybean production.

The overarching "Land Capability" layer, determined from soil, terrain and climate attributes as provided in the DAFF digital mapping product was extracted for the study area (DAFF, 2018b). This extracted raster dataset was subsequently simplified to land assessment polygons.

Once a final classification was determined and mapped for all survey points and for each applicable method, the points were upscaled to land assessment polygons as per the survey norms and standard as recommended by DAFF (2018b). Polygon delineation was based either on soil boundaries and/or topographic breaks, depending on which method was being applied.

For all spatial related analyses ArcGIS 10.5 (ESRI, 2016) was employed and this included *inter alia,* polygon generation, geoprocessing, delineation classification and mapping.

Slope	7%
Soil Form and Family	Avalon 1100
Soil Functional Group	D
Effective Depth	900 mm
Topsoil Clay Content	28%
Permeability Class	4
Crusting	Moderate
Rockiness	Zero
Wetness Class	W1

Ecotope	D.2.1

KZN LC Class	II

(1) Soil Observation Point → **(2) Soil and Land Attribute Data** → **(3) Land Assessment Classification**

Figure 3-5 An example of the Secondary Land Assessment Classification Process for KwaZulu-Natal Ecotope and Land Capability Classification (Soil Profile Photo: SCWG, 1991)

3.2.4 Arability analysis and agreement analysis

The polygon maps, generated from the five land assessment methods, were simplified into two classes, namely arable and non-arable. Ecotope and VSA methods were not specifically designed to classify arability, thus classification breaks were developed for this purpose (Table 3-2). Existing agriculture infrastructure, placed on potentially arable land, such as housing or sheds, did not contribute to non-arable class.

Table 3-2 Arability classification breaks for each land assessment method

Method	Arable Classes	Non-Arable Classes
RSA Land Capability	I - IV	V - VIII
KZN Land Capability	I - IV	V - VIII
Ecotope Classification	Crop Ecotopes suitable for dryland maize, soybean or improved pasture	Crop Ecotopes unsuitable for dryland maize, soybean or improved pasture
DAFF Land Capability	7 - 11	1 - 6
VSA Maize and Annual Field Crops	“Medium” and “High” Aggregated Soil Quality Index	“Low” Aggregated Soil Quality Index

To more accurately compare modelled arability to land utilisation, all polygon maps were converted to a 50 m grid, where the majority land use or modelled arability was used to assign its classification. These standardised grids were then used as part of a confusion matrix agreement analysis between actual land use and modelled arability as produced by the five assessment methods.

A confusion matrix is a machine learning and statistical classification method whose aim is to visualise the performance of a supervised learning algorithm (Brownlee, 2016). Each confusion matrix was identically constructed, resulting in four classes where the Top Left value is *True Positive*, Bottom Right is *True Negative*, Top Right is *False Positive* and the Bottom Left is *False* Negative (Table 3-3). First, the resulting confusion matrixes were used to spatially indicate where modelled arability differed to actual land use across the study area. Second a statistical analysis was conducted using a set of confusion matrix indices, including Classification Accuracy, Misclassification Rate, Classification Precision, Classification Sensitivity and the Matthews Correlation Coefficient (MCC).

Table 3-3 Confusion Matrix structure used in land arability classification

Actual / Predicted	Actual Positive	Actual Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

The compilation of the various confusion matrixes allows for additional statistical metrics to be extracted. The first is Classification Accuracy (CA) which is defined as

$$CA = \frac{(TP+TN)}{n} \tag{Eq. 3-1}$$

Where TP is True Positive and TN is True Negative and n is the total count across all classes.

The inverse of Classification Accuracy is Misclassification Rate (MR) or False Positive Rate which is defined as

$$MR = \frac{(FP+FN)}{n} \tag{Eq. 3-2}$$

Where FP is False Positive and FN is False Negative and n is the total count across all classes. These two statistics provide a measure of classification accuracy and error.

Classification precision (Eq 3-3) indicates what proportion of arable predications were in fact correct, a land assessment method that produces no False Positives has a Precision score of 1.0.

$$Precision = \frac{(TP)}{(TP+FP)} \tag{Eq. 3-3}$$

Classification sensitivity, also known as recall (Eq 3-4) is the probability that a land assessment method correctly predicts an actual true value. In this case when a pixel was cultivated did the land assessment model predict that it was arable.

$$Sensitivity = \frac{(TP)}{(TP+FN)} \tag{Eq. 3-4}$$

An overarching performance coefficient was calculated from the six Confusion Matrixes. Matthews Correlation Coefficient (MCC) (Eq. 3-5) incorporates all the individual matrix values to statically evaluate how well the classification performed as compared to a randomly generated classification.

$$Matthews\ Correlation\ Coefficient\ (MCC) = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{Eq. 3-5}$$

3.2.5 Precision yield data collection and processing

Seasonal precision yield data was collected by a John Deere Combine Harvester mounted with a continuous precision yield monitor (GS23:2630). A review by Lyle et al. (2014) suggests studies on continuous yield monitors report an accuracy of between 93 and 99.5% but are dependent on amongst other factors monitor brand, calibration and environmental conditions. The harvester offers onboard automatic calibration but is manually calibrated through the use of a weigh wagon. Mass point precision yield data was extracted, processed and exported using a combination of John Deere Apex Software 3.85 and more recently its successor John Deere Operations Centre (*www.deere.com*). Yield data was subsequently exported to ArcGIS for further processing and cleaning.

Commercial yield mapping systems have been in operation since 1992 and all require some level of post-processing to remove data artefacts and reduce error (Blackmore & Moore, 1999). Post processing of exported yield data included data screening and removal (e.g. Simbahan & Dobermann, 2004; Sun et al., 2013) which included removing unreasonable and distribution outliers and positional errors. Data removed were generally near major field edges, erroneous points caused by harvester re-routing, duplicate positions or points collected when the yield monitor was off.

Five years of yield data for growing seasons 2015-16, 2016-17, 2017-18, 2018-19 and 2019-20, for both dryland maize and soybean, were collected and processed to dry yield mass per hectare. Commercial farms in the region generally use a three-year rotation for maize and soybean, with two maize cycles followed by a soybean crop. This ratio is however adaptable, depending on overriding grain prices and climate forecasts. Other crops or areas under irrigation were not included in this assessment. For this study, yield results were analysed across the five years seasons as well as on an individual season-by-season basis to determine the impact of rainfall on yield variation across soil groups and terrain units. On average 680 000 individual cleaned yield points for maize and soybean were analysed every season. These yield points were then spatially joined to the pertinent land assessment polygon classifications.

3.2.6 Statistical software and methods

A combination of Microsoft Excel 365, Microsoft Access 365, IBM SPSS (IBM, 2021) and Statisica (TIBCO Software Inc., 2018) were used to manage and statistically analyse the large precision yield datasets and land assessment outputs.

To determine if any statistically significant differences exist between land evaluation classes and crop yield, the data was subjected to several statistical analyses. First, the test of homogeneity of variances was used to determine if the level of variance for crop yield is constant across the land evaluation classes of a specific method. If the assumption of homogeneity was not violated then a one-way ANOVA, the Analysis of Variance, was employed at a significance level of 95% ($\alpha = 0.05$). The ANOVA determines if there is significant yield difference across the various classes and if this difference exists then either a Tukey HSD or Tamhane (for unequal variances) post hoc test, for multiple comparisons, was employed to identify exactly which classes differed significantly from each other ($p=<0.05$).

If the test of homogeneity of variance was violated then non-parametric tests, including the Independent Median and Kruskal-Wallis Tests, were employed to determine any significant yield differences. Yield differences for individual classes were further explored by employing Dunn's post hoc pairwise comparison using for multiple tests ($\alpha = 0.05$). The resulting pairwise significance was not adjusted using the Bonferroni adjustments, as this study did not meet the specific conditions where adjustments should be considered, as reported by Armstrong (2014). In his review, Armstrong (2014). states that Bonferroni adjustments should only be considered in very particular situations if:
"(1) a single test of the 'universal null hypothesis' (H_o) that all tests are not significant is required, (2) it is imperative to avoid a type I error, and (3) a large number of tests are carried out without preplanned hypotheses". This is supported by earlier work undertaken by (Perneger, 1998) who concluded that the persistent use of Bonferroni adjustments "*are at best, unnecessary and, at worst, deleterious to sound statistical inference*". Consequently, adjusted significance values were avoided during the pairwise comparison for non-parametric tests.

3.3 Results and Discussion

The results for this Chapter are divided into three sections. The first investigates the generation of land assessment polygons, for each of the five assessment methods. The second examines the relationship between actual land utilisation and modelled land use, through an arability and agreement analysis. The final analysis compares land assessment polygons to dryland maize and soybean yield, to assess the relationship between land assessment classification and actual productivity.

3.3.1 Land assessment polygon maps

Six land assessment polygon maps were generated from the five land assessment methods (Figure 3.6a-f). The result from the RSA LC classification produced a total of four classes (Figure 3-6a), with two arable classes identified. The majority of the study area falls within Land Capability Class III and covers 62% of the surveyed area.

Class III land is considered to be of moderate potential with severe permanent limitations that restrict land use and intensity of crop production (Scotney et al., 1991). The consolidation of a wide variety of soil forms across multiple terrain units, into two arable classes, highlights the significant aggregation of this classification method, where the most limiting land capability factor between soils, terrain and climate is used to determine the final land capability class. Although this farm is a successful commercial operation, no areas were classified as Land Capability Class I or II, which are considered to be potentially, highly productive. This is primarily due to inherent partiality of the system towards soil conservation rather than production potential (Laker, 2004; DAFF, 2018a).

The KZN LC Classification produced a total of six classes (Figure 3-6b). Areas with more favourable soil and terrain properties for crop production correspond to Class II, which covers 52% of the study area. Similarly, to the National system the regionalised KZN method does not classify any areas as Land Capability Class I. Wetland areas are divided into Classes V(a) and V(b), which provides an improved delineation between seasonal and permanently wet areas. Classes III and IV, are generally located along the foot slope of the mesa landform, in the northeast, as well as in the lower lying areas central areas.

Compared to the other methods, the Ecotope Classification produced the most classes, with 29 unique crop ecotope classes (Figure 3-6c) based on a combination of soil functional group, topsoil clay and effective depth. The ecotope classification is generally recognised as the land capability unit classification, which is the finest realistic classification resolution possible (New Zealand Ministry of Works, 1969; Camp et al., 1995). The resulting map (Figure 3-6c) is colour coded to match with the various soil functional groups. For example, the differing shades of green correspond to the well-drained soils with differing clay contents and effective depths.

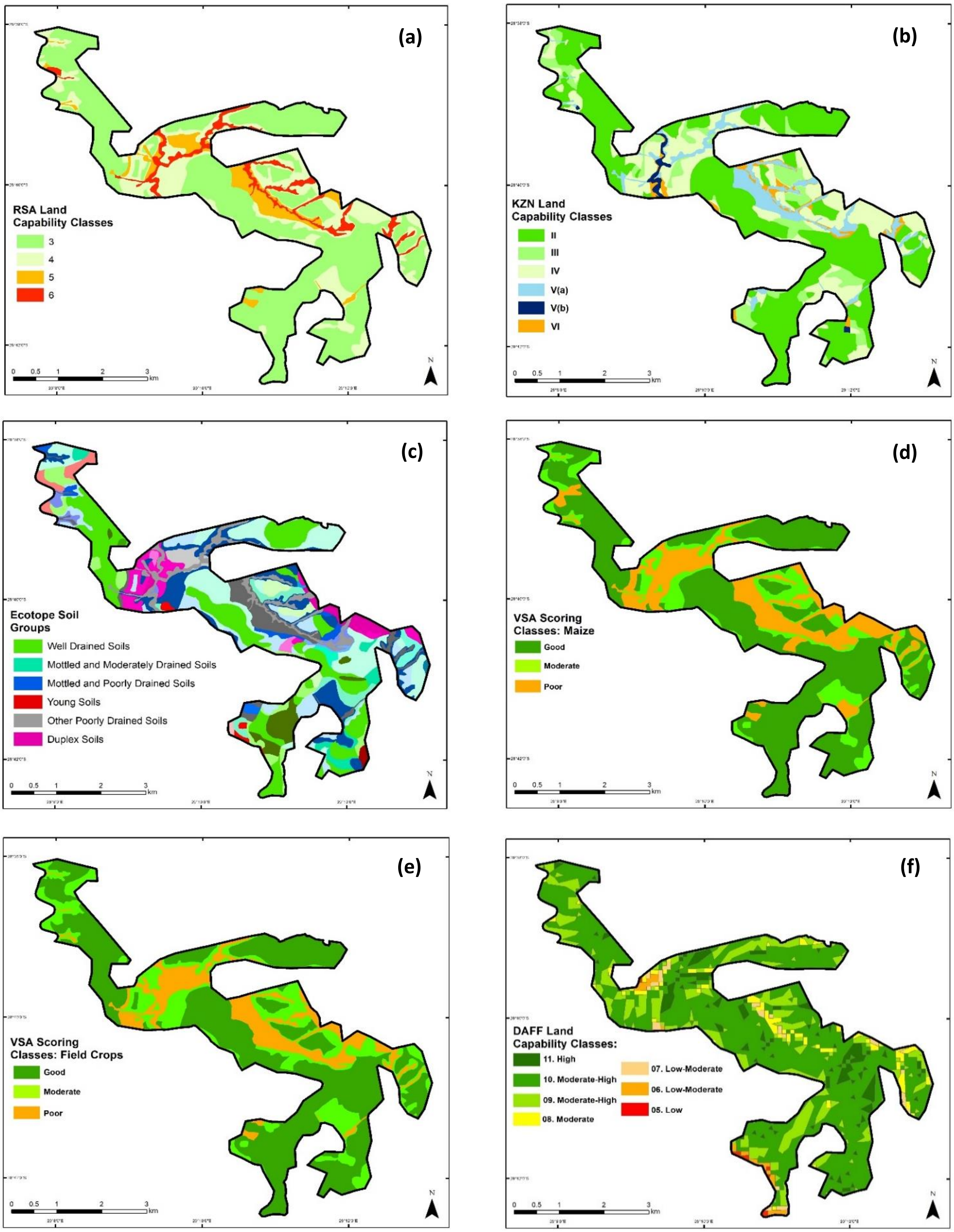

Figure 3-6 Land assessment polygon classification maps – (a) RSA Land Capability Classification (b) KZN Land Capability Classification (c) KZN Ecotope Classification colouration based on broad functional grouping (d) VSA Maize Classification (e) VSA Annual Field Crops Classification (f) DAFF Digital Land Capability.

Generally deep, well-drained soils are located along gentle sloping, upland areas (Figure 3-6c). Conversely, mottled and poorly drained soils, shown in the various shades of blue, correspond to area associated with areas with a fluctuating water table and wetness limitations. Duplex soil groups (Solonetz and Luvisols) are located in both upland and lowland areas, with the majority associated with effective soil depth between 300 and 500 mm. The identified soil and ecotope patterns correspond well to the descriptions provided in the broader BRUP inventories (Camp, 1999).

Both VSA methods (Figure 3-6d and e) provide analogous results, where most of the surveyed area was classified as “Good” from a soil quality perspective. VSA for Maize classifies 60% of the study area as “Good” for maize production, 18% is considered “Moderate” and the remaining 22% is seen as “Poor”. Similarly, VSA for Annual Field Crops classifies 58% of the study area “Good” for soybean production, 25% is considered “Moderate” and the remaining 17% is seen as “Poor”. This spatial similarity is somewhat expected as there is a significant overlap between the soil quality indicators used for each method (Shepard *et al*., 2008; Shepard, 2010). Generally, apedal soils, located on mid-slope, were classified as “Good” for both maize and annual field crops.

The DAFF LC Classification classifies the study area into seven classes, ranging from class 11 (high) to class 05 (low). Class 10 (moderate – high) is the dominant class covering 45% of the study area. The resulting map, shown in Figure 3-6 (f) indicates that resolution of the data, used to create this layer (90 m x 90 m grids for terrain and soils variables), is not fine enough to accurately classify and provide continuity between local wetlands and drainage features. The method also appears fairly course, in terms resolution, and aggregates heterogeneous areas into single land capability classes.

3.3.2 Arability and agreement analysis

The arability analysis (Figure 3-7) provides a comparative overview between modelled arability and actual land utilisation across the study area (Figure 3-2). In terms of land utilisation, 77% of the total study area is cultivated, while the remaining 23% is non-cultivated, virgin land. All the land assessment methods indicate that more than 77% of the study area is arable, in effect the landowner is not fully utilising all the viable arable land for cultivation. If a method provided arability results well below actual cultivation rates, it would indicate the landowner was cultivating in, what method would consider, unsuitable areas. Even in a commercial environment, full utilisation of potentially arable land is difficult, due to operational practicalities and management specific objectives. For example, some areas currently under virgin veld, could be ploughed and sustainably cultivated but are currently being utilised for a management specific purpose, in this case the remaining veld plays a critical role in grazing and livestock rotation. Admittedly, land use, particularly in modified agricultural systems, is a function of anthropogenic influence and as such, it cannot be viewed as a perfect reference or baseline indicator, but in this instance, it does provide a satisfactory departure point to judge modelled arability performance.

In terms of overall arability, the RSA and KZN LC methods provide nearly identical results; respectively indicating that 86% and 87% of the study area is potentially arable (Figure 3-7). This result is expected as both methods use similar slope and soil related class breaks to determine broad arability. Arability, in terms Ecotope classification for maize, soybean and pastures indicate that 84% of the study area is suitable for at least one of these land use options. Equally, the Ecotope classification deems the remaining 16% of land unsuitable. These unsuitable areas correspond to soils with shallow effective rooting depths and more marginal soil functional groups. These arability figures, provided by the infield ecotope classification, are significantly higher than estimated arability in the corresponding BRUP inventories, where only 55% of the larger, encompassing Bioresource Unit is considered suitable for annual cropping (Camp, 1999). Ultimately, when placed with a regional context, the study area should be considered above average, in terms of arability and cropping potential.

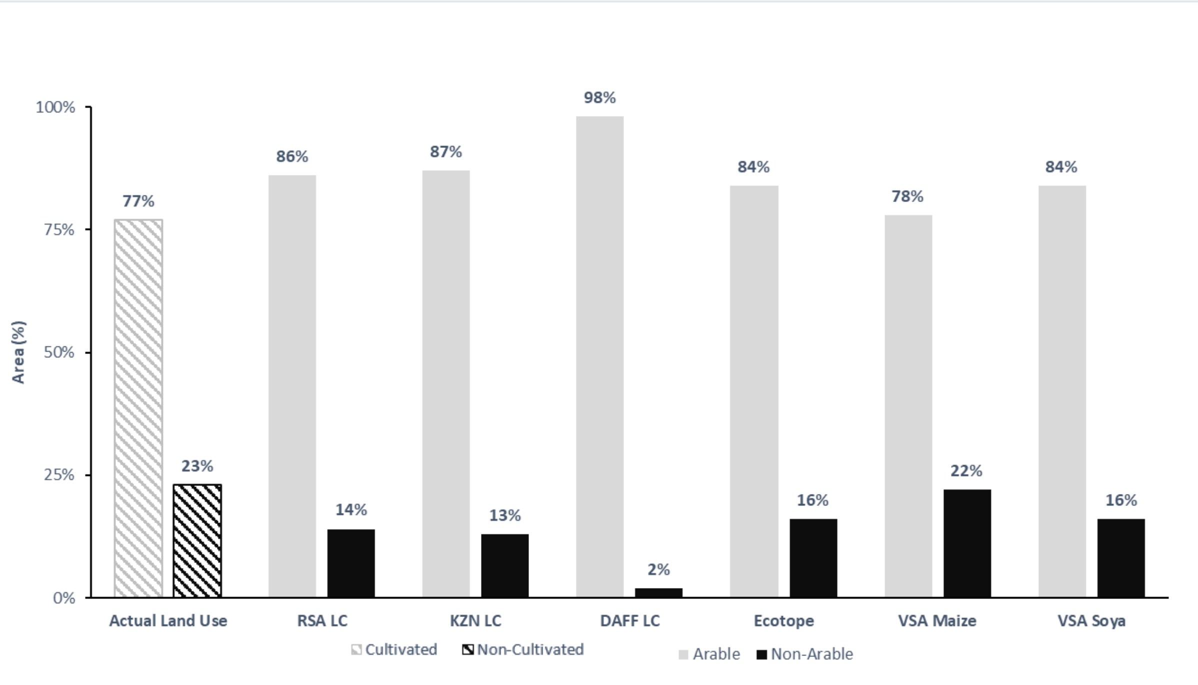

Figure 3-7 Modelled arability for each land assessment method compared to broad actual land utilisation

VSA results for Maize indicate that 78% of the land is classified as either “Good” or “Moderate” in terms of soil quality for maize production. This increases to 84% for potential soybean production using the non-crop specific scoring criteria associated with annual field crops. Comparing the scoring and class breaks for VSA Maize with VSA Annual Crops reveals that for a particular soil to be considered “Poor” for maize production it must score less than 37%, equating to a soil quality index of less than 20 out of a possible 54 points (Shepherd, 2010). While for generic annual fields crops this score drops to 31% or less than 15 out of a possible 48 points (Shepherd et al., 2008). For a soil to be considered “Good” for maize production it must score greater than 69%, while the same soil only requires a score greater than 63 % to be considered “Good” using the annual crops scoresheet. Essentially the soil quality class breaks for Annual Crops, in this case soybean production, are more forgiving and thus an increase in potential arability is expected when compared to that of Maize.

The digital DAFF LC layer indicates that virtually the entire study area (98%) is suitable for arable agriculture, with only 2% of the study area being classified as non-arable. This method fails to classify the majority of limiting soil and terrain features at a local scale, thus overestimating potential arability. The combination of a course 90 m x 90 m digital elevation model resolution (SRTM) and low density of regional soil observation points, used to derive the terrain and soil indices, which ultimately inform the DAFF land capability classification, are not sufficiently accurate determine arability at a local scale in this context.

The graphed results (Figure 3-7) were mapped to spatially highlight areas of modelled arability across the study (Figure 3-8). The polygon maps produced from the five assessment methods were rasterised and reclassified into two distinct classes, arable and non-arable, as per Table 3-2. This rasterization process was performed at a grid size of 50 m x 50 m and produced a total of 7 836 pixels for each assessment method. The resulting maps are provided in Figure 3-8 (a-f). The RSA LC, KZN LC, Ecotope Classification and VSA methods all spatially designate similar areas of non-arability, which include low wetland lying areas, shallow duplex soils, drainage lines and steeper areas (Figure 3-8 a-e).

The broad arability map for the DAFF LC layer (Figure 3-8 f) provides a spatial context to the results shown in Figure 3-8. The DAFF classifies virtually the entire study area as suitable for arable agriculture, with only small areas in the central and southern portions being highlighted as non-arable, spatially confirming the limitations highlighted in results of modelled arability (Figure 3-7), where local non-arable features are not classified correctly.

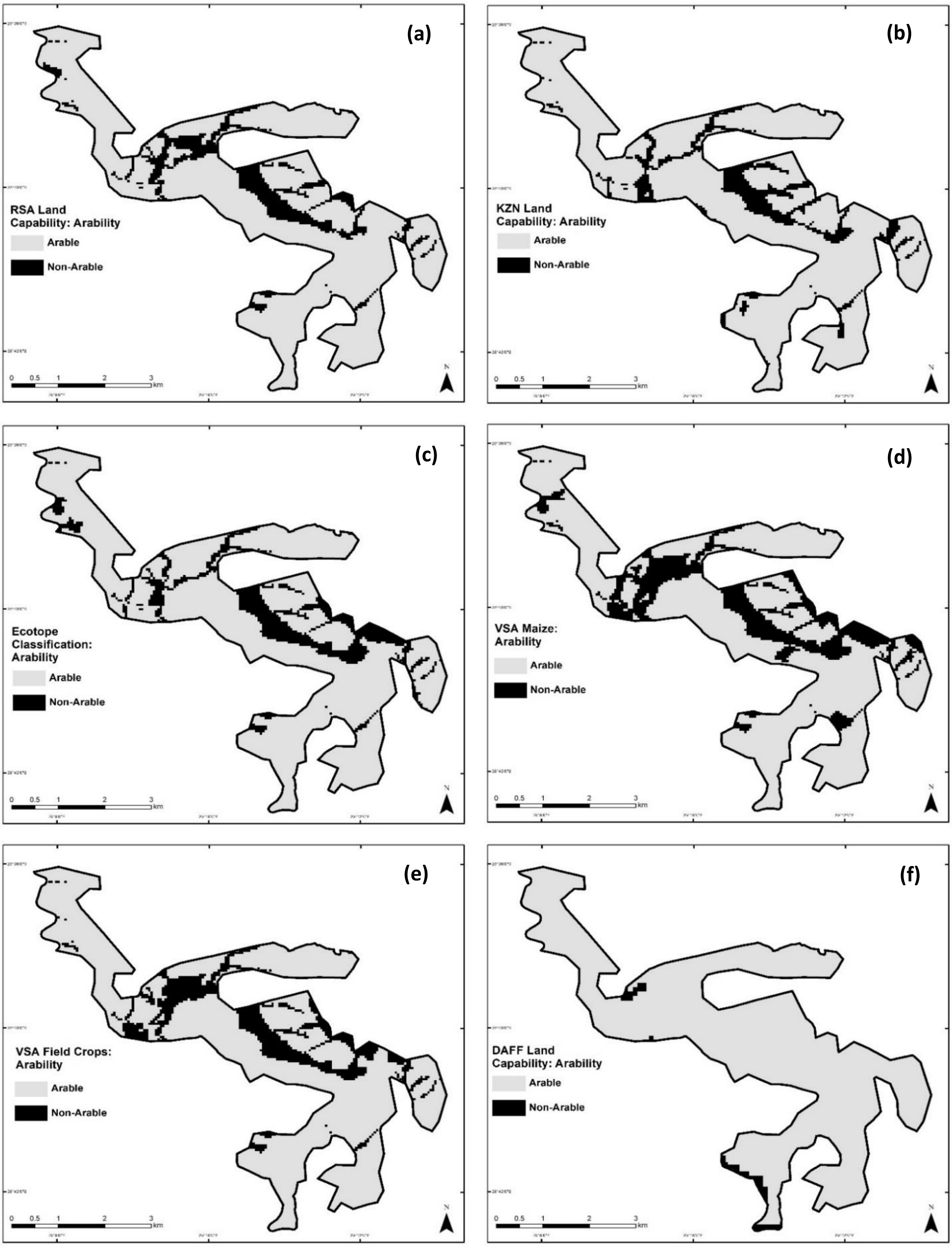

Figure 3-8 Arability classification maps – (a) RSA Land Capability Classification (b) KZN Land Capability Classification (c) KZN Ecotope Classification (d) VSA Maize Classification (e) VSA Annual Field Crops Classification (f) DAFF Digital Land Capability

To spatially determine areas of agreement and disagreement the gridded arability maps, for each method, were compared to actual gridded land utilisation for the study area. The result of this comparative analysis was summarised by six confusion matrices, which assess the accuracy of each classification method against the two land utilisation classes. Ultimately six agreement maps, unique to each of the land assessment methods (Figure 3.9 a-f), spatially represented these matrixes.

A *True Positive*, representing a match between modelled arable and cultivation, dominates most of the study area as cultivation is the dominant land use. These areas are found on gentle to moderate slopes with good soils for arable agriculture.

True Negatives, representing a match between modelled non-arable and a non-cultivated land use, are found in low lying wetland areas, steeper areas, drainage lines and marginal soils with shallow effective rooting depths. Again, the RSA LC, KZN LC, Ecotope Classification and VSA methods all provide similar results in this *True Negative* Class (Figure 3-8 a-e). DAFF LC (Figure 3-8 f) has virtually no areas classified as True Negatives due to the overwhelming bias towards predicted arability.

False Positives, represent areas which are modelled as arable but are in fact non-cultivated. These are generally areas which could, in reality be cultivated but are left fallow for management reasons, for example additional grazing areas for livestock. In this study, land classified as False Positives generally indicates where actual land use is not in fact a good indicator for potential arability.

One positive emanating from this noted disjuncture, is that these areas are delineated relatively consistently for four of the five methods, excluding DAFF Land Capability. Ultimately adding to our confidence in RSA LC, KZN LC, Ecotope, VSA Maize and VSA Annual Crops (Figure 3-8 a-e) to model potential arability. Conversely, DAFF LC has significant areas classified as False Positive, far exceeding realistic potential arability (Figure 3-8 f). This result corresponds to the overarching Provincial arability results provided in DAFF (2018), where nearly 70% of the KwaZulu-Natal Province is classified as land capability class 7 and above, far exceeding other Provincial arability estimates (Schoeman et al., 2002; Smith, 2006).

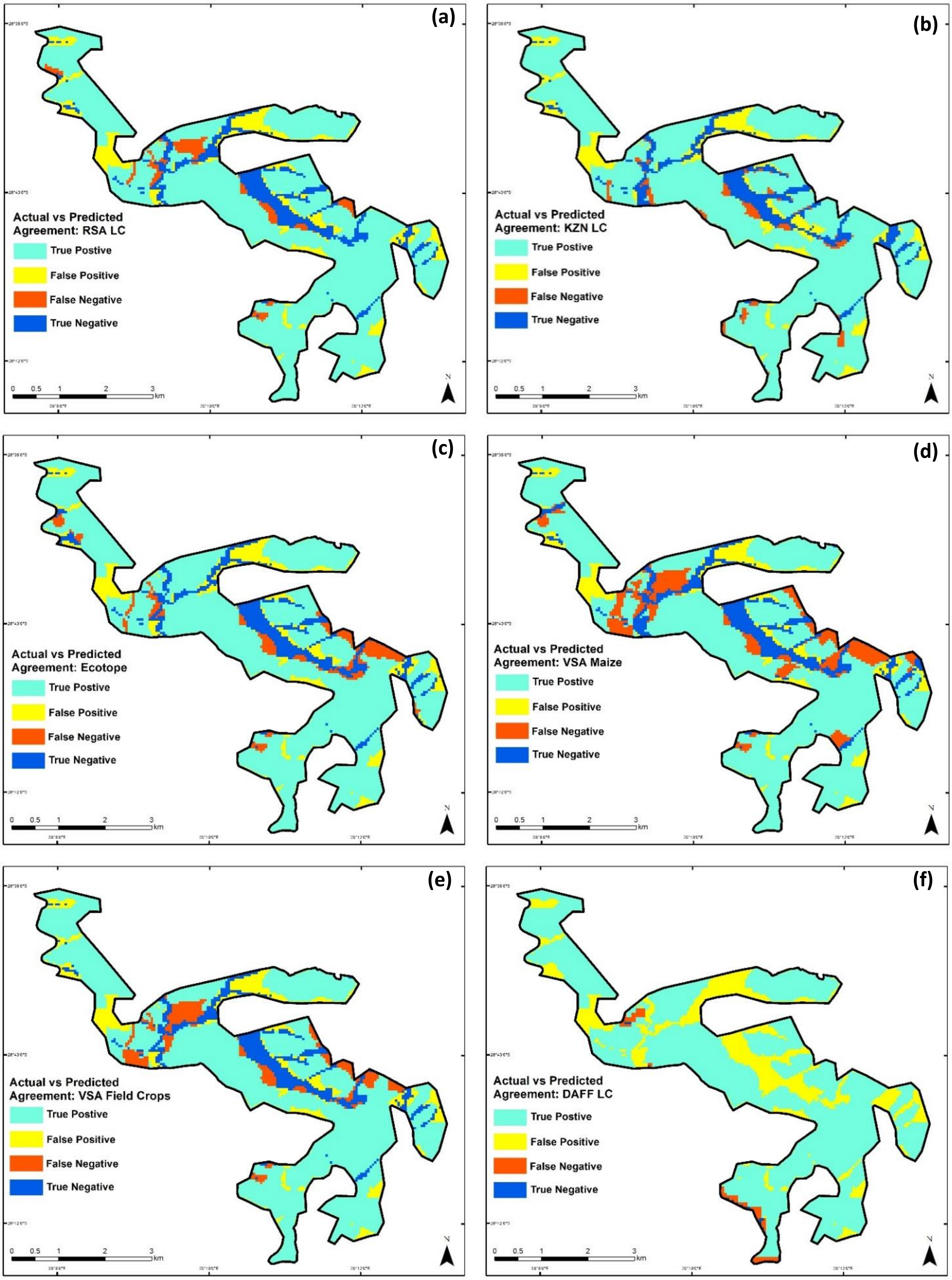

Figure 3-9 Agreement Maps based on confusion matrix results – (a) RSA Land Capability Classification (b) KZN Land Capability Classification (c) KZN Ecotope Classification (d) VSA Maize Classification (e) VSA Annual Field Crops Classification (f) DAFF Digital Land Capability

Finally, a *False Negative represents* where the model predicts a non-arable land use but is in fact cultivated. These areas are generally found in marginal production areas, where some inherent soil or topographic limitation exists but is cultivated regardless, with some risk to yield loss. These can include localised areas associated with significant limitations such as shallow rock shelves and shallow ground water but are still cultivated as part of the larger field, as it is impractical to ringfence these small areas from a management perspective. This again highlights the role of anthropogenic land use and management's influence on the overall accuracy of this arability analysis. Alternatively, the model could simply be wrong, leading to a misclassification of arability.

The compilation of the various confusion matrixes, allows for additional statistical metrics to be extracted, including CA, MR, precision and sensitivity per land assessment method (Eq. 3-1 – 3.5). These resulting metrics are summarised in Table 3-4 and indicate that the KZN Land Capability method has the highest CA with 85.21%. Similarly, RSA Land Capability and Ecotope Classification methods obtained classification accuracies of 84.41% and 83.15%, respectively.

Table 3-4 Classification accuracy (CA), misclassification rate (MR), precision, sensitivity and Matthews Correlation Coefficient (MCC) per land assessment method

Method	CA (%)	MR (%)	Precision (%)	Sensitivity (%)	MCC
RSALC	84.41	15.59	86.19	95.25	0.59
KZN LC	85.21	14.79	86.25	96.40	0.64
Ecotope	83.15	16.85	86.22	93.30	0.53
VSA Maize	78.73	21.27	86.25	86.49	0.39
VSA Annual FC	81.70	18.30	85.84	91.63	0.47
DAFF LC	76.62	23.38	77.80	97.95	-0.03

VSA for annual field crops has a marginally better CA than the VSA Maize, due to its slightly more generic scoring system. DAFF LC has the highest MR of all the methods assessed with 23% of pixels being misclassified, again suggesting its course inherent resolution leads to an increase in MR across the study site.

Statistics in terms of classification Precision and Recall were also provided in (Table 3-4). Classification Precision is the ratio of positive predicted values to actual true values. In other words when the land assessment model predicted arable compared to non-arable, how often was it actually cultivated. RSA Land Capability, KZN Land Capability, Ecotope Classification and VSA all obtained high Precision scores, of approximately 86%. The DAFF Land Capability layer had a lower Precision value of 77.8%, meaning that only 77.8% of the land the model predicted as arable was in fact cultivated.

Classification Recall is the probability that a land assessment method correctly predicts an actual true value. In this case when a pixel was cultivated did the land assessment model predict that it was arable. All assessment methods scored well in this regard (> 86%). The DAFF Land Capability scored predictably high (> 97%) due to the over predication of arable land.

Finally, an overarching performance coefficient was calculated from the six Confusion Matrixes. Matthews Correlation Coefficient (MCC) (Eq. 3-5) incorporates all the individual matrix values to statically evaluate how well the classification performed as compared to a randomly generated classification. MCC is generally considered the best performance coefficient to use in confusion matrix (binary applications) especially when compared to Cohen’s kappa (Delgado & Tibau, 2019) and F1 Score (Chicco & Jurman, 2020). The MCC returns values between -1 and +1. A result of -1 represents a total disagreement between predicted arability and cultivation, 0 is no better than a random predication and +1 score represents a perfect prediction.

The accuracy classification results using the MCC (Table 3-4) show that KZN LC classification is the best performing arability classification when comparing to actual land use with a coefficient score of 0.64. MCC can be interpreted similarly to Pearson Correlation Coefficient where a scores between 0.4 and 0.69 are considered a strong positive relationship (Mukaka, 2012). Therefore, the coefficient scores of RSA LC, KZN LC, Ecotope and both VSA methods represent a strong positive relationship to actual land use. The DAFF LC method scores are very close to 0 and this is considered to have negligible relationship or no agreement (Mukaka, 2012) between predicated arability and actual land use.

The arability analysis indicates that the KZN LC method was consistently the best performer. It achieved the highest CA, precision, recall and MCC of all the land evaluation methods tested. Its performance, in terms of arability prediction, is linked directly to the most fundamental objective, common to all land capability-based methods, which is to determine the lands most basic intensive use (Smith, 1997). Essentially, the performance of the KZN LC should be expected, as this is what the method was designed to achieve, predict broad land use within the confines of KwaZulu-Natal Province. The nationalised land capability system, RSA LC, similarly performed well but whose broader class breaks and factors, compared to that of the KZN system, slightly reduced its CA and overarching MCC scores. This again, should be expected from a broader, nationalised system. The Ecotope classification and both VSA methods are not necessarily designed for ability assessments but with some adaptions, in terms, arability breaks, performed adequately.

Overall, the KZN LC slightly outperforms RSA LC, ecotope and both VSA classification but all five methods performed well and could be confidently used in future arability assessments in similar environments. Results from the DAFF LC digital product indicates a severe overestimation of arability and based on the various analyses should not be used for farm level arability assessments. This result mirrors the Provincial results as provided by DAFF (2018a) which indicates that of the 69.89% of the KwaZulu-Natal Province is potentially arable. This is considerably higher than the 41.80% estimated by Schoeman et al. (2002) as well as other regional estimates, where only 16% of land in KZN is considered suitable for annual cultivation and an additional 8% for permanent crops (KZNDARD, 2018). The primary intended use of the DAFF spatial product is regional land use planning, for both local and district municipalities within a holistic areas based approach (Collett, 2019, 2021). At municipal level, ringfencing contiguous agricultural land is of paramount importance and thus an overestimation of potential arability should not necessarily be seen in a negative light as it should safeguard more and fragment less agricultural land. However, these results indicate DAFF LC is not a reliable indicator of arability at local scale and should not be used for farm planning.

3.3.3 Land assessment polygons and productivity

The aim of this final section of results, is to compare land assessment classification polygons to maize and soybean yield and determine if their resultant land classification is related to actual crop productivity.

3.3.3.1 Productivity contextualisation

It is important to first contextualise the production levels of the study area. The farm portions under analysis should be considered the benchmark, in terms of commercial crop production. Already, the arability analysis (cf Chapter 3.3.2) reveals that 84% of the study area is arable, significantly higher than the regional average of 55% (Camp, 1999). Management, fertilisation and variety selection are all optimised, in an attempt to maximise yields and the commercial operation used in this study is viewed as one of the most successful in the region. This is reflected in the significantly higher yields obtained, across the study, during the five-year analysis period (2016-2020), when compared to Provincial and National yield averages.

According to Maize and Soybean Quality reports produced by *The South African Grain Laboratory* (SAGL, 2015-20), the year-on-year farm average for dryland maize is some 61% above Provincial average and 95% above the National average over the analysis period[1]. While the year-on-year farm average for dryland soybean production is 16% above Provincial average and 70% above the National average. Based on these comparisons, the study area should be considered extremely favourable, in terms of actual productivity and potential.

3.3.3.2 Rainfall and dryland productivity

This analysis only investigates dryland fields thus seasonal rainfall, within the growing season (October and March), becomes a critical factor when assessing yield performance both inter- and intra-annually. Figure 3-10 (a and b), compares the annual average maize and soybean yield to the total seasonal rainfall, as recorded by an on-farm rain gauge. A correlation analysis over the five seasons reports a significant 0.86 correlation (α = 0.9) between maize yield and seasonal rainfall and a 0.97 correlation (α = 0.95) between soybean yields. The analysis indicates that both maize and soybean yields are dependent on rainfall depth. Year-on-year, across the five seasons, it requires an average of 66 mm of seasonal rainfall to produce one tonne of maize per hectare of land. Comparatively, it requires an average of 204 mm of seasonal rainfall per tonne of soybean. National maize guidelines (du Plessis, 2003) reports a maize yield of 3.1t/ha^{-1} requires between 350 and 450 mm of rain per annum, this suggests the yields reported in the study area are significantly above water use norms used Nationally.

Based on long-term rainfall records for Bioresource Unit Wxc5 (Schulze, 1997; Camp, 1999), 2016 and 2019 harvests are considered dry years, 2017 and 2018 are marginally wet and 2020 is considered average (Figure 3-10). Notably, the 2019 season also received very late rainfall with over 60% of the season's rain falling in the last two months of the growing season, negatively impacting yields.

Over the five year period the average yield obtained from all yield observations was 9.7 t.ha^{-1} for maize and 2.9 t.ha^{-1} for soybean. For wet (2017 and 2019) and the average (2020) seasons the average dryland yield obtained increased to 10.7 t.ha^{-1} for maize and 3.2 t.ha^{-1} for soybean. While in the two drier seasons (2016 and 2019) the average yields dropped to 6.8 t.ha^{-1} for maize and 2.4 t.ha^{-1} for soybean, equating to an average decrease of 30% for maize yield and 17% for soybean yield (Figure 3-10). These results are akin to those reported by Wang et al. (2020), who reviewed long term climate and yield studies, between 1961 and 2017, across

[1] Average maize yield was calculated by area weighting white and yellow maize production as reported by SAGL (2015-2020)

China. The findings of these comparative studies, indicate that during severe drought years, yield losses doubled for soybean but increased more than four-fold in maize crops. Essentially soybean appears to be a more drought resistant crop, compared to that of maize.

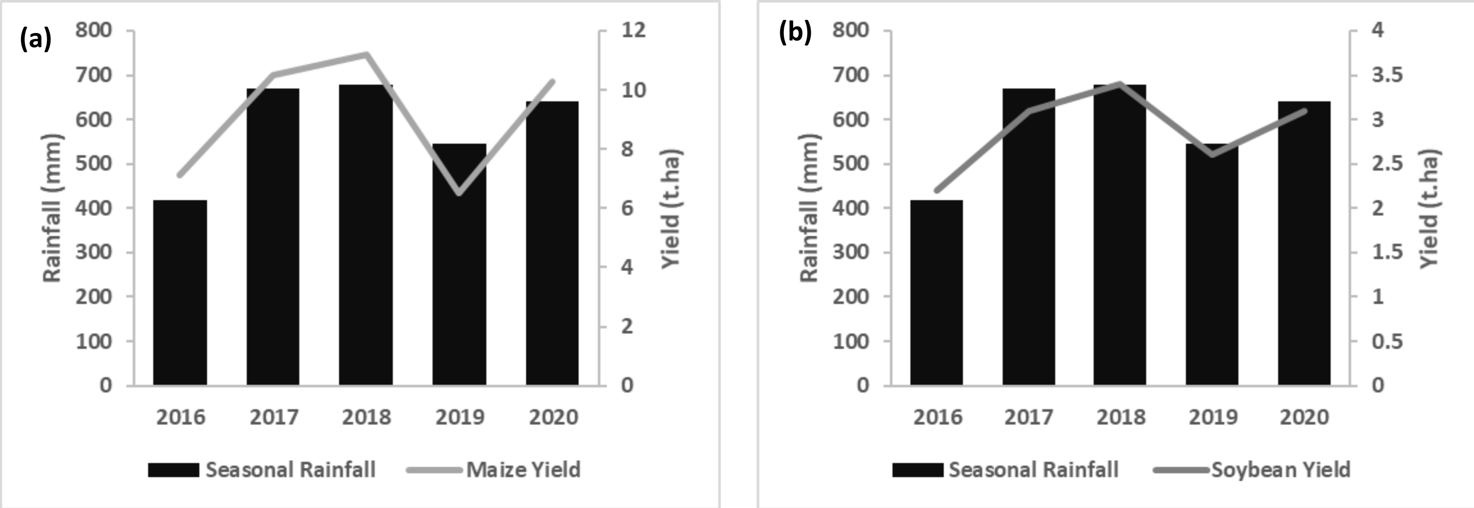

Figure 3-10 Seasonal rainfall (mm) compared to (a) dryland maize yield (t.ha-1) and (b) dryland soybean over five growing seasons

3.3.3.3 Rainfall and crop planting ratios

Forecasted seasonal rainfall, prior to planting, plays a significant role in the ratio of maize to soybean planted in the forthcoming season. Commercial farmers in this region have adapted to seasonal forecasting, by reducing the amount of maize, compared to that of soybean, planted in a predicted dry year, in an attempt to mitigate against yield losses.

For the study area, in forecasted drier years, the ratio between the amount of maize and soybean planted, is relatively equal while in wetter years the ratio is approximately 4:1 in favour of maize production. A crop distribution and yield overview map for 2018, a wet year, is provided in Figure 3-11 (a) as well as a dry year in 2019 in Figure 3-11 (b), illustrating this adaption strategy.

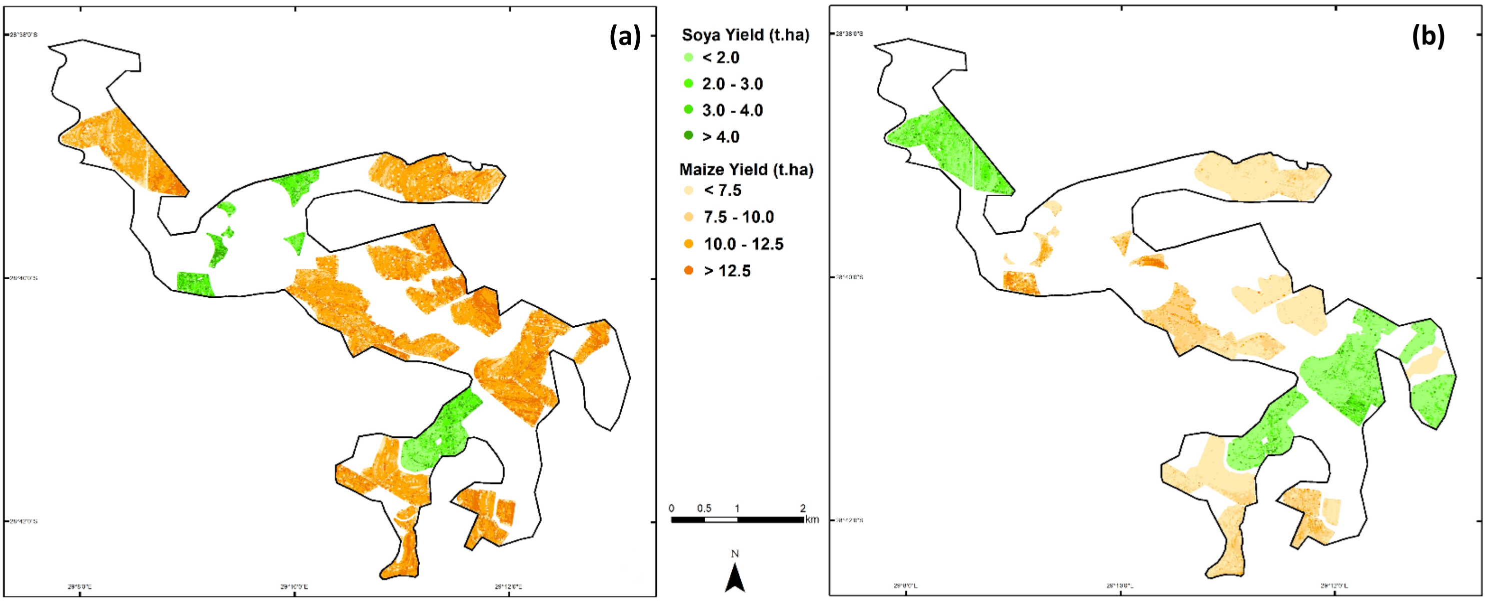

Figure 3-11 Maize and Soybean crop distribution and yield overview for (a) 2018 (wet year) and (b) 2019 a dry year

3.3.3.4 Land assessment methods versus productivity across five growing seasons

To analyse the performance, of each of the five land assessment methods, their resulting land assessment classifications were spatially joined to processed annual maize and soybean yields, for the five growing seasons, between 2016 and 2020. Importantly, the analyses presented below are based on polygon delineation, thus the average yield differs marginally across the various methods. Two seasons, 2016 and 2018 were selected for additional analyses, due to their differing rainfall regimes; the 2016 season recorded below average rainfall while above average rainfall was recorded for the 2018 season. Ultimately, the tables provide an overview of classification performance across all five growing seasons, while the box and whisker plots allow seasonal trends to be highlighted.

- **South African Land Capability Classification**

The primary objective of land capability classification is to arrange land, based on its most intensive use and indicate its permanent hazards, however its classification should also provide a basis for soil productivity and allow for the identification of high potential agricultural land (Scotney et al., 1991).

For maize RSA LC Class III, the highest land capability classification observed in the study area also produced the highest yields, some 2.06% above the average yield (Table 3-5). The remaining three classes (IV, V and VI) all produced below average yields across the five-year period. The results indicate there is a significant difference (p = 0.014) in the distribution of average maize yield across the four classes with significant differences in average maize yields between RSA LC Class III and the non-arable classes V ($p=0.034$) and VI ($p=0.012$). There was also a significant difference RSA LC Class IV and VI ($p=0.035$). The results for maize presented in Table 3-5 broadly indicate that RSA LC classification does relate to actual maize production, with yield decreasing with increasing RSA LC class. One would expect yields to decrease as one moves to lower land capability classes, with non-arable classes producing significantly less than arable ones. This is indeed the case with the performance between Class III land being statistically different from non-arable classes. However, a statistically significant yield difference particularly between arable RSA LC classes (III and IV), would indicate an improved correlation between class breaks and productivity.

Table 3-5 Average dryland maize yield performance per South African Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield ($p>0.05$).

RSA Land Capability Class	n	Average Maize Yield ($t.ha^{-1}$)	Standard Deviation	% Difference Between Class and Method Average Yields
III	341	9.53^{a}	2.36	2.06
IV	151	9.23^{ab}	2.61	-1.22
V	26	8.16^{bc}	3.21	-12.65
VI	11	7.75^{c}	2.01	-17.08

In terms of soybean (Table 3-6), a one-way ANOVA analysis revealed no significant difference in soybean yield across the four RSA LC classes ($p=0.413$). With average soybean yields only varying 0.13 $t.ha^{-1}$ across the four capability classes, with RSA LC Class III actually being the poorest performer across the five growing seasons.

Table 3-6 Average dryland soybean yield performance per South African Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield ($p>0.05$).

RSA Land Capability Class	n	Average Soybean Yield ($t.ha^{-1}$)	Standard Deviation	% Difference Between Class and Method Average Yields
III	241	2.70^{a}	0.72	-1.46
IV	149	2.81^{a}	0.69	2.55
V	23	2.78^{a}	0.77	1.46
VI	6	2.93^{a}	0.81	6.93

Although, both maize and particularly soybean do not show a statistically significant differences between yield and all RSA LC Classes, they do provide some important trends. For maize, yields consistently decrease as land capability deteriorates. Further, maize crops harvested within areas considered “arable” by the RSA LC classification, outperformed those areas classified as “non-arable”. With non-arable areas producing below average maize yields, across the five growing seasons. For soybean, yields tended to increase in lower capability classes, with non-arable areas producing above average yields across the five growing seasons. This suggests that, even though RSA LC is not considered crop specific, RSA LC class breaks may be better suited to maize production compared to that of soybean. Soybean is also seen to be more robust when planted in areas considered to be marginal, by the RSA LC Classification. This is supported by a general consensus that soybeans are more resilient than maize growing in suboptimal soil conditions (DAFF, 2010). Both maize and soybean performed well above the national yield average in areas considered non-arable, suggesting that the RSA LC system's arability classification breaks may be too conservative in this environment.

Box and whisker plots (Figure 3-12 a-d) summarise maize and soybean yield performance for the dry 2016 and wet 2018 seasons. A broad comparison between the seasons shows a relative yield increase, across all classes, from the dry 2016 to the wet 2018 season for both maize and soybean, which is attributed directly to increased rainfall (cf Chapter 3.3.3.2).

For maize the 2016 season (Figure 3-12 a) shows a steady decrease in maize yields between RSA LC Classes III (7.1 $t.ha^{-1}$), IV (6.4t $.ha^{-1}$) and V (6.0 $t.ha^{-1}$). However, the post hoc tests indicates, that even though there is a decrease in yields across these classes the difference is not statistically significantly ($p=<0.720$). The above average yield of 8.4 $t.ha^{-1}$, achieved in RSA LC Class VI in 2016 emanates from a single polygon and this value is most likely an outlier. For the wet 2018 season (Figure 3-13 b) maize yields were relatively constant for classes III, IV and VI, with no statistically significant difference between the average yields achieved. A post hoc test found that the average maize yield for 2018 was only significantly lower in RSA LC Class VI when compared to yields achieved in RSA LC Class III ($p=0.01$) and IV ($p=0.03$). This significant drop in yields in Class VI in the wet 2018 season is mostly attributed to waterlogging. Class VI areas are dominated by wetlands and drainage lines, which become saturated during these wetter seasons, ultimately decreasing yields. Kaspar et al. (2004), whose study focused on the relationship of maize and soybean yields to soil and terrain properties, similarly reported crop yields were negatively affected in closed depressions and low-lying areas in wet years.

The results for soybean (Figure 3-12 c and d) do not provide any clear trends in terms of production versus RSA LC Class, with yields often increasing in poorer capability classes. Again, suggesting that the RSA LC classification and associated class breaks are more suited to maize production than that of soybean. Further there is no statically significant difference in soybean yield across the RSA LC Classes in 2016 ($p=0.119$) and in 2018 the only significant yield difference was between RSA LC Class III and Class IV ($p=0.022$). This result again suggests that expected yield trends do not apply to soybean in this production environment when using the RSA LC classification system at polygon level.

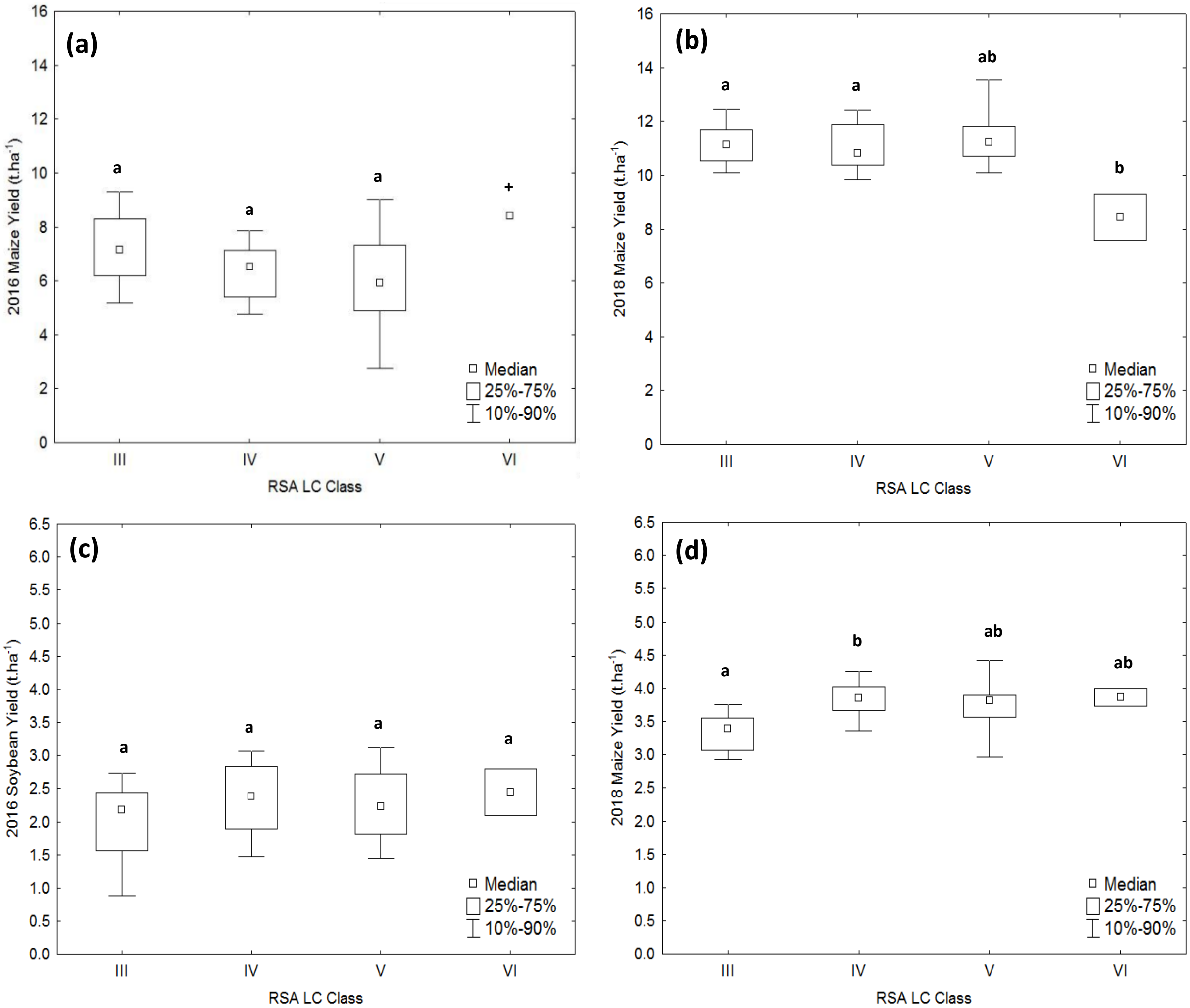

Figure 3-12 Box and Whisker Plots per South African Land Capability Class for (a) 2016 (low rainfall) Dryland Maize Yield (b) 2018 (high rainfall) Dryland Maize Yield (c) 2016 Dryland Soybean Yield (d) 2018 Dryland Soybean Yield. The same letters indicate statistically insignificant differences in yield ($p>0.05$). (+) Denotes classes not included in the ANOVA analysis due to lack of observed samples.

Notably, across all actively cultivated land, only four capability classes were classified using the RSA LC system, consisting of two arable (III and IV) and two non-arable classes (V and VI), while a total of seventeen individual soil forms and four primary slope units were encountered during the resource surveys, exhibiting a wide variety of limiting layers, soil textures, permeabilities and depths. Yet, four land capability classes are able to encapsulate this diversity, across all actively cultivated land. This indicates significant aggregation of soil related properties during the classification process, which may mask more subtle relationships between contributing factors and yield.

As highlighted previously, in the polygon and arability results (cf Chapters 3.3.1 and 3.3.2), one of the most concerning aspects of the RSA LC system is the lack of Class I and II land identified during classification process, especially, if one compares the farm yields against Provincial and National averages (cf Chapter 3.3.3.1). This is an extremely productive farm, and absence of Class I and II lands highlights the problems with RSA LC in that it is too severe in terms of class downgrades, which in this case may not reflect actual productivity. This has ramifications with regards to land release applications via Act 70 of 1970, where productivity potential may detrimentally be trumped by soil conservation and hazard limitations.

➢ **KwaZulu-Natal Land Capability Classification**

The KZN LC classification (Table 3-7), provides a total of five capability classes for the farm which includes one additional arable class (II), compared to that of the RSA LC classification. However, similarly to the National methodology, no Class I land could be identified. As aforementioned, the results should ideally demonstrate a steady decrease in yields, as one moves into poorer capability classes. The ANOVA analysis indicates there is significant difference ($p=<0.001$) in maize yields across the KZN LC Classes, over the five growing seasons (Table 3-7). This trend is clearly evident for maize production across the delineated KZN LC polygons, with yields decreasing from Class II through to Class VI (Table 3-7). Class II is the best performer across the five growing seasons, yielding nearly 6% above the average yield and was found to be significantly different to classes IV, V(a) and VI. Class II is also the most dominate class in terms of spatial coverage, with 49% of cultivated polygons falling into this capability class. Only capability Classes II and III achieved above average yields, while Classes IV, V(a) and VI achieved below average yield. KZN LC Classes V(a) and VI, which are considered non-arable, both obtained below average yields of -14.35 and -20.74% respectively.

The overall results for maize indicate that the KZN LC method is performing adequately, with yields steadily decreasing with land capability class with significant yield differences between most arable and non-arable classes. However, when compared to the national average, the yield performance for maize in these non-arable areas is still high, again suggesting that the arability breaks associated for land capability classification are too conservative in this production environment.

Table 3-7 Average dryland maize yield performance per KZN Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

KZN Land Capability Class	n	Average Maize Yield (t.ha-1)	Standard Deviation	% Difference Between Class and Method Average Yields
II	235	9.53[a]	2.31	5.90
III	56	9.16[ab]	2.43	1.78
IV	174	8.60[bc]	2.66	-4.33
V(a)	15	7.70[bc]	2.03	-14.35
VI	25	7.13[c]	2.72	-20.74

The five-year summary for soybean yields per KZN LC Class, is provided in Table 3-8. Unlike the results for maize, there is little distinction in soybean yield performance across all land capability classes with than 0.12 $t.ha^{-1}$ separating the five delineated classes. This was confirmed by a one-way ANOVA analysis which indicated that there is no statically significant difference in soybean yield across the five KZN LC classes (p=0.744). Class II land, the expected top performer, produced below average a yields across the five growing seasons. With the poorer rated lands, Classes IV, V(a) and VI lands, producing above average yields.

Table 3-8 Average dryland soybean yield performance per KZN Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

KZN Land Capability Class	n	Average Soybean Yield ($t.ha^{-1}$)	Standard Deviation	% Difference Between Class and Method Average Yields
II	143	2.64[a]	0.76	-2.17
III	43	2.65[a]	0.73	-1.77
IV	145	2.76[a]	0.84	2.23
V(a)	14	2.70[a]	0.74	0.06
VI	23	2.75[a]	0.58	2.01

The box and whisker plot for maize production, in the dry 2016 season (Figure 3-13 a), best illustrates the anticipated trends, where yields generally decrease with capability class. A one-way ANOVA analysis indicates there is significant difference (p=<0.001) in 2016 maize yields across the KZN LC Classes, with yields in KZN LC Class II being significantly higher than the yields obtained in Classes IV and VI. While the yields obtained in the wetland class, V(a) was not significantly different, suggesting physical limitations, not associated with severe wetness limitations, are more pronounced during low rainfall years. In the higher rainfall season (Figure 3-13 b), there were no statistical difference between maize yields across three arable classes (KZN LC Classes II-IV). However, the two non-arable classes, V(a) and VI were found to be significantly different from the arable classes, using the non-parametric Kruskal-Wallis Test. The wetland class, V(a), in particular is linked to a significant decrease in maize yields for 2018 season (>3.0 $t.ha^{-1}$). This yield decrease is most likely due to soil saturation and associated waterlogging associated with an above average rainfall season. Similarly, local production guidelines recognise that soils prone to waterlogging result in poor maize yield (CADI, 1993).

The soil component of both the KZN and RSA LC classifications is based primarily on physical soil properties and these properties generally become more pronounced and limiting during drier cycles, resulting in lower maize yields. The 2018 season with higher rainfall appear to mask these physical soil limitations with consistent maize yields (>10 t.ha^{-1}) being achieved in the land capability classes II (KZN LC) and III (RSA LC). Guo et al. (2012) reported similar results for cotton yields where yield and soil properties had a stronger correlation in drier years compared to years with above average rainfall.

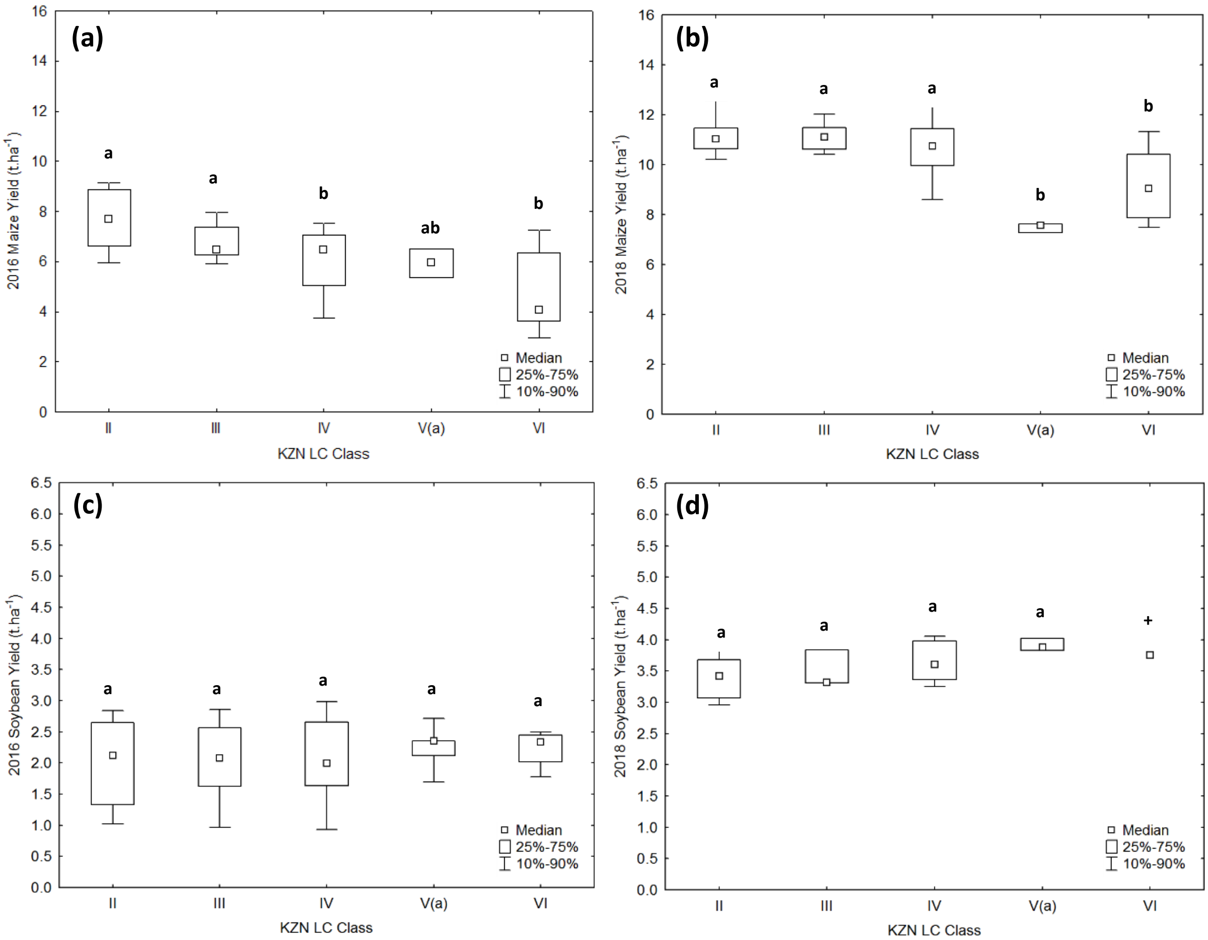

Figure 3-13 Box and Whisker Plots per Kwa-Zulu Natal Land Capability Class for (a) 2016 (low rainfall) Dryland Maize Yield (b) 2018 (high rainfall) Dryland Maize Yield (c) 2016 Dryland Soybean Yield (d) 2018 Dryland Soybean Yield. The same letters indicate statistically insignificant differences in yield (p>0.05). (+) Denotes classes not included in the ANOVA analysis due to lack of observed samples.

The soybean results for 2016 and 2018 (Figure 3-13 c and d), are akin to that of the National classification, where yields do not vary considerably across land capability classes, with yields often increasing in poorer capability classes. The yield variation was confirmed by a one-way ANOVA analysis which indicated that there is no statically significant difference in soybean yields across all KZN LC Classes in either the dry 2016 (p=0.93 a) or wet 2018 (p=0.95) seasons.

- **KwaZulu-Natal Land Crop Ecotope Classification and Productivity**

When compared to traditional land capability classification, evaluating the relationship between crop productivity and ecotope class is more complex, due to the number of classes generated during the land evaluation process. Table 3-9 and Table 3-10, summarise the yield performance of each ecotope class, over the five growing seasons, for maize and soybean. Importantly, crop ecotopes linked to fewer than 10 individual polygons or containing fewer than 2 000 individual yield points, over the five growing seasons, were omitted due to insufficient data.

One of the advantages of the ecotope classification is that it can link to yield data through the BRUP and its associated yield models (Smith, 1997; Camp, 1999). Predicted maize and soybean yields were extracted from the BRUP inventories, allowing additional performance metrics to be calculated, unique to the ecotope classification. These results are provided in the final column of Table 3-9 and 3-10.

The polygon results for maize (Table 3-9) indicate that top two performing ecotopes across the five seasons are Ecotope D23 and B11. Ecotope D23 is a mottled and moderately drained profile (e.g. plinthosol), containing between 15 and 35% topsoil clay, with an effective depth of between 300 and 500 mm, predominately due to wetness limitations. Although it is often seen as a limiting factor in the land evaluation process the presence of soil wetness indicators, at sufficient depths can also benefit dry land crop production, particularly in the more arid part of South Africa or during times of drought (Camp et al., 1995). While, Ecotope B11 is a well-drained profile also containing greater 35% topsoil clay, with an effective depth of greater than 800 mm. This ecotope is considered a high potential soil, with very few physical limitations (Camp et al., 1998). In this environment, topsoils with higher clay contents (>35%), overlying well drained apedal subsoils are generally high producing due to their favourable water holding characteristics (Camp et al., 1995).

Over the five-year period the average yield varied between Ecotope Classes (p=0.008) Well and moderately drained ecotopes (soil groups B and D) were the highest performers, with the top seven performing ecotopes emanating from these groups and all producing above average yields. These soils typically have an apedal subsoil horizon, underlying a weakly structured topsoil with greater than 15% clay content. This combination of soil properties is considered optimal for maize production, as outlined in the maize production guidelines for South Africa (du Plessis, 2003). In terms of yield differentials across the various ecotopes, the most significant differences occur between Poorly Drained Soils (Group E) and Duplex Soils (Group J) and the well and moderately drained (Groups B and D). Poorly drained ecotopes (e.g. E23), and duplex soils (e.g. J22) consistently produced below average yields. In the case of Ecotope E23 this is most likely due to intermittent waterlogging, while for Ecotope J22 the subsoil is characterised by a marked enrichment of clay in the subsoil, limiting rooting depths (Fey, 2010).

Across all ecotopes the maize yield recorded was consistently higher, than the predicted yield benchmark provided in the BRUP inventories (Table 3-9). For higher potential ecotopes (B and D soil groups), this difference varied between 13 and 113%. Effectively, the modelled yield provided in the BRUP inventories, is significantly under estimating yields compared to what is currently being achieved in commercial environments. The crop yield models used in the BRUP were developed by Smith (1997) and have not been significantly updated to take into account recent genetic improvements and technological advancements, such as precision agriculture, which are currently being applied. Egli (2008) reports significant yield gains in recent times for of both maize and soybean, these gains need to be incorporated into local yield models. All ecotopes considered to be unsuitable for maize production by the BRUP, generally due to depth restrictions, all produced below average yields.

Table 3-9 Average dryland maize yield performance per crop ecotope class polygons over five growing seasons (2016-2020), (>2000 yield points and >10 individual polygons). The same letters indicate statistically insignificant differences in yield (p>0.05). Dotted lines separate primary soil groups.

Crop Ecotope Class#	n	Average Maize Yield (t.ha-1)	Standard Deviation	% Difference Between Class and Method Average Yields	% Difference Between Class and BRU Benchmark Yield
B11	20	10.37^{ab}	2.25	6.92	31.29
B21	79	9.72^{ab}	2.26	0.18	34.97
B22	36	9.83^{ab}	2.33	1.33	51.22
B23	13	9.19^{ab}	2.91	-5.21	58.53
D11	21	10.30^{ab}	2.34	6.16	43.02
D21	65	9.70^{ab}	2.08	0.02	12.82
D22	66	9.34^{b}	2.33	-3.75	60.97
D23	15	10.67^{a}	2.60	10.03	113.40
E22	12	8.95^{b}	2.20	-7.74	108.12
E23	63	8.69^{b}	2.75	-10.45	0.00*
J23	34	8.93^{b}	2.42	-7.96	0.00*
J24	12	7.87^{b}	3.06	-18.91	0.00*

Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class
* Denotes ecotopes which are deemed unsuitable for soybean production in the BRUP Inventories

The polygon soybean results indicate that five-year average yields are more evenly distributed across a wide range of crop ecotopes (Table 3-10). Consequently, there was no statistically significant difference between soybean yield and ecotope (p=0.164). This unpredictability in terms of performance is further highlighted, where the top three performing ecotopes, all come from different soil groupings.

Although statistical significance could not be determined some important trends were observed. Deep, well drained soils (Group B) performed poorly while ecotopes traditionally considered limiting, from a cropping perspective, performed well. For example ecotope J24, which is characterised by a strongly structured, clay enriched subsoil horizon (planosols), underlying a less structured topsoil with shallow effective soil depth of between 200 and 300 mm deep, produced above average yields. Pedocutanic and prismacutanic B horizons are typically classified in this ecotope grouping and whose associated soil structure grade is often considered limiting to root growth and water movement. However, the growing guidelines for soybean states that soybeans are generally better adapted to heavier soils and better able to utilise water at lower soil depths than most other crops, including maize (DAFF, 2010). Similarly, a study by Cox et al. (2003), investigating the relationship between soil properties and soybean yield, reported that higher clay contents resulted in higher yields, across three fields.

Ecotope E24, a poorly drained soil, is the poorest performer in terms of soybean yield and this result is reflected in the Soybean growing guidelines for South Africa (DAFF, 2010). This guideline reports that maximum seed yield is possible where water in the root zone is kept above 50% plant-available, while waterlogged conditions, as one would expect for gleyed soils, will have a negative effect on the crop yield (DAFF, 2010).

As with maize, comparing actual soybean yields to the modelled BRUP inventories indicates all arable ecotopes are producing well above BRUP benchmarks. Ultimately, the yield results indicate the crop models for both maize and soybean in the BRUP require revision to better reflect current production norms.

The ecotope box and whisker plots, for maize and soybean production, for the 2016 and 2018 seasons are provided in Figure 3-14 a-d. For maize production in a low rainfall year (Figure 3-14 a), the median yield for all soil functional groups is below 9 $t.ha^{-1}$. There are significant yield differences in both the 2016 (p=0.026) and 2018 (p=0.001). Yields, generally correspond to the long-term trends (Table 3-10), where well drained and mottled soils (Soil Groups B and D) are the highest performers. However, soils within these groups, with shallower effective depths (300 – 500 mm), are associated with lower yields in this low rainfall year. In a drier year, such as 2016, deeper profiles allow roots access to larger water reserves, thereby alleviating water stress under dryland conditions (Shepherd, 2010).

Table 3-10 Average dryland soybean yield performance per Crop Ecotope Class polygons over five growing seasons (2016-2020), (>2000 yield points, >10 polygons). The same letters indicate statistically insignificant differences in yield (p>0.05). Dotted lines separate primary soil groups.

Crop Ecotope Class #	n	Average Soybean Yield (t.ha^{-1})	Standard Deviation	% Difference Between Class and Method Average Yields	% Difference Between Class and BRU Benchmark Yield
B11	13	2.58^{a}	0.78	-4.18	3.10
B21	54	2.67^{a}	0.73	-0.74	16.09
B22	35	2.43^{a}	0.83	-9.56	21.64
B23	14	2.92^{a}	0.73	8.51	62.17
D11	14	2.94^{a}	0.50	9.14	17.44
D21	27	2.77^{a}	0.55	2.82	2.44
D22	54	2.49^{a}	0.85	-7.29	38.55
D23	11	2.74^{a}	0.83	1.85	71.23
E23	55	2.99^{a}	1.15	11.07	0*
E24	10	2.35^{a}	0.65	-12.82	0*
J23	37	2.71^{a}	0.71	0.78	0*

Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class
* Denotes ecotopes which are deemed unsuitable for soybean production in the BRUP Inventories

Both the Soil Groups D and E soil perform well in a drier year, where a fluctuating water table, found close to the surface, acts as reservoir for the plant roots. Duplex soils, Soil Group J, associated with a strongly-structured subsoil horizon, perform poorly in a dry year, as these clay soils restricted water abstraction from plant roots (Asgarzadeh et al., 2010). In a wetter season, (Figure 3-14 b), yield performance across ecotopes is not as definitive, with virtually all soil groups attaining a median yield of greater than 10 t.ha^{-1}. Generally, functional Soil Groups B and D are still the top performers, however clay content and effective depth do not appear to influence yield significantly. However, waterlogging and associated yield loss in the gleyed ecotope EI4 was significant in a wet season when compared to higher yield ecotopes.

In a dry year (Figure 3-14c), the non-parametric Kruskal-Wallace Test indicates a significant (p=0.005) difference in soybean yield across the Ecotope Classes. Crop Ecotope I24 is the top performer, this Ecotope is a poorly drained soil with signs of wetness, in most cases a soft plinthic B subsoil, very close to the surface. The effective depth is estimated to be between 200 and 300 mm deep. The yield performance is unexpectedly high, as even in a dry year surface ponding and a water logging is a distinct possibility. Compared to maize, duplex and strongly structured soils (Group J) performed above average production levels for soybean for both the dry 2016 and wet 2018 seasons.

In wet 2018 season (Figure 3-14d) Soil Group B, well drained profiles, are the poorest performers, particularly the shallow variants (B22 and B23). The employed ad hoc tests for multiple comparisons could find a significant difference between soybean yield and Ecotope Class. However, the results do indicate that in a wet year the expected trends are almost perfectly inverted with well drained and moderately drained soils groups being the poorest performers, and the more physically limiting E and J functional groups performing significantly above expectations. The seasonal results for soybean do not follow expected yield trends as outlined in the BRUP inventories. In terms of general trends, first one would expect yields to be highest within the B and D soil groups and taper off within the more physically limiting the E, H, and J soil groups. Second, within a particular soil group, yields should respond to changes in topsoil clay and effective depths. Finally, different soil groups should respond to changes in rainfall, for example mottled and poorly drained soils should fare better in lower rainfall seasons (Figure 3-14 a and c), while a high rainfall year should correspond to water logging and potential yield loss (Figure 3-14 b and d).

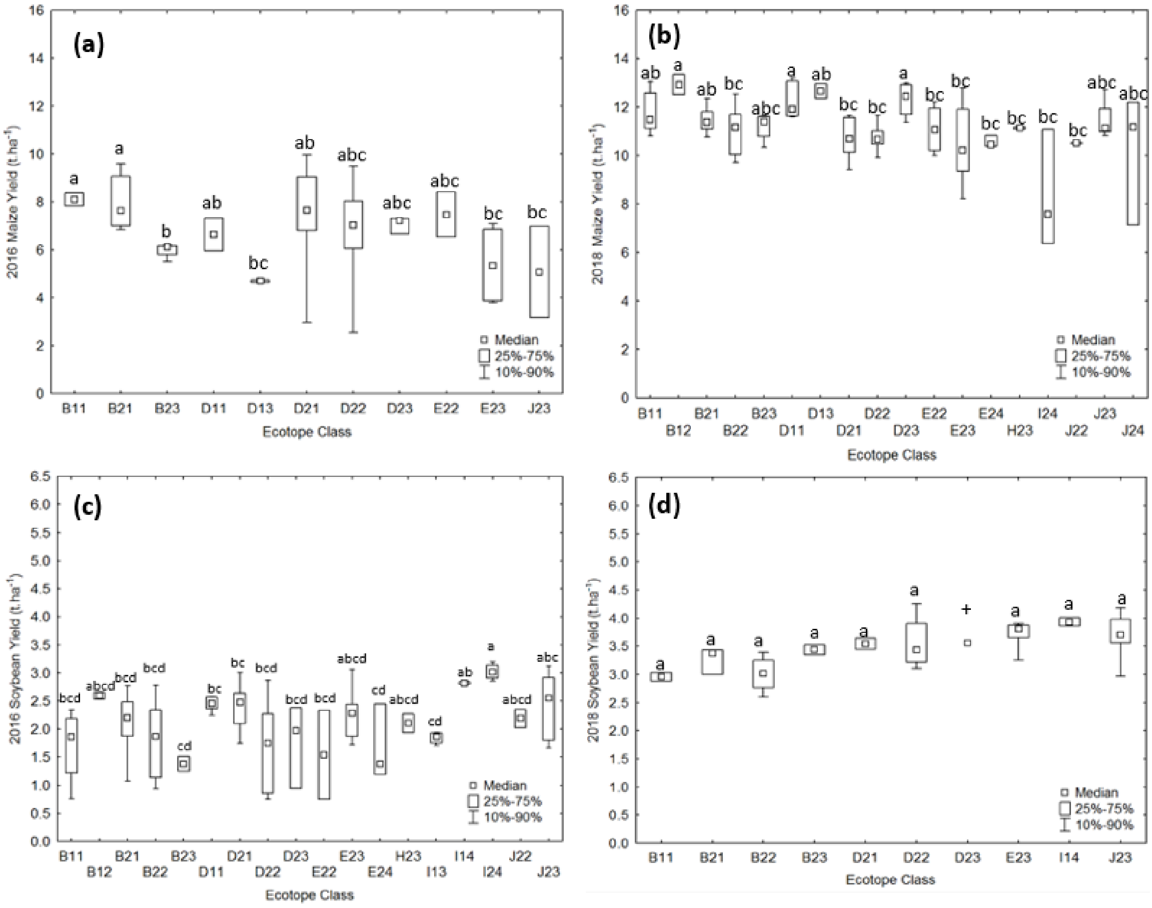

Figure 3-14 Box and Whisker Plots per Ecotope Class for (a) 2016 Dryland Maize Yield (b) 2018 Dryland Maize Yield (c) 2016 Dryland Soybean Yield (d) 2018 Dryland Soybean Yield. The same letters indicate statistically insignificant differences in yield (p>0.05). (+) Denotes classes not included in the ANOVA analysis due to lack of observed samples.

➢ Visual Soil Assessment and Productivity

The results of the Visual Soil Assessments, undertaken in the study area, provide an indication of both soil quality and potential plant performance (Shepherd, 2010). The field guides for maize (Shepherd, 2010) and annual field crops (Shepherd et al., 2008), state that higher scoring soils will generally be in better condition and soils with "Good" VSA scores will, by in large, give the best production performance.

The average dryland yield performance of maize, per VSA class, is provided in Table 3-11. Over 64% of the actively cultivated polygons, across the five growing seasons, fall within the "Good" VSA class, the highest class possible. This indicates that the VSA scoring methodology for maize is able to determine that the majority of the cultivated soils on this farm are of high potential, matching the farms actual high productivity benchmark. These soils also produced the highest yields of the three VSA classes and performed slightly above the farm average (3.4%). The results of the one-way ANOVA indicate a statistically significant result (p=0.01), with the Tukey HSD post hoc test showing that the "Good" VSA class yielded significantly more maize when compared to the "Moderate" (p=0.040) and "Poor" (p=0.004) classes. This result shows that the VSA indicators for maize are linked to actual production performance when using large-scale polygons.

Table 3-11 Average dryland maize yield performance per Visual Soil Assessment Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

VSA Class	n	Average Maize Yield ($t.ha^{-1}$)	Standard Deviation	% Difference Between Class and Method Average Yields
Good	308	9.61^{a}	2.47	3.40
Moderate	84	8.85^{b}	2.66	-4.82
Poor	87	8.61^{b}	2.69	-7.37

The soybean yield performance, per VSA class, over the five seasons is provided in Table 3-12. A one-way ANOVA analysis revealed no significant difference in soybean yield across the three VSA classes (p=0.142). The "Moderate" and "Poor" soil quality classes, for soybean, are the highest performers, with both classes scoring above average yields, across the five seasons. However, similarly to RSA LC, KZN LC and Ecotope Classifications, there is little distinction in soybean yield performance across the various classification methodologies, with only 0.26 $t.ha^{-1}$ separating the three VSA classes. The results also indicate that, compared to maize (36%), soybean is generally planted on more marginal soils in terms measurable soil quality for annual crops, with nearly half of the planted soybean falling within a polygon classified as either "Moderate" and "Poor" in terms of soil quality.

Table 3-12 Average dryland soybean yield performance per Visual Soil Assessment Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

VSA Class	n	Average Soybean Yield (t.ha^{-1})	Standard Deviation	% Difference Between Class and Method Average Yields
Good	184	2.76^{a}	0.78	-2.09
Moderate	115	2.83^{a}	0.81	0.69
Poor	41	3.02^{a}	0.82	7.45

Box and whisker plots (Figure 3-15 a-d) summarise both maize and soybean yield performance for the 2016 and 2018 seasons, across the three VSA Classes. The results are consistent with the other land evaluation methodologies, where the maize yields in the dry 2016 season (Figure 3-13 a), best illustrate the anticipated trends, with yields decreasing with VSA Class. However, in this case the VSA classification also indicates that for the 2018 season there is also significant difference in maize yields with the yields produced within the "Good" Class being significantly higher than the yields in "Poor" soil quality class. As with previous land evaluation methodologies the higher rainfall experienced in the season 2018 may mask various soil physical limitations. Consequently, the negative impact of determinantal physical characteristics, such as those linked to textural or structural problems will be lessened, ultimately reducing yield variation across the VSA classes.

The expected yield trends are not present in either of the 2016 or 2018 soybean harvests (Figure 3-15 c and d). The seasonal soybean results indicate an inverse relationship between yield and visual soil quality indicators with yields increasing significantly between the "Good" and "Poor" VSA Classes for both the 2016 ($p=0.021$) and 2018 ($p=0.001$) seasons. The comparative seasonal results and yield trends, between maize and soybean, suggest the scoring system, designed specifically for Maize (Shepherd, 2010) is more capable than the more generic annual field cropping scoring system (Shepherd et al., 2008), as used for soybean.

Ideally, from a farm management perspective, soils which fall within the "Poor" or "Moderate" VSA classes would be identified and suitable interventions prescribed. These interventions would hopefully improve soil quality attributes and ultimately improve yield performance. The consequence of the "Poor" class achieving the highest soybean yield suggests that spending money on soil quality improvements does not always equate to yield improvements, essentially, sending a message to farmers that "Poor" quality soils are able to perform as well or better as "Good" quality ones, increasing the hesitancy of applying expensive remediations to improve or sustain soil quality. Improvements could include subsoil drainage, incorporation of additional organic matter and maintenance of good cover. This may have detrimental effects on long-term sustainability and ultimately increase land degradation.

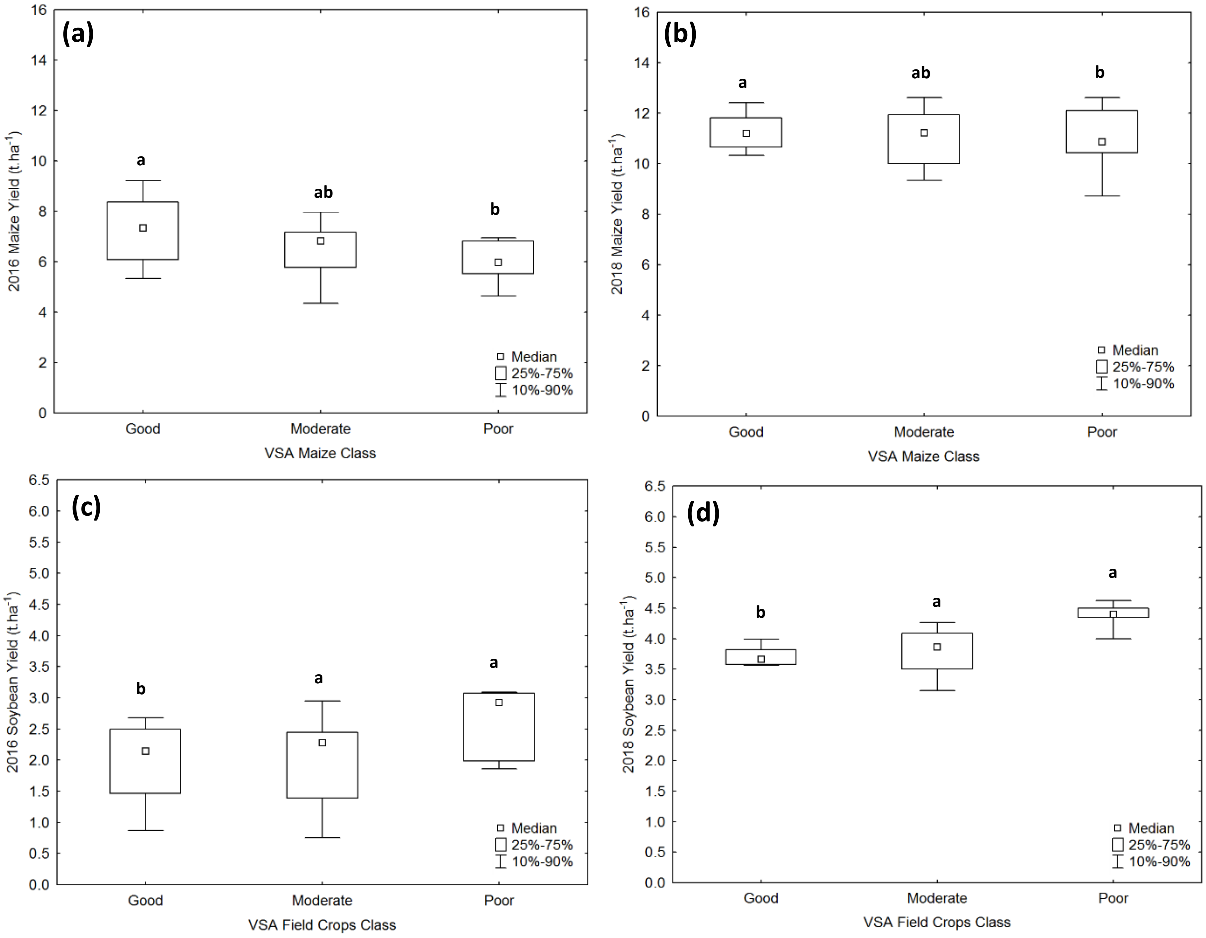

Figure 3-15 Box and whisker plots per Visual Soil Assessment Class for (a) 2016 (low rainfall) dryland maize yield (b) 2018 (high rainfall) dryland maize yield (c) 2016 dryland soybean yield (d) 2018 dryland soybean yield. The same letters indicate statistically insignificant differences in yield (p>0.05).

➢ DAFF DIGITAL LAND CAPABILITY AND PRODUCTIVITY

Although consisting of many more contributing attributes, the DAFF digital land capability classification, is based on the same premise as the classical eight-class systems, but whose rating is expanded to fifteen classes and inverted. This inversion translates to a classification where the higher the capability class the more production potential. Table 3-13 provides a summary of the average maize yield for all relevant DAFF LC across the five growing seasons.

A total of seven DAFF LC classes occurs in the study area, however nearly 70% of actively cultivated areas fall within classes 9, 10 and 11 *Moderate to High* and *High*. The highest yielding classes are 5 (*Low*) and 6 (*Low to Moderate*). However, the low number of contributing polygons, in these specific land capability classes, may be influencing these results and not adequately reflecting production potential. A one-way ANOVA indicates there is no significant difference between maize yields and DAFF LC classification across all classes. Overall, there is no clear relationship or trends between DAFF LC Class and maize production across the five growing seasons, with less than a 10% maize yield differential, across all observed classes.

Table 3-13 Average dryland maize yield performance per Department of Agriculture, Forestry and Fisheries Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

DAFF Land Capability Class	n	Average Maize Yield (t.ha-1)	Standard Deviation	% Difference Between Class Yield and Farm Average Yield
5 Low	12	9.60^a	2.11	7.37
6 Low - Moderate	16	9.78^a	2.35	9.37
7 Low - Moderate	37	8.64^a	2.71	-3.44
8 Moderate	81	8.47^a	2.74	-5.34
9 Moderate - High	160	8.72^a	2.72	-2.53
10 Moderate - High	52	8.73^a	2.74	-2.46
11 High	129	9.54^a	2.61	6.62

The five-year results for soybean (Table 3-14), are akin to the maize results, whereby the majority of contributing polygons (>67%) come from DAFF LC Classes 9, 10 and 11. Only 0.34 $t.ha^{-1}$ separates the highest and lowest yielding class and these low yield differentials and makes performance analysis difficult. This is confirmed by a one-way ANOVA which found no significant yield difference (p=0.804) between the seven DAFF LC Classes.

Table 3-14 Average dryland soybean yield performance per Department of Agriculture, Forestry and Fisheries Land Capability Class polygons over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences in yield (p>0.05).

DAFF Land Capability Class	n	Average Maize Soybean (t.ha-1)	Standard Deviation	% Difference Between Class Yield and Farm Average Yield
5 Low	9	2.67^a	0.64	-0.41
6 Low - Moderate	14	2.65^a	0.62	-0.99
7 Low - Moderate	28	2.43^a	0.60	-9.33
8 Moderate	58	2.63^a	0.72	-1.84
9 Moderate - High	120	2.77^a	0.92	3.21
10 Moderate - High	40	2.66^a	0.73	-0.73
11 High	63	2.69^a	0.77	0.47

Box and whisker plots (Figure 3-16 a-d) summarise both maize and soybean yield performance for the 2016 and 2018 seasons, across the seven DAFF LC class. Similarly, to the five year analysis no definitive trends can be observed between crop production and DAFF LC Classes in either a dry (2016) or wet year (2016). For maize the yield difference across all

DAFF LC Classes was significant, for both 2016 (p=0.019) and 2018 (p=0.49), but the post hoc tests could only provide a significant yield difference between two arbitrary classes (Figure 3-16 i-ii).

For soybean no significant difference between yield and the DAFF LC Classes could established for either 2016 or 2018 (Figure 3-16 iii-iv). As similarly stated in the arability exercise (cf Chapter 3.3.2) the combination of a course digital elevation model resolution and a low density of regional soil observation points, which are used to derive these DAFF LC Classes, appear not to be sufficiently accurate to determine maize or soybean production levels at a farm scale.

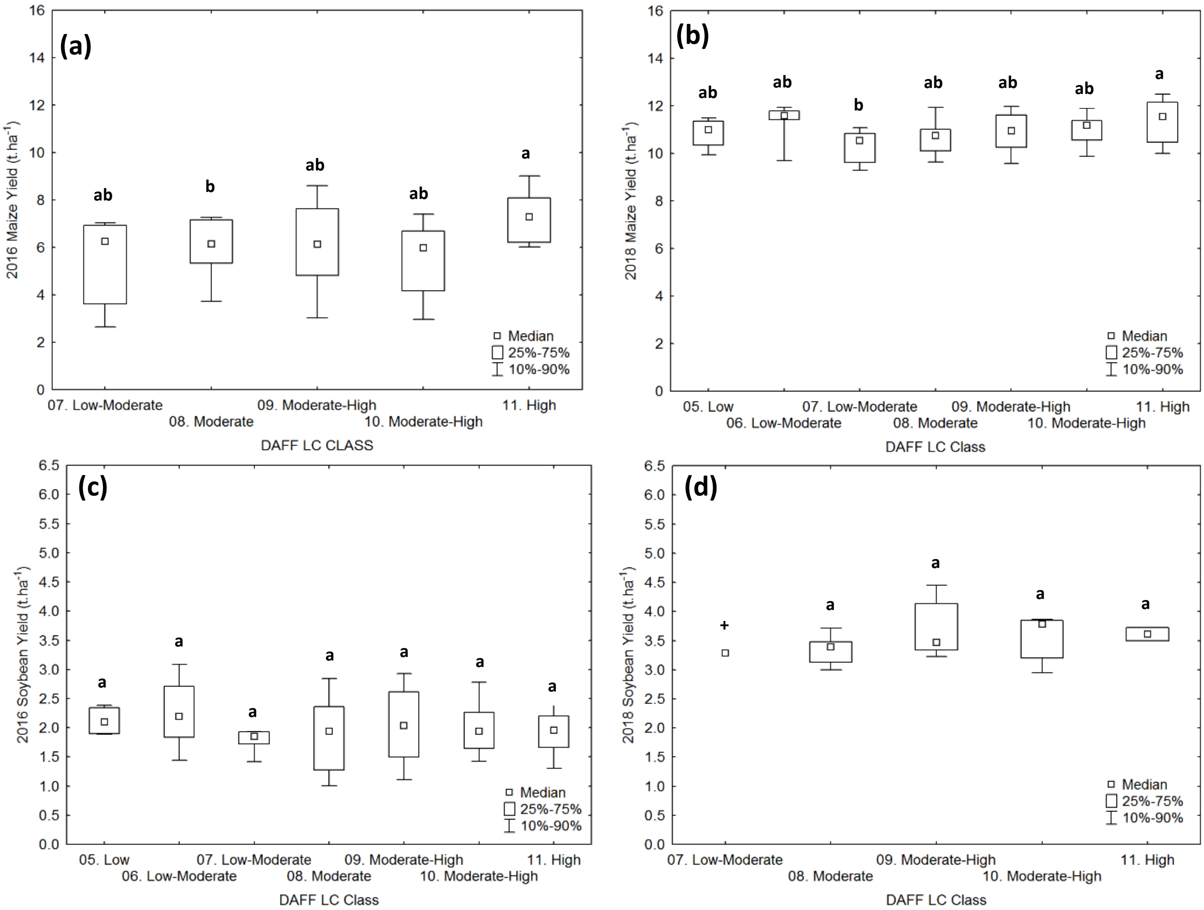

Figure 3-16 Box and Whisker Plots per Department of Agriculture, Forestry and Fisheries Land Capability Class for (a) 2016 (low rainfall) Dryland Maize Yield (b) 2018 (high rainfall) Dryland Maize Yield (c) 2016 Dryland Soybean Yield (d) 2018 Dryland Soybean Yield. The same letters indicate statistically insignificant differences in yield (p>0.05). (+) Denotes classes not included in the ANOVA analysis due to lack of observed samples.

3.3.3.5 Assessment Deficiencies

One of the most highlighted land evaluation deficiencies, identified by the literature review (cf Chapter 2.7), is the absence of local verification studies. This polygon-based analysis followed the recommended methodology for land evaluation in South Africa (DAFF, 2018), wherein soil observation densities and soil survey standards were adhered to. However, during the course of the verification exercise three factors were considered limiting.

First, is the lack of practical validation metrics, which can be employed to effectively evaluate the performance of the various methodologies. Ultimately in this analysis, land use and crop productivity, were selected as the primary validation metrics. Both, however, are significantly influenced by anthropogenic factors, which detract from the overall accuracy of the analysis.

Second, not all of the selected land assessment methodologies were developed to specifically determine arability or crop productivity performance. Consequently, some of the model deficiencies, identified through these validation processes should be expected.

Third, land evaluation polygons are generally delineated from soil maps, which in turn are created by upscaling soil point observations. In this analysis, yield points were linked to resulting land evaluation polygons, however yield performance and variation can only be explained at or near the observation point. As one moves away from each observation point, differing yield drivers cannot reliability be determined or analysed. This deficiency is addressed in Chapter 4.

3.4 CONCLUSIONS

The aim of this Chapter was to determine whether any of the five unique land evaluation methodologies selected, could adequality reflect actual land utilisation and production levels at polygon level.

The arability analysis indicates that the KZN LC method was consistently the best performer. This methods performance, in terms of arability prediction, is expected as this method was designed for this exact purpose, wherein climate, soil and terrain variables are combined to classify land into broad capability class, within the confines of the KwaZulu-Natal Province. Although KZN LC slightly outperforms RSA LC, ecotope and both VSA classifications all four methods performed well and could be confidently used in future arability assessments in this production environment. The arability results for the DAFF LC digital product indicates a severe overestimation of arability and based on the various analyses and performance indicators, this method should not be applied in future farm level arability assessments.

The analysis comparing land assessment polygons to maize and soybean yields, produced mixed results but important principles emerged. 1) Seasonality does influence the relationship between land classification and yield, with physical factors becoming more apparent in drier years. 2) A single land evaluation class, often dominates the sample distribution, making deductions difficult. 3) The analysis highlights the danger of utilising non-crop specific methodologies, as results and seasonal trends differ significantly between maize and soybean. 4) Maize yields had stronger relationship to land evaluation polygons, compared to that of soybean where significant yield differences were rarely established 5) The highest maize yields generally corresponded to the best land evaluation class or class with highest cropping potential. 6) Although, the ecotope methods produced significantly more classes, greater detail could be extracted from the resulting classification. 6) Modelled yields based on ecotope classification and BRUP inventories consistently underestimate yields and these models should be updated to reflect contemporary varieties and management. 7) Long term yield studies tend to average out and mask important trends, leading to small yield differentials between land evaluation classes. 8) All land evaluation methodologies and associated verification metrics have limitations. Both should be established and quantified so that the most suitable method(s) are applied, under the correct conditions and ultimately provide the most accurate and reliable results to decision makers. 9) Finally, land evaluation polygons, linked to precision yields can provide a general overview of method performance. However, yield performance and variation, across land evaluation methods and classes, is only explicit on or near a soil observation point where measurements are taken. Thus, a point-based verification of land assessment methodologies is required to better understand the drivers affecting crop performance.

4. VERIFICATION OF LAND ASSESSMENT RESULTS AND INDIVIDUAL ATTRIBUTES THROUGH BUFFERED POINT YIELD ANALYSIS

4.1 Introduction

A fundamental aim of many land evaluation methodologies is to provide an indication of agricultural potential, specifically in terms of crop productivity (Mueller et al., 2010). Methodologies come in many forms; *inter alia* numeric rating schemes, such as Soil Potential Ratings (USDA, 2011) simple categorical data, as provided by many land capability systems as well as crop specific suitability classifications linked to crop yield models, such as the KZN Ecotope Classification (Camp et al., 1998; Smith, 2006).

Yield, usually expressed by the amount of crop harvested per unit area, is recognised as the most commonly used indicator to assess agricultural productivity (Diskin, 1997; Wineman et al., 2019). Merits of crop yield as an indicator include its ease of calculation, widespread applicability and intuitive interpretation (Reynolds et al., 2015). Accordingly, crop yield is recurrently used to benchmark the performance of various production-based models *viz.* agricultural (e.g. Pakawanich et al., 2020), food security (e.g. Nicholson et al., 2021), agronomic (Rodrigues et al., 2022) and land assessment models (e.g. USDA, 2011). Still, the use of crop yield as benchmark indicator is not without its complications with a review by Klompenburg et al. (2020) finding its use and prediction as one of the most challenging problems within the field of Precision Agriculture (PA). Nevertheless, it is widely recognised that the “crop is the best sensor of its environment” (e.g. Legg and Stafford, 1998). Consequently, crop yield remains an important indicator whose accurate quantification represents the complex relationship between soil, climate, terrain and management.

Crop yield measurement, through the use of precision yield monitors, provides an accurate measure of production performance in an agricultural environment (Lyle et al., 2014). The output from a yield monitor is a high-density point data file with thousands of observations per hectare (Córdoba and Balzarini, 2021). Comparing this resolution to that of a typical land assessment survey a clear disjunction becomes apparent. Land assessment variables, particularly those related to soil properties are collected at a point scale through the use of representative soil pits and augers, at a typical scale of one observation per ten hectares (DAFF, 2018). Therefore, it is unrealistic to assume the scale and associated results of a typical land evaluation survey can account for all inter- and intra-field yield variability. Nor is it reasonable for a land evaluator to conduct a precision level soil survey when undertaking a land release application, terms of Act 70 if 1970. However, the use of technologies associated with PA, such as high-resolution yield should be applied to validate, refine and ultimately improve land evaluation methodologies.

In Chapter 3, land evaluation procedures were verified by linking thousands of individual yield points to land evaluation polygons, across five growing seasons. These polygons were delineated primarily from soil observation points and major terrain breaks. The observed soil points were interpolated (upscaled), either through conventional pedological mapping or spatial interpolation technique within a GIS interface, to ultimately create polygon layers of various land assessment methods. Regardless of what interpolation technique is applied, all upscaling processes, are associated with spatial predication and uncertainty (Phillips & Marks, 1996). The concept of spatial autocorrelation is fundamental in spatial analysis (Getis, 2008) and is based on the premise that geographic elements, which are located closer together are generally more alike than those located farther apart (Sadler et al., 1998). Consequently, yield points located closer to the soil observation would be more correlated than located further away. One of the conclusions, emanating from the analysis undertaken in the previous chapter is that the use of large polygons, delineated primarily from soil and terrain properties, contained significant spatial variability. This variability masked some important trends and drivers, within each resultant land evaluation polygon layer. This conclusion is supported by detailed research undertaken by Hattingh (2018), which found that South Africa soils exhibit significant spatial variability at both the macro- and micro-scale levels because of the interaction of soil and topographical properties. To reduce spatial variability and more accurately verify land assessment methods, their individual components and isolate land evaluation-based productivity drivers, a point-based approach should be applied. The use of a point-based approach will also eliminate errors related to interpolation, whereby only measured locations are utilised to drive the verification process, ultimately increasing verification accuracy. Consequently, the aims of the Chapter are to:

1. Determine a spatially relevant yield buffer for maize and soybean harvests, in order to calculate a representative yield, for each measured soil observation point;
2. Normalise the annual yield of maize and soybean harvests to improve comparability analyses across multiple growing seasons;
3. Assess the performance of pertinent land assessment methodologies using representative maize and soybean yield buffers; and

4. Analyse the performance of individual land assessment factors to inform the development of new productivity-based approaches for land evaluation.

4.2 Materials and Methods

4.2.1 Study area, soil survey, land assessment classification and yield data

The same study site, soil survey points, land assessment classification methods and yield data, as provided in Chapter 3, were used as the basis for this point-based verification study.

4.2.2 Buffer determination and yield extraction

To determine a spatially representative yield, a circular buffer with a radius of 30 m, was initially created around each soil observation point, within ArcGIS 10.5 (ESRI, 2016). In long-term experiments investigating spatial yield relationships, Sadler et al. (1998), concluded that significant differences in yield measurements may occur in distances as short as 10 m. Consequently, this 30 m buffer was considered large enough to incorporate the optimum buffer size, which was expected to fall well within this initial buffer. Processed precision yield points (cf Chapter 3.2.5), across the five growing seasons (2016-2020), for both dryland maize and soybean, were subsequently extracted using this 30 m buffer. The co-ordinates of each yield point, within the buffer, was determined, as well as the distance between each extracted yield point and the soil observation point, which lies at centre of each individual buffer. The yield semivariance, an autocorrelation statistic, was calculated for each yield point using the R Statistical Package (R Core Team, 2013) and defined as (e.g. Robertson, 2008):

$$\gamma(h) = \left[\frac{1}{2N(h)}\right] \Sigma \left[z_i - z_{i+h}\right]^2 \qquad \textit{(Eq. 4-1)}$$

Where:

$\gamma(h)$ = semivariance for interval distance class *h*;

z_i = measured sample value at point *i*;

z_{i+} = measured sample value at point *i*+*h*; and

$N(h)$ = total number of sample couples for the lag interval *h*.

The median value of individual yield points was used to summarise and provide a single, representative yield value for each soil observation. The median value was selected as it provided an improved relationship between yield semivariance and distance across the 30 m buffer, compared to that of the mean. Individual semivariograms, comparing median crop yield semivariance with distance, were developed for each soil observation point. These were subsequently aggregated to develop an average semivariogram for each of the five seasons. Importantly the purpose of the semivariance analysis is not to model at unsampled locations but rather present, using experimental semivariograms, the spatial relationship between crop yield and the soil observation point. Consequently, spatial interpolation via kriging and associated model fitting did not form part of the analysis.

A secondary analysis, for validation purposes, calculated the mean and median yields and associated standard deviations, at varying distance intervals within the 30 m buffer, for both maize and soybean across the five seasons.

After determining the representative yield buffer, through the semivariogram analysis the yield points falling within this smaller buffer were extracted. However, before being used as an input in further analyses, each individual yield buffer was manually examined and extracted buffers exhibiting problems, such as extreme yield variability, missing harvester passes or containing edge effects, were removed.

4.2.3 Yield normalisation

To improve comparability across growing seasons and reduce yield variation caused by differing varieties, planting populations densities and abnormal factors, the representative yield calculated for each observation point was normalised. Normalisation is a commonly used data pre-processing technique which minimises bias, removes outliers and improves the classification performance of predictive models (Singh & Singh, 2020, 2022). Dryland maize and soybean yield data were normalised to obtain Standardised Normal Values (SNV), on an annual basis, using the following formula, as adapted from Ingeli *et al*. (2015):

$$SNV = \frac{(\mu_{1/2} - \bar{x})}{SD}$$

(Eq. 4-2)

Where:

SNV = standardised normal values

$\mu_{1/2}$ = median yield within the soil buffer

$\bar{x}$ = arithmetic mean of the obtained annual yield

SD = standard deviation of the obtained annual yield

SNV, which can also be expressed as a percentage, results in an intuitive yield normalisation where below average yields are given negative values and above average yields positive values.

4.2.4 Assigning land assessment points a representative yield

Rather than upscaling point land classification data into polygons (cf Chapter 3.3.3) this assessment investigates land assessment classification at a point scale. Each relevant soil observation point was classified using each of the five land assessment methods and assigned a representative yield, across the five growing seasons, using a predetermined representative yield buffer.

4.2.5 Individual land evaluation attributes

Land evaluation methodologies use a combination of soil, terrain and climatic attributes to assess land performance, in terms of its requirements and potential use (FAO, 1976). Ultimately, individual factors are combined to develop an overall classification. A one-way ANOVA or where applicable, the non-parametric Kruskal-Wallis test, were employed to determine if any individual land assessment attributes were associated with any significant yield differences across its specified classes. For example the analysis investigated whether maize yield significantly varied across effective rooting depth classes. Based on this analysis, pertinent individual land evaluation factors were compared to maize and soybean yields to highlight important relationships and trends.

4.2.6 Statistical software and methods

The same statistical software and methods (cf Chapter 3.2.6.) used to determine statistical significance between crop yield and various land evaluation classes, were used in for this point-based verification study.

4.3 Results and Discussion

4.3.1 Buffer size optimisation and representative yield extraction

Across the five growing seasons each initial 30 m buffer, surrounding an individual soil observation point, contained an average of 224 yield points for maize and 243 yield points for soybean. Examples of typical yield variation patterns, within the initial 30 m buffer, are provided in Figure 4-1 (a-c).

Figure 4-1 (a) illustrates an ideal yield variation pattern where the observed yield is relatively constant throughout the 30 m buffer and there is no missing data nor abnormalities. In this example any distance, within the buffer could be selected to aggregate yield points and calculate a representative yield.

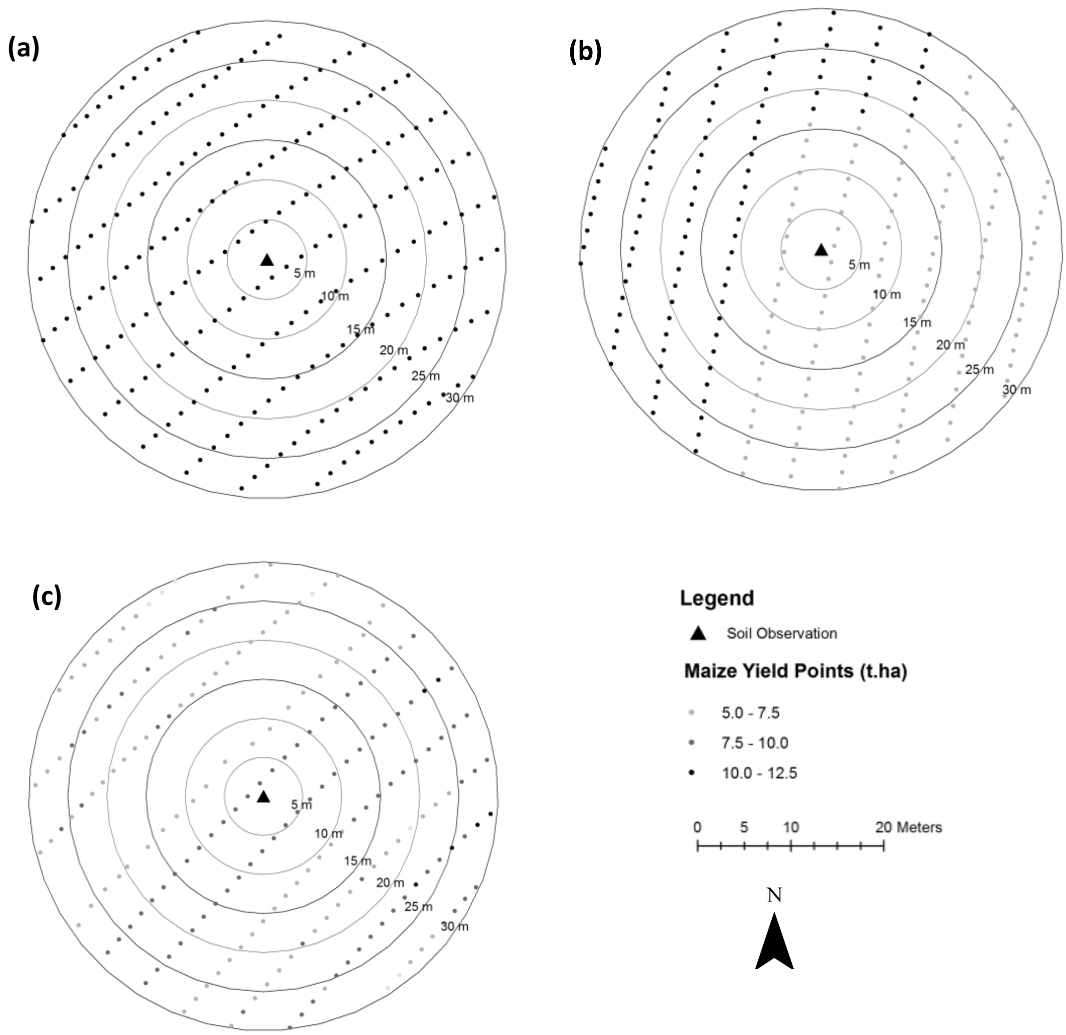

Figure 4-1 Examples of yield variation patterns within a 30 m buffer created around a soil observation point – (a) an example of an ideal yield variation pattern (b) an example of two distinct areas with differing yields (c) an example of significant yield variability and a strip of no yield observations due to a contour bank

Figure 4-1 (b) is example which illustrates two distinct yield areas, within the 30 m buffer, the eastern and southern portions record yields of between 5 and 7.5 t.ha^{-1}, while the western side records consistently higher yields, of above 10 t.ha^{-1}. The soil observation point is located in the lower yielding area and thus all soil and land evaluation properties will correspond to this yield level. The increased yield in the western half cannot be explained by this soil observation point, thus selecting a buffer over 10 m in this example will decrease the correlation between land evaluation properties and yield.

Figure 4-1 (c) illustrates an inconsistent yield pattern, where yields vary considerably across the buffer and yields on the outer edge of the buffer cannot be described by the soil observation point. The initial 30 m buffer also contains a contour bank, which is not planted to crops and is not representative of the soil observation point. If this contour bank was included in the final analysis it would again decrease the correlation between land evaluation properties and yield.

Aggregated experimental semivariograms, comparing observed median crop yield semivariance with distance from a soil observation point, are provided in Figure 4-2. Typically a semivariogram is used to characterise the degree of spatial correlation present in data (Boroumand et al., 2018). In this instance a semivariogram was also applied to determine an optimum yield buffer size, in order to extract a representative yield value for each soil observation point.

Compared to soybean the average semivariance for maize was considerably higher, due to the higher crop yields achieved as well as the greater variation in maize yield values. However, both crops show similar trends with observed yields varying across the 30 m buffer, with the average semivariance increasing rapidly within the first 4 to 6 m of the soil observation point and stabilising thereafter. Over the five growing seasons, the average semivariance consistently levels out at approximately 8 m from the soil observation point for both maize and soybean. This distance, known as the range of the semivariogram, is the maximum distance of spatial autocorrelation (Xiaohu et al., 2016). Essentially this range distance incorporates the maximum yield variability and thus an 8 m buffer, around each soil observation point was used to extract a representative crop yield for both maize and soybean.

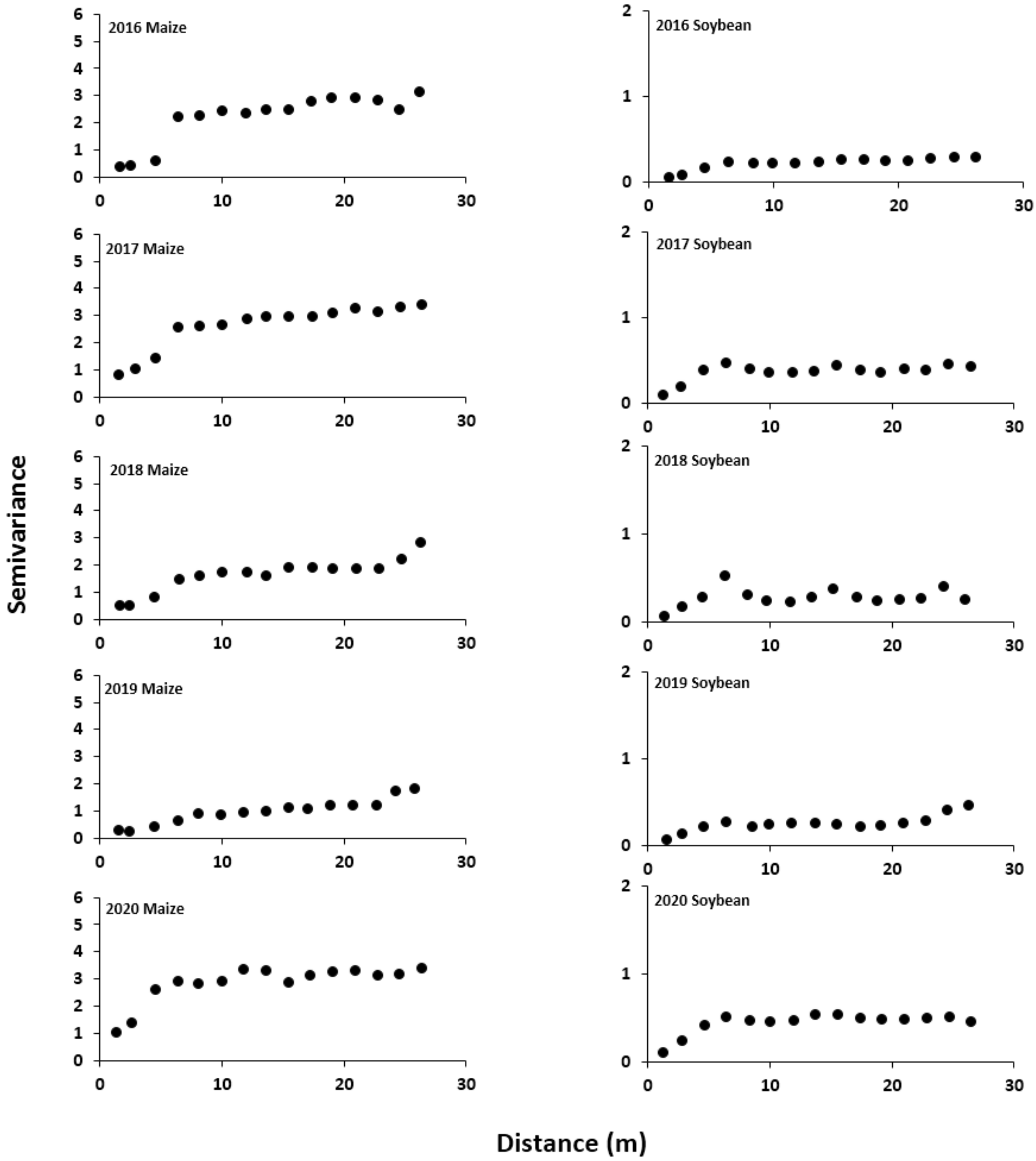

Figure 4-2 Experimental semivariograms of observed median dryland maize and soybean yields across each of the five growing seasons

An example of this 8 m buffer is shown in Figure 4-3. On average, across the five growing seasons, this 8 m buffer includes a total of 17 individual yield points for maize and 18 for soybean. This 8 m buffer was confirmed using an additional analysis, where average median yields and associated standard deviations were calculated at varying distance intervals within the 30 m buffer, for both maize and soybean across the five seasons. The results of this analysis are provided in Appendix A and indicate that median crop yields stabilise at 8 m from a soil observation point while standard deviations consistently increase with distance.

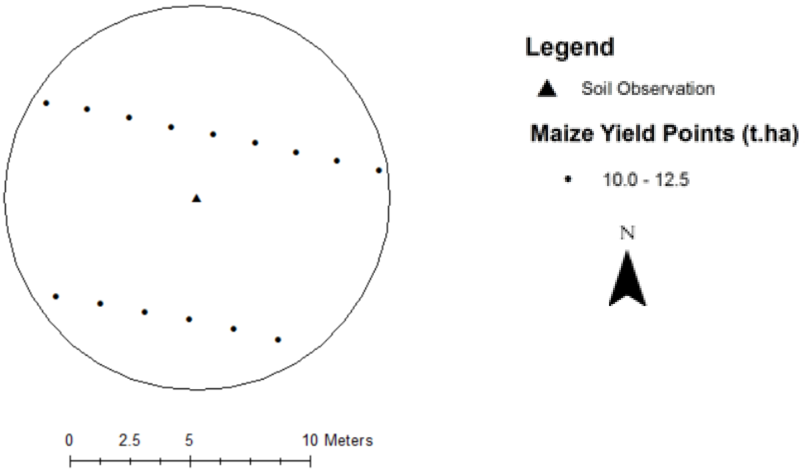

Figure 4-3 An example of an 8 m buffer around a soil observation point with individual yield points

Median maize and soybean yields, around each soil observation point, were extracted from the 8 m buffer and compared to annual farm yields (Figure 4-4). Crop yields extracted from the 8 m buffer are highly correlated to the farm average, calculated from all observed yield points for both maize (R^2 =0.9986) and soybean (R^2 =0.9794). This is an indication that the sampling density, soil observation placement and yield extracted from the 8 m buffer are representative of the yield obtained across the study area. Having a representative sample allows one to conduct secondary analyses and make conclusions that are representative for the population from which the sample is taken (D'Exelle, 2014).

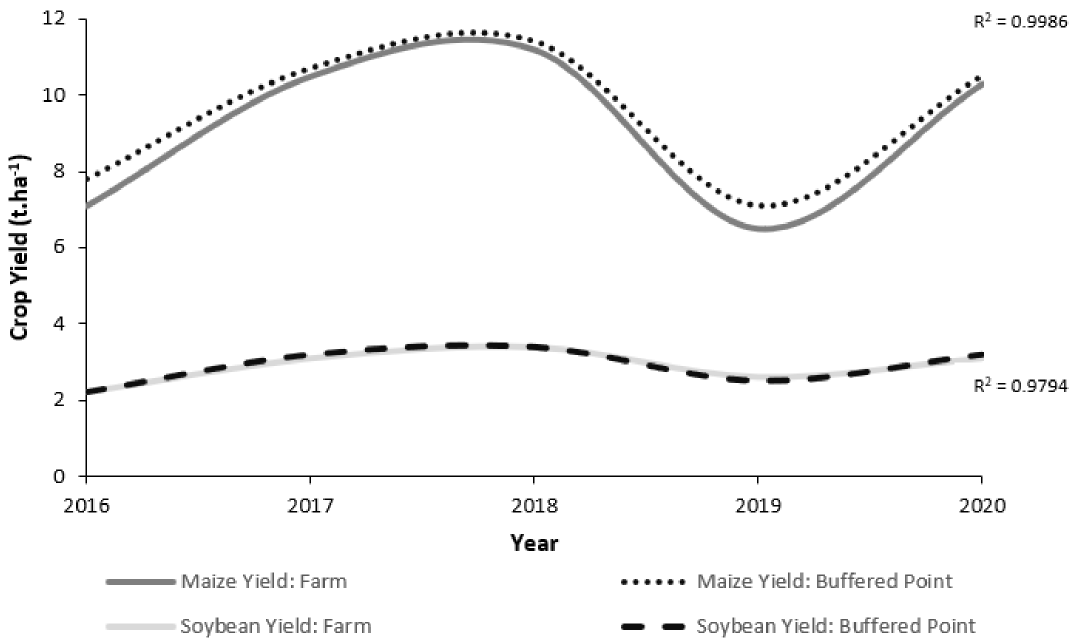

Figure 4-4 Farm and buffered point average yields for maize and soybean (2016-2021)

4.3.2 Yield normalisation

The results of the yield normalisation are provided in Figure 4-5 (a-d) where positive SNV represent an above average yield, while negative SNV represent a below average yield. The difference between SNV and zero indicates the magnitude deviation from the average yield. Each line in Figure 4-5 represents a soil observation point linked to median yield extracted using the 8 m buffer.

The influence of lower rainfall on crop yields is highlighted in Figure 4-5 (a and c) where SNV for both maize and soybean are generally below zero in 2016 and 2019. Both timelines were developed by combining all maize and soybean yields, within the 8 m buffer, across the five years and then normalising the data using the five year average yields and standard deviations. Although yield variation across multiple seasons is clearly observed this approach to yield normalisation has its limitations, where in drier seasons (2016 and 2019) virtually all soils produced below average yields, when compared to the five-year average.

Essentially soils which are traditionally considered to have a high potential would still be linked to negative SNV, a below average yield when compared to the entire five-year period. This limitation is overcome by normalising each year individually; using its annual mean and standard deviation, and then combining each year to create the five-year timeline (Figure 4-5 b and d). Annual normalisation not only provides equal importance to each growing season but also allows soils, which perform well even in low yielding years to be recognised as such. Further, annual normalisation allows for consistently above or below average performers to highlighted, ultimately providing a level playing field for soil and land assessment attribute comparisons across muliple growing seasons. The equal contribution of features within a dataset is important and ultimately improves statsitical processes (Singh & Singh, 2020).

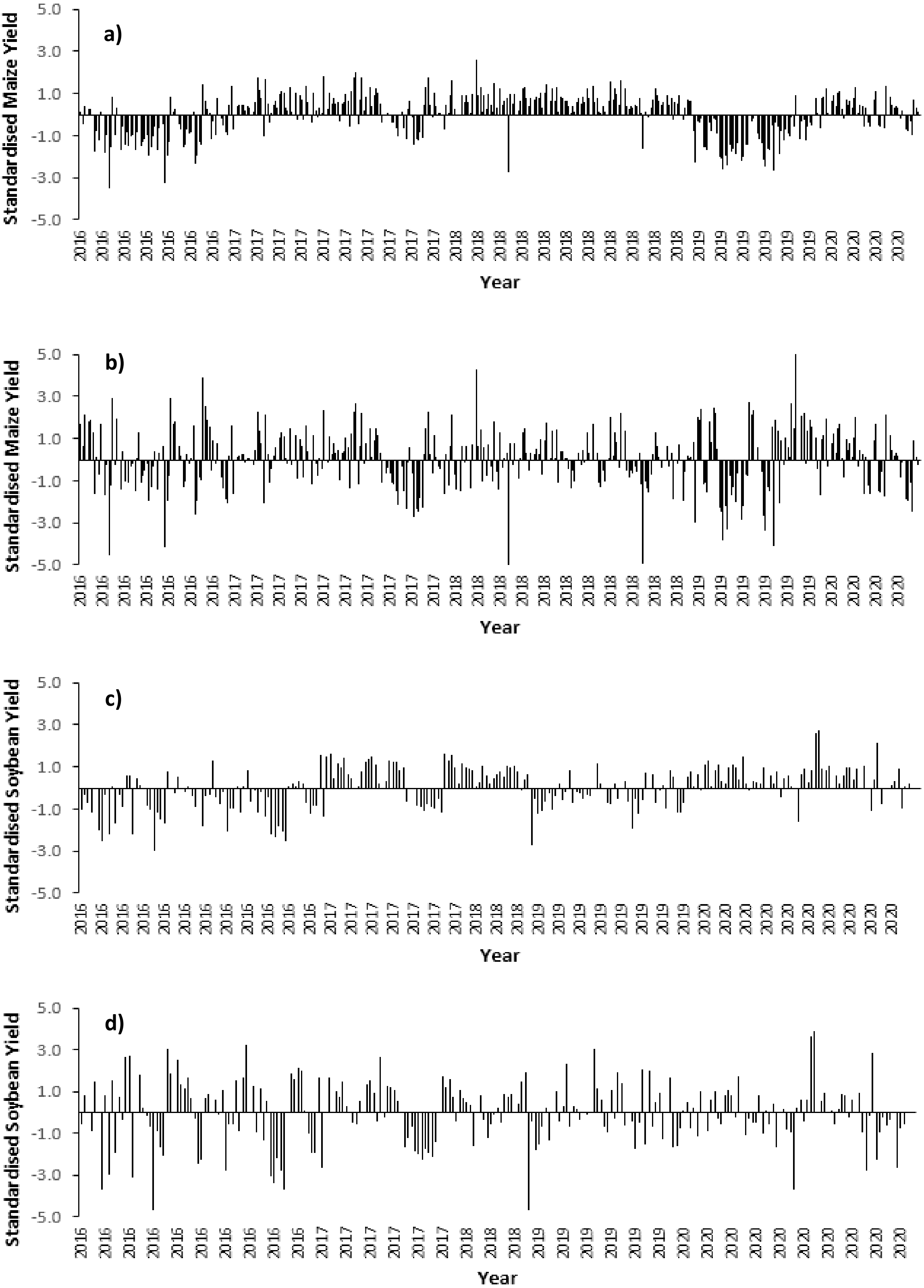

Figure 4-5 SNV of the median crop yield value extracted from all 8 m soil observation buffers between 2016 and 2020 for - (a) combined maize yields normalised using five-year averages (b) combined annual normalised maize yields (c) combined soybean yields normalised using five-year averages (d) combined annual normalised maize yields

4.3.3 Land assessment classification and buffered point yield

Rather than repeat the results and discussion presented in Chapter 3.3.3.4 which compares to productivity to the five selected land assessment methods, this section focuses on the changes caused by buffered point yield extraction and the yield normalisation. For ease of reference all the Tables presented in this section include the polygon average yields and significance lettering as determined in Chapter 3.3.3.4.

4.3.3.1 South African Land Capability Classification

The average dryland yield performance of maize per RSA LC class is provided in Table 4-1. Overall, maize yields decrease with land capability class, which indicates that RSA LC classification is related to maize production. Of the 418-point observations for maize, 83% of these are located on land classified as RSA LC Class III, with this class also producing above average normalised values. Observation points classed Class IV produced below average yields in terms of SNV with an average of -27%, while Class V, deemed non-arable, achieved an average SNV of -97% achieved across the five years.

A one-way ANOVA ($\alpha = 0.05$) found that there a statistically significant difference in median and normalised maize yield between RSA LC Classes III and IV ($p=0.035$). Although no statistically significant difference between arable (III and IV) and non-arable (V) classes could be established, 80% of the observation points located in this capability class contained below average SNV for maize yields. The low number of observed sample points ($n=5$) within RSA LC Class IV may be the cause for the non-significant result. The low sample count reduces the statistical power of the ANOVA, which may reduce the detection of realistic differences (Larson, 2008).

The application of buffered point yield has improved the yield differentials across capability classes, when compared land capability polygons (Chapter 3.3.3.4), increasing from 1.37 $t.ha^{-1}$ for polygons to 2.31 $t.ha^{-1}$ when using a buffered soil points. While the low land capability ratings remain the primary concern, the results indicate that the RSA LC classification can reasonably distinguish between maize production levels.

Similarly, to maize, 78% of all soybean observations were located on land classified as RSA LC Class III (Table 4-2). However, unlike maize yields, soybean yields do not predictably decrease with land capability class. A one-way ANOVA indicates there no significant difference in soybean yield across the RSA LC classes for either average median ($p = 0.727$)

or SNV (%) soybean yields (p = 0.095). The results reiterate the finding, that even though RSA LC is not considered crop specific, RSA LC assessment factors and class breaks may be better suited to maize production compared to that of soybean.

Table 4-1 Average median maize yield, standard deviations and average SNV per South African Land Capability Class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

RSA Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)	Polygon Avg. Yield (t.ha^{-1})
III	345	10.04^{a}	6.67^{a}	9.53^{a}
IV	68	9.24^{b}	-26.74^{b}	9.23^{ab}
V	5	7.73^{ab}	-96.52^{ab}	8.16^{bc}

Table 4-2 Average median soybean yield, standard deviations and average SNV per South African Land Capability Class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

RSA Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)	Polygon Avg. Yield (t.ha^{-1})
III	186	2.79^{a}	6.12^{a}	2.70^{a}
IV	52	2.83^{a}	-24.48^{a}	2.81^{a}
V	2	3.23^{a}	67.21^{a}	2.78^{a}

4.3.3.2 KwaZulu-Natal Land Capability Classification

The average dryland yield performance of maize per KZN LC class is provided in Table 4-3. Both average median and SNV for maize indicate a decrease in yields with increasing KZN LC class, with Class II consistently producing above average yields, across the five growing seasons. Comparing the average point and polygon yield results, shows the yield differentials across the four classes are larger in yields extracted from the buffered points. This suggests the polygons are providing a more generalised average, which is expected due the differences in scale. However, the significance between the KZN LC classes is the same between average buffered point yield, average SNV and average polygon yield, where the average value was significantly different (p=0.016) between KZN LC Class II and IV only. Land capability class V land was excluded from the ANOVA and subsequent post hoc tests as it only included single sample.

The results comparing KZN LC classes to soybean yield is provided in Table 4-4. The ANOVA analyses reveal no significant difference in soybean yield across the KZN LC classes for either average median (p=0.564) or SNV (%) (p=0.270). Although no significant differences could be identified in KZN LC Classes the average SNV (%) does show that normalised soybean yields do decrease consistency with capability class. Class II land was the highest performing in terms of SNV, while land classified as non-arable (Class V) obtained a SNV for soybean yield of -92.18 %. This is an example where normalised yields can improve our understanding of crop performance where average yield cannot.

Table 4-3 Average median maize yield, standard deviations and average SNV per KZN Land Capability Class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column. (+) Indicates the class was not include in ANOVA analysis due to lack of observed samples

KZN Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
II	291	10.12^{a}	7.87^{a}	9.53^{a}
III	44	9.51^{ab}	-5.72^{ab}	9.16^{ab}
IV	82	9.27^{b}	-22.76^{b}	8.60^{bc}
V	1	7.44^{+}	-172.62^{+}	7.70^{bc}

Table 4-4 Average median soybean yield, standard deviations and average SNV per KZN Land Capability Class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column. (+) Indicates the class was not include in ANOVA analysis due to lack of observed samples

KZN Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
II	147	2.76^{a}	7.87^{a}	2.64^{a}
III	34	2.88^{a}	-1.96^{a}	2.65^{a}
IV	58	2.88^{a}	-17.21^{a}	2.76^{a}
V	1	2.69^{+}	-92.18^{+}	2.70^{a}

4.3.3.3 KwaZulu-Natal Land Ecotope Classification

As alluded to in Chapter 3.3.3.4, evaluating the relationship between crop productivity and ecotope class is more complex, due to the high number of classes generated during the land evaluation process. Consequently, only crop ecotopes linked to 10 or more soil observations were used to compare maize and soybean yield variation, over the five growing seasons. However, the yield results for all classified ecotopes are presented in Appendix B.

For maize, only deeper (>800 mm), well drained (Group B) and mottled (Group D) ecotopes produced above average yields in terms of SNV, across the five seasons (Table 4-5). While all poorly drained soils (Group I) and duplex (Group J) ecotopes consistently produced below average maize yields in terms of SNV. A one-way ANOVA analysis was not performed on the average median maize yield as the test of homogeneity of variances was significant, thus the non-parametric Kruskal-Wallis Test was employed to test if the distribution of maize yield is the same across all ecotopes. The Kruskal-Wallis Test found that there is a significant difference (p = 0.023) in the distribution of average median maize yield across classes, with crop ecotope B11, a deep well drained soil with greater than 35% topsoil clay, producing significantly more yield than the other ecotopes. The ANOVA for SNV for maize yields also found a significant difference between individual ecotopes classes (p=0.004), with duplex soils (J23), producing significantly less yields than ecotope B11.

Table 4-5 Average median maize yield, standard deviations and average SNV per Ecotope Class over five growing Seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

KZN Ecotope*	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
B11	23	11.24^{a}	56.61^{a}	10.37^{ab}
B21	107	10.33^{b}	13.43^{a}	9.72^{ab}
B22	18	9.51^{b}	-38.24^{a}	9.83^{ab}
D11	30	9.99^{b}	14.94^{a}	10.30^{ab}
D21	79	10.07^{b}	10.39^{a}	9.70^{ab}
D22	47	9.36^{b}	-19.82^{a}	9.34^{b}
E22	10	8.40^{b}	-55.44^{a}	8.95^{b}
E23	19	9.57^{b}	-10.73^{a}	8.69^{b}
J23	14	8.79^{b}	-61.84^{b}	8.93^{b}

*Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class

For soybean the results indicate that neither average median (p = 0.210) or SNV (p = 0.126) of soybean yield varied significantly across crop ecotopes. However, shallower soils (<800 mm) in soils groups A, B and J all produced below average yields in terms of SNV.

As with the land capability classifications presented above, the use of buffered points has increased the yield differentials across ecotope classes, when compared to polygon derived averages for both maize ($1.01\ t.ha^{-1}$) and soybean ($0.19\ t.ha^{-1}$).

Table 4-6 Average median soybean yield, standard deviations and average SNV per Ecotope Class over five growing Seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

KZN Ecotope*	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
B11	11	2.98 [a]	24.90 [a]	2.58[a]
B21	49	2.84 [a]	27.51 [a]	2.67[a]
B22	17	2.52 [a]	-51.14 [a]	2.43[a]
D11	14	2.82 [a]	9.34 [a]	2.94[a]
D21	27	2.79 [a]	7.65 [a]	2.77[a]
D22	34	2.62 [a]	-12.38 [a]	2.49[a]
E23	10	3.27 [a]	14.96 [a]	2.99[a]
J23	13	2.60 [a]	-30.15 [a]	2.71[a]

*Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class

4.3.3.4 Visual Soil Assessment

The average dryland yield performance of maize per VSA class is provided in Table 4-7. Both median and normalised maize yields decrease with VSA Class, with the “Poor” VSA Class producing 54% less maize yield than average. There was a significant difference (p = 0.011) in median yield with both the “Good” and “Moderate” VSA Classes being significantly higher than the “Poor” class. While for the SNV for maize yield indicates that there is a statistically significant difference between VSA Class Good and Poor (p = 0.02), while there was no statistically significant difference between other VSA Classes.

The average dryland yield performance of soybean per VSA class is provided in Table 4-8. One-way ANOVA analyses reveal no significant difference in soybean yield across the VSA classes for either average median (p=0.358) or SNV (%) soybean yields (p=0.340). Although no significant differences could be identified in VSA Classes the average SNV (%) does show that normalised soybean yields do decrease consistently with VSA. Soils classified as “Good”, in terms of soil quality, were the highest performing in terms of SNV, while soils classified as “Poor” obtained a SNV for soybean yield of -30.16 %.

Overall, although not all yield variation, in terms of SNV, was statistically significant the VSA Classes did produce predictable yield trends for maize and soybean. Further, as previously noted in other methods the yield differentials across the classes are higher where the yields were extracted using the buffered point rather than the more generalised polygon layer.

Table 4-7 Average median maize yield, standard deviations and average SNV per Visual Soil Assessment class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

VSA Class	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
Good	297	10.10[a]	8.43[a]	9.61[a]
Moderate	88	9.67[a]	-8.13[ab]	8.85[b]
Poor	33	8.51[b]	-54.19[b]	8.61[b]

Table 4-8 Average median soybean yield, standard deviations and average SNV per Visual Soil Assessment class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

VSA Class	Buffered Point n	Buffered Point Avg. Median Yield ($t.ha^{-1}$)	Buffered Point Avg. SNV (%)	Polygon Avg. Yield ($t.ha^{-1}$)
Good	173	2.78 [a]	5.21 [a]	2.76[a]
Moderate	53	2.94 [a]	-9.03 [a]	2.83[a]
Poor	14	2.67 [a]	-30.16 [a]	3.02[a]

4.3.3.5 DAFF Digital Land Capability

The average dryland yield performance of maize per DAFF LC class is provided in Table 4-9. Only the soil observations classified as DAFF LC Classes 10 (moderate-high) and 11 (high) obtained positive normalised yield values across the five growing seasons, with 67% of all soil observations falling within the DAFF LC Class 10 (moderate-high). The results indicate that across the seven DAFF LC class only two, 8 (moderate) and 11 (High), produced significant variation (p=0.029) in terms average median yield. Similarly, only classes 5 (low) and 11 (high) were significantly different (p=0.017) when comparing average SNV for maize yields. There was no statistically significant difference across the remaining DAFF LC Classes.

The average dryland yield performance of soybean per DAFF LC class is provided in Table 4-10. Soil observations classified as DAFF LC Classes 5, 9, 10 and 11 obtained positive normalised yield values across the five growing seasons. The above average performance of Class 5 land appears to be a misnomer, which is highlighted by the high standard deviation obtained in this class. One-way ANOVA analyses reveal no significant difference in soybean yield across the DAFF LC classes for either average median (p=0.308) or SNV for soybean yields (p=0.073).

Broadly, the DAFF LC classes were somewhat related to maize and soybean yields with normalised yield values producing more predictable trends, where the highest normalised yield being found in classes 10 (moderate-high) and 11 (high). However, of the five selected land assessment methods the DAFF LC is unique insofar as it is not based on point observations but rather is an existing digital layer, at a 90 m resolution. The purpose of the 8 m buffer was to better relate yield to land assessment factors at a point scale. This is not the case for the DAFF LC dataset, where this yield buffer is still being related to a larger 90 m grid, even more so when the scale of the underlying soil and climate layers are considered. The differing spatial scales between points and grided pixels is highlighted in Atkinson et al. (2010) where more variation exists in the larger spatial element, in this case the DAFF LC digital layer. Atkinson et al. (2010) further advocates that some form of upscaling of the point data should be performed before comparisons are made. Consequently, the utility of the DAFF LC product and supporting spatial layers, at a point scale, is questionable.

Table 4-9 Average median maize yield, standard deviations and average SNV per DAFF Land Capability class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

DAFF Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)	Polygon Avg. Yield (t.ha^{-1})
5 Low	7	8.08^{ab}	-98.42^{b}	9.60^{a}
6 Low - Moderate	2	9.22^{ab}	-76.92^{ab}	9.78^{a}
7 Low - Moderate	11	9.09^{ab}	-28.51^{ab}	8.64^{a}
8 Moderate	18	8.51^{b}	-48.67^{ab}	8.47^{a}
9 Moderate - High	68	9.91^{ab}	-4.77^{ab}	8.72^{a}
10 Moderate - High	278	9.95^{ab}	3.63^{ab}	8.73^{a}
11 High	34	10.61^{a}	39.63^{a}	9.54^{a}

Table 4-10 Average median soybean yield, standard deviations and average SNV per DAFF Land Capability class over five growing seasons (2016-2020). The same letters indicate statistically insignificant differences (p>0.05) per column.

DAFF Land Capability Class	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)	Polygon Avg. Yield (t.ha^{-1})
5 Low	4	2.90^{a}	40.67^{a}	2.67^{a}
6 Low - Moderate	2	2.22^{a}	-88.38^{a}	2.65^{a}
7 Low - Moderate	8	2.21^{a}	-74.85^{a}	2.43^{a}
8 Moderate	11	2.56^{a}	-57.65^{a}	2.63^{a}
9 Moderate - High	49	2.86^{a}	3.33^{a}	2.77^{a}
10 Moderate - High	146	2.84^{a}	4.81^{a}	2.66^{a}
11 High	20	2.82^{a}	19.12^{a}	2.69^{a}

4.3.3.6 Summary

This results section focused on the use of median and normalised yields extracted from representative buffers around soil observation points. Importantly, the application of buffered points and normalised yields do not resolve the problems inherent in the various classification systems, such as severe downgrading, scaling issues and factor aggregation (Chapter 3.3.2 and 3.3.3). Rather, these techniques provide the best opportunity for the methods to be linked directly to crop productivity. Ultimately, the application of a more spatially representative yield resulted in greater yield differentials between classes and less yield generalisation. While the use of SNV for maize and soybean yields also provided greater insight into yield variability and performance across the various land assessment methodologies.

4.3.4 Individual land assessment attributes in a production environment

The use of a more spatially relevant yield buffer was found to improve the link between crop productivity and land classification (cf Chapter 4.3.3). Correspondingly, land evaluation methods need to be disaggregated into their individual components, to not only reduce the impact of factor aggregation but allow for the improved identification and analysis of key yield drivers in this production environment. Consequently, the results presented below drill down and focus on individual land assessment components and investigates their relationship to actual maize and soybean production, across the five growing seasons. Moreover, the utility of pertinent land attributes is discussed within the context of developing new productivity-based approaches, with the view of supplementing existing methodologies.

The significance levels (p) for each land assessment attribute for maize and soybean is provided in Table 4-11. A total of 25 attributes were determined to have significant ($\alpha = 0.05$) yield variation across their specific attribute classes. For maize, 21 factors were identified compared to only 4 for soybean, indicating that even at this detailed level of analysis the factors influencing soybean productivity are far more difficult to determine. A summary of the pertinent land assessment factors is provided in Table 4-12 and further analysed in Sections 4.3.4.1 – 4.3.4.6. It should be noted that some non-significant results were included in these sections to illustrate the impact of various attributes and associated classes on the two different crops.

Table 4-11 Individual land assessment attributes and their associated significance levels of SNV for maize and soybean yields

Method / Dataset	Attribute / Factor	Maize p	Soybean p	Reference
Visual Soil Assessment for Maize and Field Crops	Soil Texture	< 0.001*	0.500	Shepherd et al. (2008); Shepherd (2010)
	Soil Structure	0.020*	0.527	
	Soil Porosity	0.020*	0.530	
	Soil Mottling	0.014*	0.505	
	Soil Colour	0.119	0.522	
	Earthworm Presence	#	#	
	Soil Smell	0.145	n/a	
	Soil Rooting Depth	0.020*	0.031*	
	Soil Ponding	< 0.001*	0.017*	
	Soil Crusting	0.803	0.211+	
	Soil Erosion	0.224	0.500	
DAFF Digital Land Capability	Soil Capability	0.07	0.423	DAFF (2018a)
	Terrain Capability	0.08	0.573	
	Climate Capability	#	#	
RSA Land Capability	Erosion Hazard	0.002*	0.440	Scotney et al. (1991)
	Flood Hazard	0.109	0.010*	
	Effective Soil Depth	0.014*	0.464	
	Soil Texture	0.002*	0.024*	
	Internal Drainage	0.023*	0.525+	
	Mechanical Limitations	0.037*	0.608	
	Other Limitations	0.128	0.252+	
	Climatic Factors	#	#	
KZN Land Capability	Slope Class	0.049*	0.757	Camp et al. (1998); Smith (2006)
	Topsoil Texture	0.038*	0.530	
	Effective Rooting Depth	0.028*	0.363	
	Upper Soil Permeability	< 0.001*	0.108	
	Wetness Limitations	0.133	0.463	
	Soil Crusting	0.803	0.127	
	Rockiness	0.032*	0.218	
KZN Ecotope	Soil Group	< 0.001*	0.632	Camp et al. (1999); Smith (2006)
	Topsoil Clay Content	0.038*	0.530	
	Effective Soil Depth	< 0.001*	0.064	
	Slope	0.049*	0.757	
	Rockiness	0.032*	0.218	

* Denotes a significant variation in yield within that particular attribute (α = 0.05)
\+ Denotes the use of the non-parametric Kruskal-Wallis Test where the homogeneity of variance was violated
\# Denotes were ANOVA or non-parametric tests were not employed due to lack of classes

Table 4-12 A summary of significant land assessment factors

Land Assessment Factor	Description	Common to	Reference
Slope	The percentage of inclination of the land relative to the horizontal as determined by an infield Abney Level. Influences soil erosion potential and arability.	KZN LC, KZN Ecotope, DAFF LC	Camp et al. (1995, 1998); Smith (2006)
Erosion Hazard	A compound land assessment factor incorporating slope, leaching status, soil erodibility and the textural difference between top- and subsoil horizons.	RSA LC	Scotney et al. (1991)
Flood Hazard	A general estimate of both the frequency and duration of flooding events based on the interpretation of soil properties and supplementary evidence.	RSA LC	Scotney et al. (1991)
Effective Soil Depth	Is the depth of soil to which plant roots could potentially penetrate before reaching a barrier to root growth. Barriers to root growth include physical or chemical soil properties. An important dry land soil property effecting moisture supply to crops.	RSA LC, KZN LC, KZN Ecotope VSA	Shepherd et al. (2008); Shepherd (2010); Camp et al. (1998);
Soil Texture	Defines the size of soil particles and refers to the relative proportion between sand, silt and clay. Plays an important role in determining water availability, aeration, drainage, workability and nutrient supply. Is often estimated infield using ball and thread methods.	RSA LC, KZN LC, KZN Ecotope, VSA, DAFF LC	Scotney et al. (1991); Shepherd et al. (2008); Shepherd, (2010)
Soil Rockiness / Mechanical Limitations	The estimated proportion of stones, rocks and bedrock exposures which could influence tillage operations. Rockiness also influences infiltration and the cultivatable soil surface. For RSA LC slope and erosion is combined with rockiness to create the compound mechanical limitation factor.	RSA LC, KZN LC, KZN Ecotope	Scotney et al. (1991); Camp et al. (1995, 1998); Smith, (2006)
Soil Functional Group	Soil profiles were classified using the Taxonomic Soil Classification System (SCWG, 1991) and reclassified to soil functional groups, where soils within a particular group will produce similar yields.	KZN Ecotope	Smith, (2006); Camp et al. (1998);
Soil Structure	A general estimate of soil structure ranging from fine aggregates to course soil clods. Structure influences aeration, infiltration, nutrient supply and root penetration.	VSA	Shepherd et al. (2008); Shepherd, (2010)

Soil Porosity	Is closely linked to soil structural properties, influencing aeration and drainage. Porosity is estimated by visually inspecting the soil aggregates to estimate the size and distribution of soil pores.	VSA	Shepherd et al. (2008); Shepherd, (2010)
Wetness Limitation / Internal Drainage / Surface Ponding	Provides an indication the depth of hydromophic soil properties including the presence mottling and gleying. This factor provides an indication of waterlogging internal drainage characteristics and risk of surface ponding.	RSA LC, KZN LC, KZN Ecotope, VSA	Camp et al., 1998; Shepherd et al., 2008; Shepherd, 2010
Subsoil Permeability	Subsoil permeability provides an indication of the rate of water absorption and movement through the soil. It is measured infield by recording the time it takes for water to be absorbed into either the face of the soil pit or into a clod of soil. This rate is converted into a permeability class ranging from 1 (Impermeable) to 7 (Extremely Rapid).	KZN LC, VSA	Camp et al., 1995; Smith, 1997

4.3.4.1 Terrain Factors: Erosion hazard, slope and flood hazard

Four terrain factors were found to have a significant impact on yield variation across the five seasons. For maize, the erosion hazard factor, sourced from the RSA LC, and slope classes common to both the KZN LC and Ecotope methods were found have significant impact on yield variation (Table 4-11). While for soybean, only the classes associated with the flood hazard factor, were found have a significant relationship to yield.

The erosion hazard factor is a compound land assessment factor, unique to the RSA LC system (Scotney et al., 1991). Unlike, in the KZN LC classification, which uses slope as a direct input, the National system combines slope, leaching status, soil erodibility and the textural difference between the top and subsoil horizons to classify land on its overarching erosion hazard (Table 4-12). Erosion hazard ultimately combines with the flood hazard factor to create the two terrain factor considered in the National classification (Scotney et al., 1991). Jenny (1941) identifies topography (relief) as one of the five factors of soil formation and it has been established that topography influences (micro)climatic and meteorological characteristics, drainage, runoff and spatial distribution of vegetative cover (Florinsky, 2012). Consequently, topography and its associated variability influences the spatial distribution of soil properties (Rabia et al., 2021). Five different erosion hazard classes were identified in the study area, ranging from Class E1, land with low water and/or wind erosion hazard to Class E5, land with high water and wind erosion hazard when cultivated (Figure 4-6).

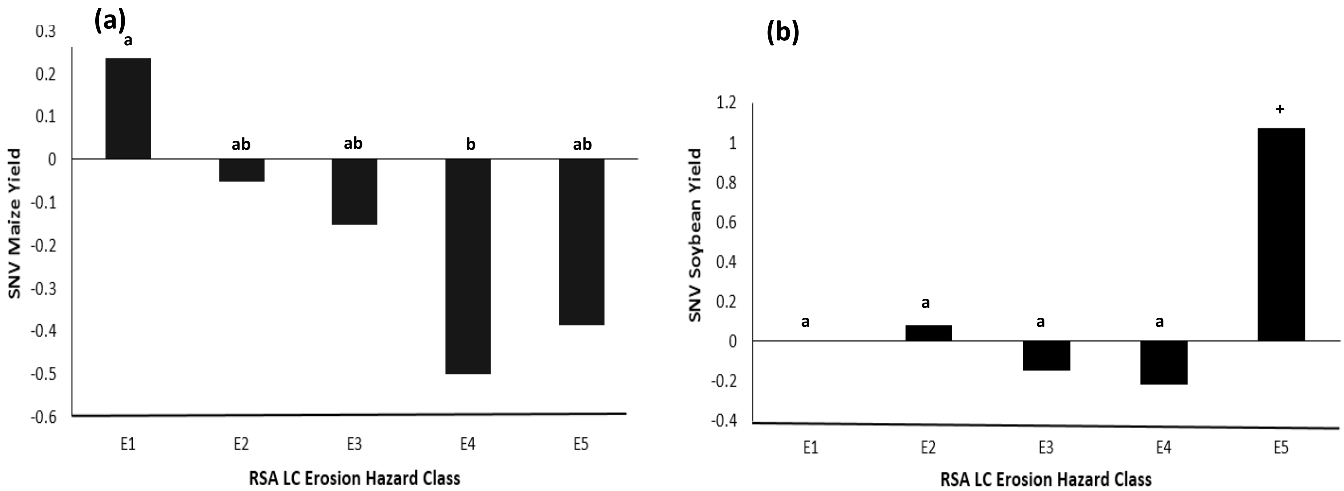

Figure 4-6: SNV Yield and RSA LC Erosion Hazard Classes for (a) Maize and (b) Soybean. The same letters indicate statistically insignificant differences in yield (p>0.05). + Denotes classes excluded from ANOVA due to low sample numbers. Erosion hazard ratings increase in severity from Class E1 (Low Erosion Hazard) through to Class E5 (High Erosion Hazard)

Maize yields (Figure 4-6 a) were, on average, 25% higher in the more favourable erosion hazard class 1, while SNV generally decreased steadily with erosion hazard class, reflecting a negative relationship between productivity and slope and soil erodibility potential. However, these results are not always transferrable to soybean (Figure 4-6 b). Instead, the results illustrate an often-inverse relationship, with soybean yield rather increasing with erosion hazard classes E2 and E5. This suggests that land evaluation factors, in terms of their weighting and importance, are not consistent across different crop types. A study by Leuthold et al. (2022), investigating terrain and yield relationships, similarly concluded that terrain drivers differed for each crop type.

Slope, as a single factor is commonly used as a topographic attribute in land evaluation methodologies, such as in the KZN LC and KZN Ecotope. Within the field of land assessment slope is important when considering soil conservation measures such as terracing and contour banks and plays a critical role in terms mechanisation, where steeper slopes reduce its potential application (Sys, 1985). Slope and associated topographic features has also been found to influence crop yields (e.g. Franz et al., 2020; Guo et al., 2012; Kaspar et al., 2003). The results obtained at FCL farming, comparing slope with maize normalised yield, similarly indicate that slope percentage influences production (Figure 4-7). For both crops the SNV are positive for slopes between 3 and 6% but drop off significantly where slopes approach 12%. Maize, however, exhibits a stronger relationship within moderate slopes, where SNV reached 5% above average yield compared to only 2% for soybean. Whereas soybean yields suffered more than twice the relative yield loss, when grown on steeper gradients, compared to that of maize. The results indicate that steeper gradients (>8%) showed a negative correlation to production across the five growing seasons for both maize and soybean. This observation corresponds to the work of Marques da Silva and Silva (2008) who also observed negative correlations between maize yield and slope, while Leuthold et al. (2022) found that that soybean yield was similarly negatively correlated to slope gradient. It was noted that even though the results for soybean were not statically significant, they do provide important information regarding yield performance which could influence future productivity-based assessment approaches.

The results obtained at FCL also indicate the relationship between slope and yield is not linear, yields do not decrease linearly from flat areas (<3%) to steeper areas (>8%). A trend line fitted by a polynomial function performed significantly better than the linear based function for both maize (Figure 4-7 a) and soybean (Figure 4-7 b). Although these trends may be influenced by ordinal data and associated binning as highlighted by Liddell & Kruschke (2018), the recognition of non-linear production drivers is important for future land assessment systems.

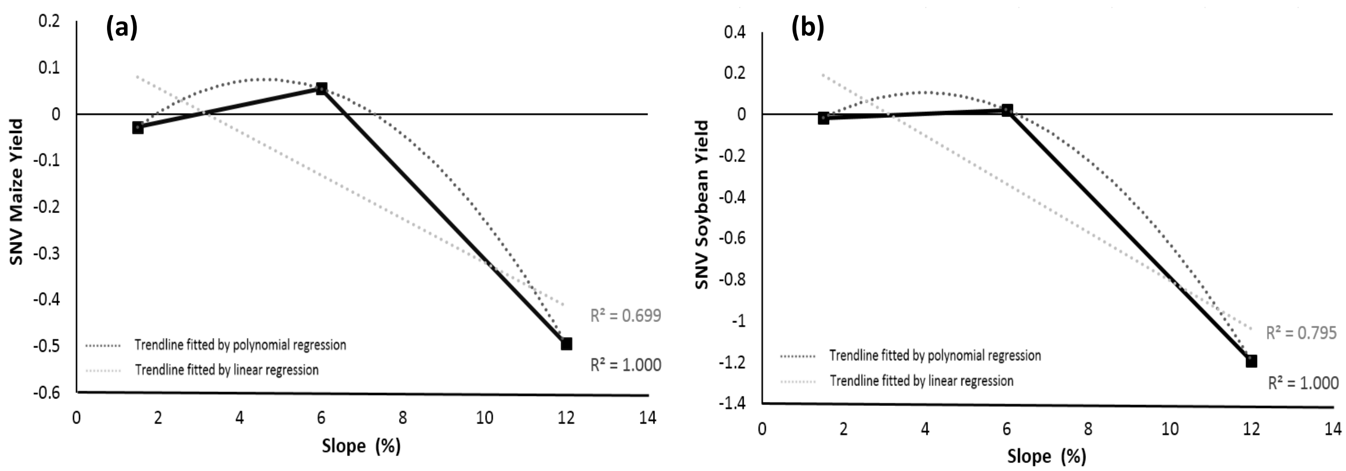

Figure 4-7: SNV Yields and Midpoint of KZN LC Slope Classes for (a) Maize and (b) Soybean

The results comparing slope classes to yield (Figure 4-7) also provide an insight into yield dynamics that is not often considered in local evaluation systems, which predominately uses terrain to determine the level of soil conservation inventions required (e.g. Manson et al., 1995; Scotney et al., 1991; Smith, 1997). For example, in very flat areas (<3%) which is generally considered the best land for arable land use practices (Camp et al., 1998), the yield SNV are in fact negative, indicating below average crop yields for both maize (Figure 4-7 a) and soybean (Figure 4-7 b). In this production environment flat areas are restricted to very particular topographical positions such hilltops and bottomlands. These areas typically contain soils which inherently restrict crop growth, where flat bottomlands are a mix of duplex and poorly drained soils, while flat hilltops are dominated by shallow soils, with parent rock close to the surface. Comparably, Kravchenko and Bullock (2000) reported that various terrain variables, including slope, impacted maize and soybean yield but only when crops were grown in extreme topographical locations, such as depressions or eroded hilltops. Additional studies, comparing maize and soybean yields with slope, could determine if these trends are localised or extended to different locales. Alternatively, the incorporation of more advanced terrain classification could potentially delineate these particular topographical positions.

The results comparing flood hazard class to soybean productivity (Figure 4-8) indicate that yields significantly decrease in areas associated with a high risk of frequent flooding with long durations of water inundation (Scotney et al., 1991). Flood hazard classification in the RSA LC is a broad estimation based on soil properties, vegetation and onsite flooding evidence. However, it is stated in Scotney et al. (1991) that the flood hazard ratings do not provide a high degree of accuracy. Accordingly, this factor could be improved by employing a terrain unit classification (e.g. Jasiewicz & Stepinski, 2013), to more consistently identify flood prone areas.

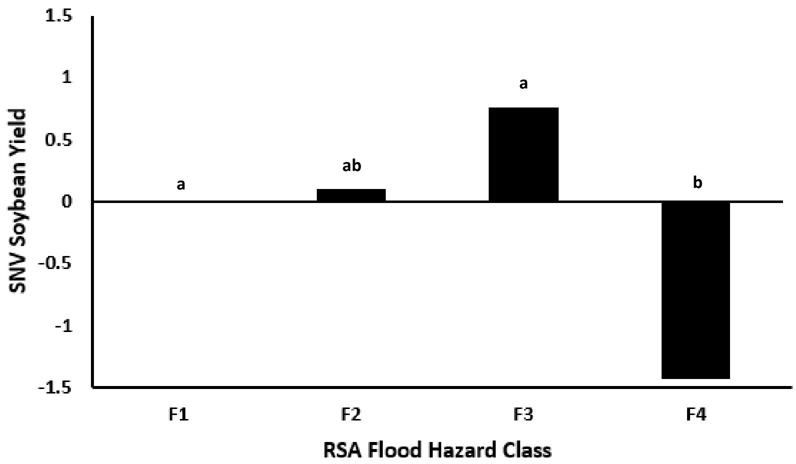

Figure 4-8: SNV Yield and RSA LC Flood Hazard Classes for Soybean. The same letters indicate statistically insignificant differences in yield (p>0.05). Flood hazard ratings increase in severity from Class F1 (No flood hazard) through to Class F4 (Frequent and long flood hazard)

Overall, the results for terrain attributes found that by combining interrelated factors, for example where RSA LC combines slope, leaching status, soil erodibility and textural contrast, novel trends between a particular crops and land evaluation factors can established. Further, slope and terrain related factors can be used to not only determine soil conservation requirements but also crop production potential. Finally, newer technologies such as terrain classification could be employed to improve less reliable factors and improve the link between terrain and production.

4.3.4.2 Effective soil depth, rockiness and mechanical limitations

As with slope gradient, effective soil depth is commonly used in land evaluation methodologies with Le Roux et al. (2013) considering soil depth the most import property of soil bodies, when evaluating land. Due to the nature of agricultural land evaluation, its methodologies are generally limited to a small pool of assessment factors, which tend to be practical and easy to measure during an in-field resource survey. Consequently, methodologies often share common land evaluation attributes, but differ in terms of the number classes and class breaks used in each method. This can be seen in Table 4-11, where effective rooting depth, derived from different methodologies, were found to incorporate significant maize yield variation, indicating that both the method of determination and associated class breaks are important when assessing its relative impact on productivity. Indeed, the ANOVA results indicate that crop yields were significantly impacted by effective soil depth, with the class breaks presented in the VSA methodology providing significant results for both maize and soybean (Table 4-11).

Effective soil depth is the depth to which plant roots can penetrate and determines water and nutrient availability to crops and it greatly influences soil and land capability (Scotney et al., 1991). In turn water availability is recognised as a major determinant of crop yield, when water supply is suboptimal, crop growth and yield is similarly reduced (Singh et al., 2012). Shepherd (2010) similarly recognises that in drier periods, deeper soils allow plant roots to access greater water reserves, thereby reducing water stress in non-irrigated crops. Importantly, effective rooting depth includes both physical and chemical layers which restrict root growth. Common subsoil horizons considered limiting to plant roots included strongly structured cutanic horizons, lithocutanic horizons and hydromorphic horizons.

To assess the impact of effective rooting depth the midpoint of the effective soil depth classes, provided by the VSA and KZN Ecotope methods were compared to SNV of maize and soybean yield across the five growing seasons (Figure 4-9). Additionally, continuous (unclassed) effective soil depths were compared to the same yield data (Figure 4-9 b), to analyse the difference between classed and continuous soil depth data.

The results for maize (Figure 4-9 a) shows a positive, near linear relationship (R^2=0.952) between VSA soil classes depth and yield. The fitted linear trendline estimates that effective soil depths greater than 750 mm will result in an above average maize yield. Shallower effective depths, illustrate a near perfect linear decrease in yield performance, with yields decreasing from -23.14% at 500 mm to -48.64% for 300 mm depths. When the class breaks from the KZN Ecotope method (Figure 4-9 b) are compared maize yield, the linear relationship decreases to R^2=0.758. Indicating that subtle changes in depth breaks can play a large role in determining the significance between individual land assessment factors and maize productivity.

The results for soybean (Figure 4-9 c) also indicates a positive linear relationship (R^2=0.679) between VSA depth classes and yield, however this relationship is not as strong as the corresponding maize data. SNV for soybean approaches above average values (SNV>0) at effective soil depths of 500 mm but then decreases between 500 and 700 mm and finally records above yields at depths greater than 800 mm. The class breaks associated with KZN Ecotope (Figure 4-9 d) show similar results with a positive linear relationship (R^2=0.705) between Ecotope depth classes and yield. Indicating the subtle class breaks between the two methods are not as important for soybean classification in this environment.

The results for both maize and soybean indicate that yields generally increase with effective depth. This corresponds to the general consensus among land evaluation methodologies, that increasing soil depths increase water and nutrient supply and ultimately crop production (Camp et al., 1998; Scotney et al., 1991; Smith, 2006). For both maize and soybean, soil depths greater than 750 mm resulted in above average yields. This depth is similar to the two broad depth categories i.e. deep (>700 mm) and shallow (<700 mm) as provided by Sadras and Calvinõ (2001), whose on-farm research investigated soybean profitability across differing soil depths.

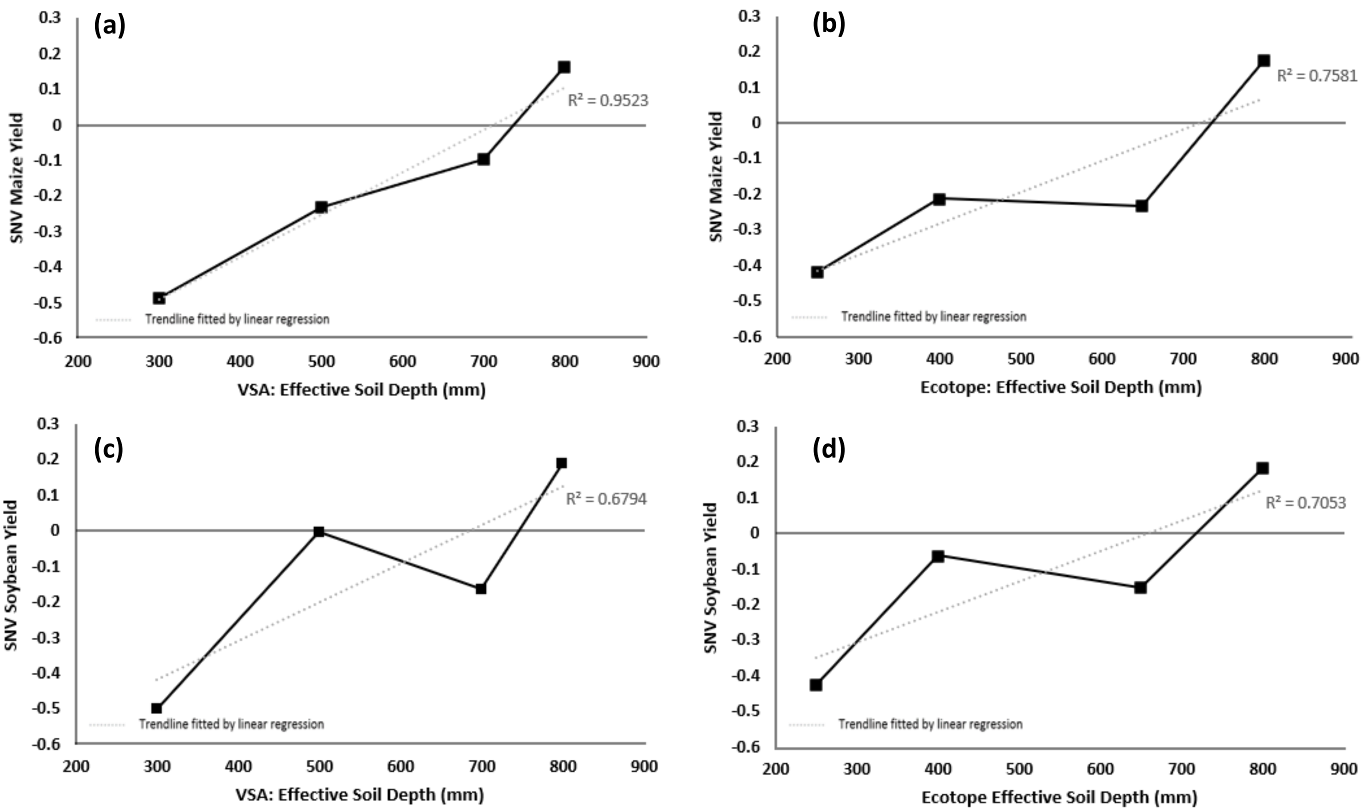

Figure 4-9 Average SNV for yield and effective soil depth for (a) Maize using the midpoint and maximum depth of the KZN Ecotope stipulated depth classes (b) Maize using continuous soil depths (c) Soybean using the midpoint and maximum depth of the VSA Annual Field Crops stipulated depth classes (d) Soybean using continuous soil depths

This research also found the average rate of grain yield reduction, with decreasing soil depth, was largest in maize and smallest in soybean, which was also observed in this study for soil depths between 400 and 600 mm. Locally, the Cedara Agricultural Development Institute (CADI) used two primary soil depth breaks for maize production in KwaZulu-Natal on most soil types, namely at 300 and 700 mm (CADI, 1993). While also stating that high-potential maize soil should allow rooting to 750 mm or more, which are similar to the results observed in this study.

Although the classed data (Figure 4-9) provided a better correlation with maize yields, the continuous data (Figure 4-10 a and b) provides greater insight into yield variability, particularly at shallower soil depths. Spatial variability research by Muller (2004) as cited by Hattingh (2018) also reported that shallower soils increased the inter-annual variability of rainfed yields due to more frequent and severe water stress periods during key agronomic stages for yield formation. Conversely, the high variability observed at deeper depths could be influenced by the low observation count at specific depths (e.g. 1 250 mm).

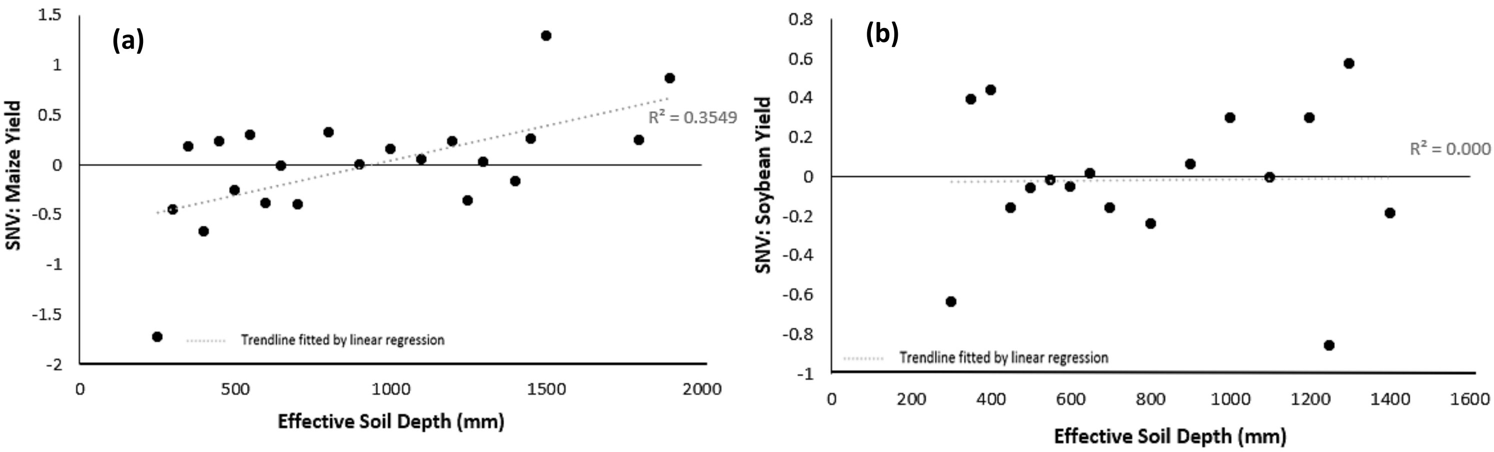

Figure 4-10 Average SNV for yield and effective soil depth for (a) Maize using continuous soil depths and (b) Soybean using continuous soil depths

Numerous research studies have shown that both minimum and optimum rooting depths vary depending on the crop grown (Kirkham et al., 1998; Sadras and Calvinõ, 2001), however many land evaluation methods, particularly land capability such as RSA LC and KZN LC, do not provide crop specific rooting depths. Even the methodology produced for VSA (Shepherd, 2010; Shepherd et al., 2008), which differentiates maize and field crops, uses the same depth classes for each scoring system. The results from this study indicate that maize and soybean crops respond differently to effective soil depths and these differences should be taken cognisance of in crop-specific land evaluation methodologies. The data also highlights importance of class breaks in land evaluation systems.

4.3.4.3 Soil texture and topsoil clay content

An estimation of soil texture is common to all selected land assessment methodologies, and four of them found that soil texture significantly impacts maize yields, while only the soil textural classes, presented in the RSA LC, significantly impacted soybean yields (Table 4-11). Two different textural classifications, from two different land assessment methods were used to compare the influence of soil texture on maize and soybean yield.

For maize the VSA scoring system provided the most significant ANOVA results ($p < 0.001$) and produced four classes with scores ranging from 2 “Good” to 0.5 “Poor” (Figure 4-11 a). The VSA methodology allocates the highest score of 2 to a soil with a silt loam texture, while a soil with clay loam texture score 1.5 (Shepherd, 2010). Soils with either a silt loam or clay loam topsoil texture produced above average yields and although the yields between these classes were not significantly different, clay loams produced the highest maize yields of more than 15% above average. Based on the textural description provided by Shepherd (2010) the clay loam consists of finer material creating a sticker feel which provide more favourable water holding characteristics in this environment, compared to the silt loam. Soils with topsoil textures of either loamy silt or sandy loam scored a 1 within the VSA classification and produced significantly below average yields when compared to soils with a clay loam topsoil. These soils are associated with less than 20% clay and poorer water holding characteristics (Schulze, 1995). Accordingly, crops growing in these soils may be susceptible to water stress, particularly during extended drier periods (Camp et al., 1995). Silty clay and clay topsoil textures were the poorest performers in terms of maize yields (Figure 4-11 a) and are considered to have, as per generic growing guidelines, “*air and moisture regimes that are sub-optimal for maize production*” (du Plessis, 2003). These topsoils generally contain high clay contents (> 55%) and are characterised by poor water permeability, low total porosity, waterlogging and high compaction rates, all reducing maize yield potential (Anikwe, 2000). These unfavourable characteristics combined to significantly depress maize yields (-53%) compared to silt loam and clay loam topsoils.

Notably, compared to the VSA methodology, the textural classes provided in the broader KZN Ecotope classification are far easier to determine in-field, yet still provide a significant ($p=0.038$) result in terms of soil texture and maize yield variability (Table 4-11). Consequently, the ease of measurement should also be taken cognisance of, when developing new production-based approaches.

The topsoil texture classification as used in the RSA LC classification was found to significantly impact soybean yield (Table 4-11). In RSA LC Classification textural groups are first defined by general erodibility / textural grouping on basis of soil form and family (Scotney et al., 1991). Once the soil group has been established texture charts using the proportion of sand, silt and clay is used to define the textural group. Consequently, textural groups can include topsoils with both high clay and high sand contents depending on initial soil group. However, texture group 1 should be seen as the most advantageous for arable agriculture, followed by texture group 2 and finally group 3. Textural groups 1 and 2 were highest performers in terms of soybean yield, with both groups producing above average yields (Figure 4-11 b). In this environment textural group 2 generally corresponded to non-duplex soils with clay contents above 40%. Although the topsoil clay contents were high the yields produced on these topsoil was approximately 20% above average, this aligns with production guidelines as published by DAFF (2010) which states that soybean perform better on heavier textured soils, compared to other crops. Textural class 3, generally high clay topsoils but within the highly erodible soil group, performed significantly poorer than the other two textual classes, greatly suppressing yields.

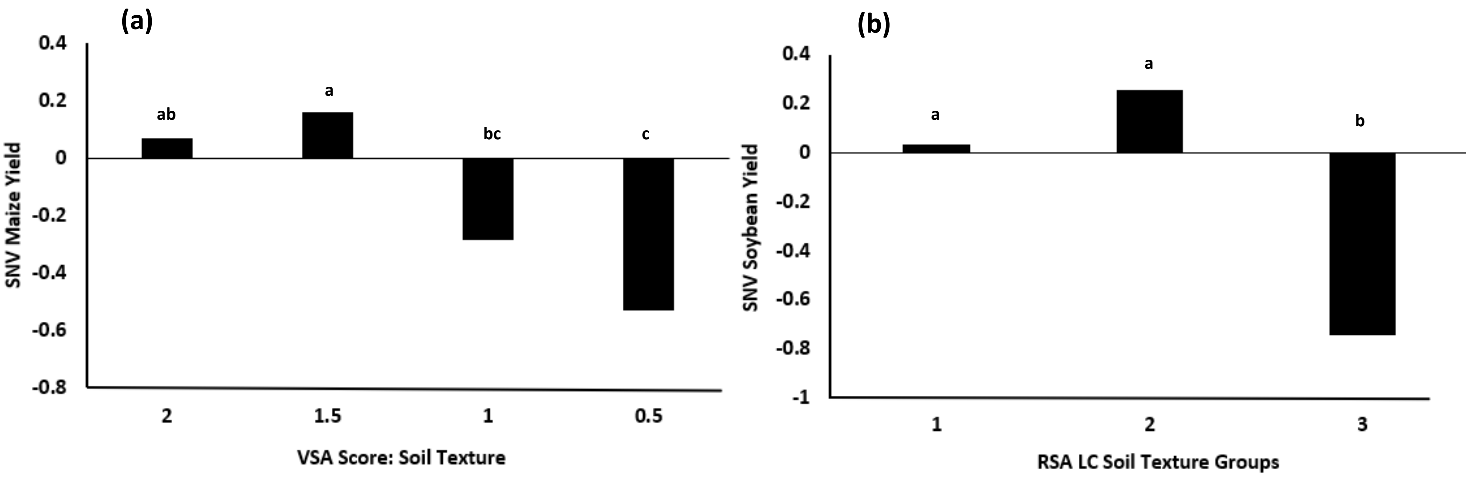

Figure 4-11 Average SNV for yield across various soil texture class for (a) Maize using the VSA Scoring System where 2 is “Good”, 1.5 is “Moderately Good”, 1 is “Moderate” and 0.5 is “Moderately Poor” (b) Soybean using the RSA LC Textural Classes. The same letters indicate statistically insignificant differences in yield ($p>0.05$).

Overall, the results for topsoil texture indicate that maize and soybean yield both vary with texture and that differing classification methods encapsulates this variation better than others. Moreover above average yields were associated with textural classes that provided a balanced water and air regime and suffered in heavily textured topsoils.

4.3.4.4 Soil functional groups, structure and porosity

The one-way ANOVA found significant differences ($p<0.001$) between maize yield between at least two soil functional groups (Table 4-11). Although a statistically significant yield variation, between soybean and soil functional groups, was not found soybean performance were also included in this section to assist in comparative analysis between the two crops.

The soil functional groups, as provided by the KZN Ecotope classification, were compared to SNV of maize and soybean yield across the five growing seasons (Figure 4-12). The KZN

Ecotope classification uses soil functional group as the primary break in its classification and soils within a particular functional group share similar potential and functionality from an agronomic perspective (Camp et al., 1995) for example, Group B incorporating well and moderately drained soils, which under the same management, should produce similar crop yields (Smith, 2006). Local guidelines indicate the most suitable soil for maize is one with "*favourable morphological properties, good internal drainage and an optimal moisture regime*" (du Plessis, 2003). Similarly, Provincial guidelines indicate that soils types producing satisfactory maize yields will usually produce good soybean yields (Department of Agricultural Development, 1990). Crop performance for both maize and soybean are similar for Functional Groups B, D (mottled and moderately drained soils), E (mottled and poorly drained soils) and H (young soils). Functional Group B was the only soil group which produced above average yields for both maize (+13% SNV) and soybean (+11% SNV). More notable differences were observed in soil groups, which are generally considered poor from an agronomic perspective, namely Soil Groups H, Group I (Gleyed Soils) and Group J (Duplex Soils). Maize yields were, substantially more depressed when grown in these soil groups, compared to that of soybean (Figure 4-12). Due to the relative low sample numbers in Groups H and I, only Group J soils were found to be statistically different to the yields achieved in Groups B and D.

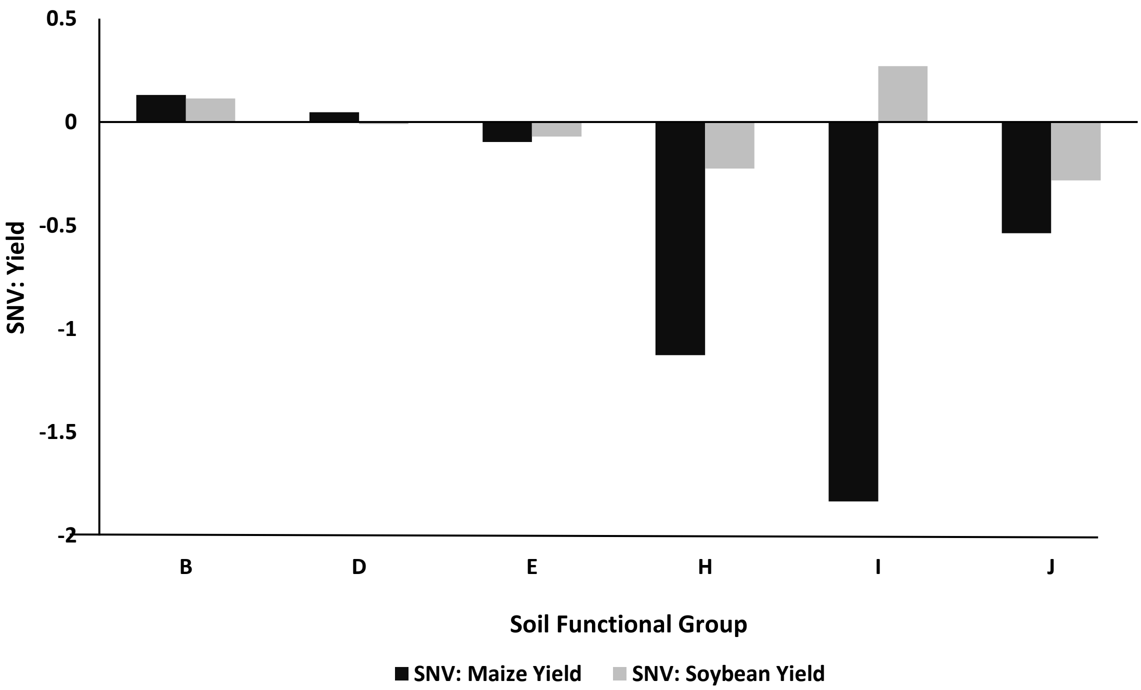

Figure 4-12 SNV yield per soil functional group for maize and soybean, where B is well and moderately drained soils, D is mottled and moderately drained soils, E is mottled and poorly drained soils, H is young soils, I is gleyed soils and J is duplex soils

Soils falling into functional Group I, are dominated by redoxymorphic features, such as mottling and gleying, which develop due to prolonged saturation (DWAF, 2008). Typically, these soils are saturated for significant periods within the wet season and waterlogging, due to poor drainage, is common. Local production guidelines recognise that soils prone to waterlogging result in poor maize yield (CADI, 1993). Soybean is also susceptible to waterlogging but once established is more tolerant to this limitation (DAFF, 2010). Over the five-year analysis period an average of 70% of the seasonal rainfall fell in the second half of the growing season. Suggesting that the soybean crops were already well established prior to significant waterlogging and thus yield losses were minimised.

Observed normalised soybean yields, although below average (-28%), were also less impacted when grown in duplex soils (Group J), compared to that of maize (-54%). Duplex soils are characterised by a marked enrichment of clay in the subsoil (Fey, 2010). In South Africa it is commonly acknowledged that soybeans, due to their combination of a long tap root and shallower lateral roots, are generally better adapted to heavier soils compared to most other crops, including maize (DAFF, 2010).

The one-way ANOVA results also indicate that soil porosity and soil structure both significantly influence maize yields (Table 4-11). It was determined however, that the scoring classification of both these soil attributes were strongly linked to soil functional group. For example for 96% of soils associated with good structure and porosity characteristics, a score of 2 in the VSA (Shepherd et al., 2008; Shepherd, 2010), were found in the well-drained functional group (Group B). Further, 100% of duplex soils (Group J) were associated with poor porosity and structural scores. Consequently, an overarching attribute such soil functional group can be used as an accurate proxy for a number of soil related attributes, reducing the number of attributes that need to be directly measured and scored in-field, saving time and reducing duplication. These potential benefits should be taken cognisance of when developing new productivity-based approaches.

4.3.4.5 Soil Wetness, internal drainage and surface ponding

Soil wetness classification provides an indication of drainage and is common to land assessment methodologies (Table 4-11). The presence of soil wetness indicators, observed surface ponding or internal drainage limitations are all used in land evaluation methodologies to highlight waterlogging risk to agronomic crop production. The soil wetness attribute groups soil profiles based on the depth to and prominence of redoxymorphic features, such as mottling and gleying, which develop due to prolonged saturation (DWAF, 2008). The closer these indicators are to the soil surface the higher risk of waterlogging in normal or above average rainfall years. Waterlogging and associated yield loss is crop dependant, with many field and in particular orchard crops being highly susceptible to phytophthora, root rot (Smith, 2006).

The soil wetness limitations such as degree of mottling and surface ponding provided by the VSA produced the most significant values according to the one-way ANOVA. The presence, abundance and colour of mottles as provided by the VSA classifications, were compared to SNV of maize and soybean yield across the five growing seasons (Figure 4-13 a and b). As with soil functional group soybean yields were not statistically significant correlated to soil mottling but were again included for comparative analysis. Both crops show similar trends, where the highest yields (+10% SNV) were observed where no mottles were present. These soils are well drained, with no signs of wetness within the soil profile.

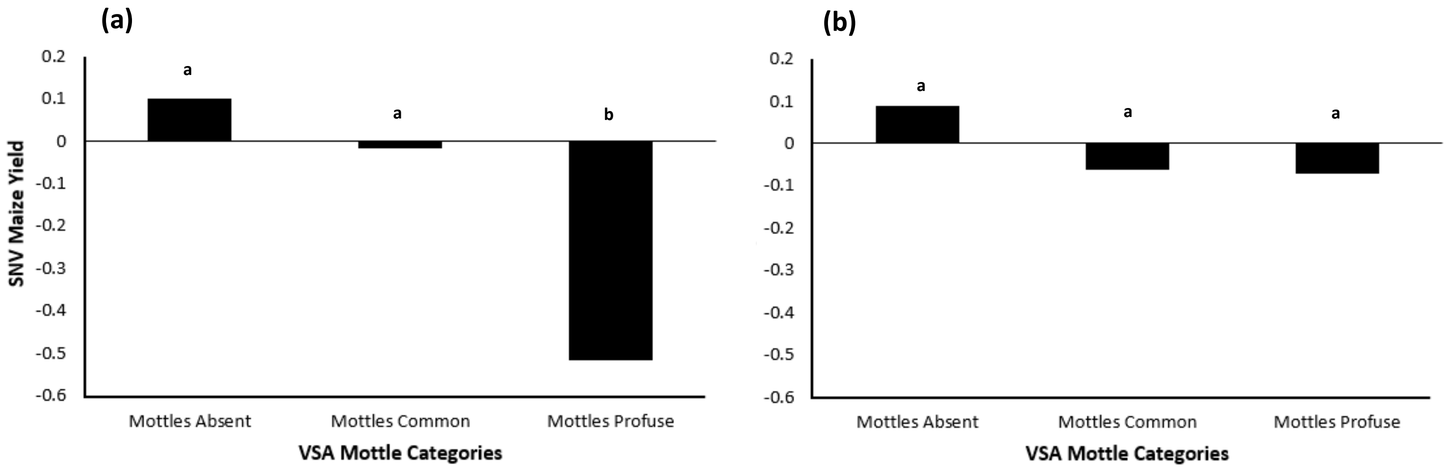

Figure 4-13 SNV Yield and abundance of soil mottles for (a) Maize and (b) Soybean

Yields decrease where the soil profile has many (10-20%) fine and medium orange and grey mottles as described by Shepherd, (2010) and Shepherd et al. (2008). Both maize and soybean yields are close average (SNV=0) in this wetness class. Soils with profuse mottling (>50%) of medium and course orange and particularly grey mottles significantly depressed maize yields, when compared to the drier classes (Figure 4-13). Maize yields produced -52% SNV in this wetness class, while soybean yields only decreased to -7% SNV. Yields are lowest in this class, which include soils that are wet for long periods of the year with mottling and gleying close to the surface. The differing yield responses to drainage limitations reaffirm that land evaluation factors and their associated class breaks are not consistent between maize and soybean and ultimately each crop should be assessed separately.

Similar trends were evident when the degree of surface ponding was compared to maize and soybean yields with yield decreasing with increased surface ponding (Figure 4-14). However, the difference between classes and SNV for maize and soybean was significantly different in the “significant ponding” class. To accurately determine surface ponding, areas need to be visually inspected after heavy rains (Shepherd, 2010). Consequently, the reliable measurement of surface ponding becomes problematic for land assessors, where it may not be practical to visit or survey a farm after heavy rains. However, areas classified as having significant ponding were isolated to certain topographical areas, such as hollows and depressions. Accordingly, the link between production and ponding, could be improved by employing terrain unit classification, to more conveniently and reliably identify areas prone to surface ponding.

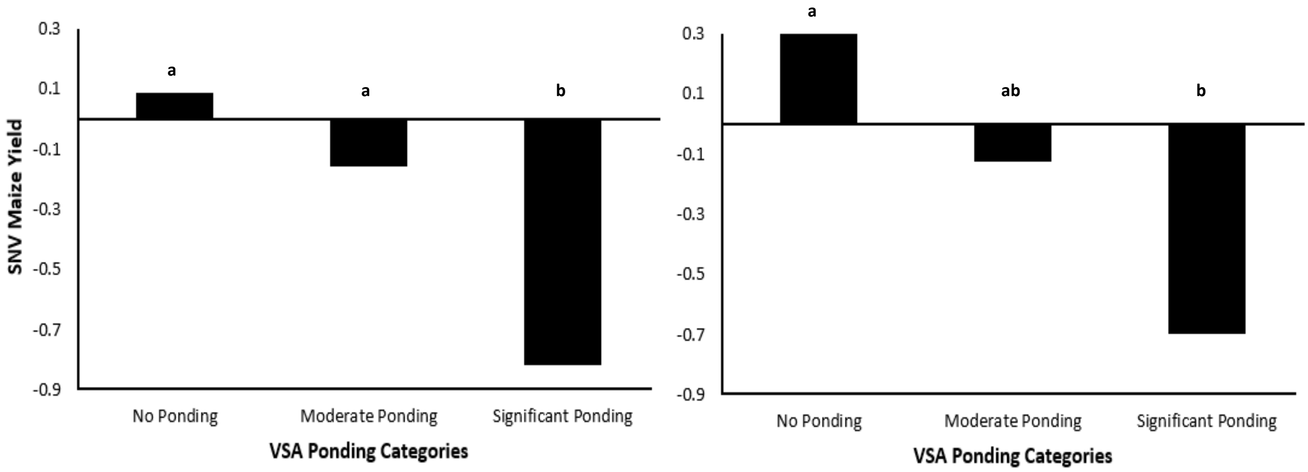

Figure 4-14 SNV Yield and KZN LC Wetness Class for (a) Maize and (b) Soybean

- **A note on soil wetness**

Although it is often seen as a limiting factor in land evaluation process the presence of soil wetness indicators, at sufficient depths can also benefit dry land crop production, particularly in the more arid part of South Africa or during times of drought (Camp et al., 1995). In this landscape notable soil wetness or internal drainage restrictions typically manifest themselves as either a moderately drained soft plinthic (Plinthosols) or poorly drained gleyed (Gleysols) horizons. Soft plinthic horizons, found at sufficient depths and underlying a freely drained subsoil horizon (e.g. Avalon Soil Form) are prized by grain farmers in the Highveld of South Africa as crop roots can tap into the water stored in the lower parts of the profile (Fey, 2010). However, unlike the drier parts of Highveld, which only receive between 500-600 mm of mean annual rainfall, the study area receives over 820 mm (Camp, 1999). Consequently, the benefit of this subsoil “*reservoir*” is not as apparent, even in drier seasons, where only soils without mottles produced above average yields for both maize and soybean (Figure 4-13). This was also confirmed in Chapter 3.3.3.4, where the highest obtained yields in 2016, a drought year, were on well-drained profiles.

4.3.4.6 Subsoil Permeability

Subsoil permeability provides an indication of the rate of water movement through the upper subsoil. Changes in subsoil permeability, as defined in KZN LC, was found to cause significant maize yield variation (Figure 4-15). An ideal permeability lies between "excessively rapid", where drainage is excessive and "impermeable", where drainage is impeded by either rock or very strong structure (Smith, 2006). The study area did not produce permeability timings below 4 seconds, which is considered "rapid" or 1 second which is considered "very rapid".

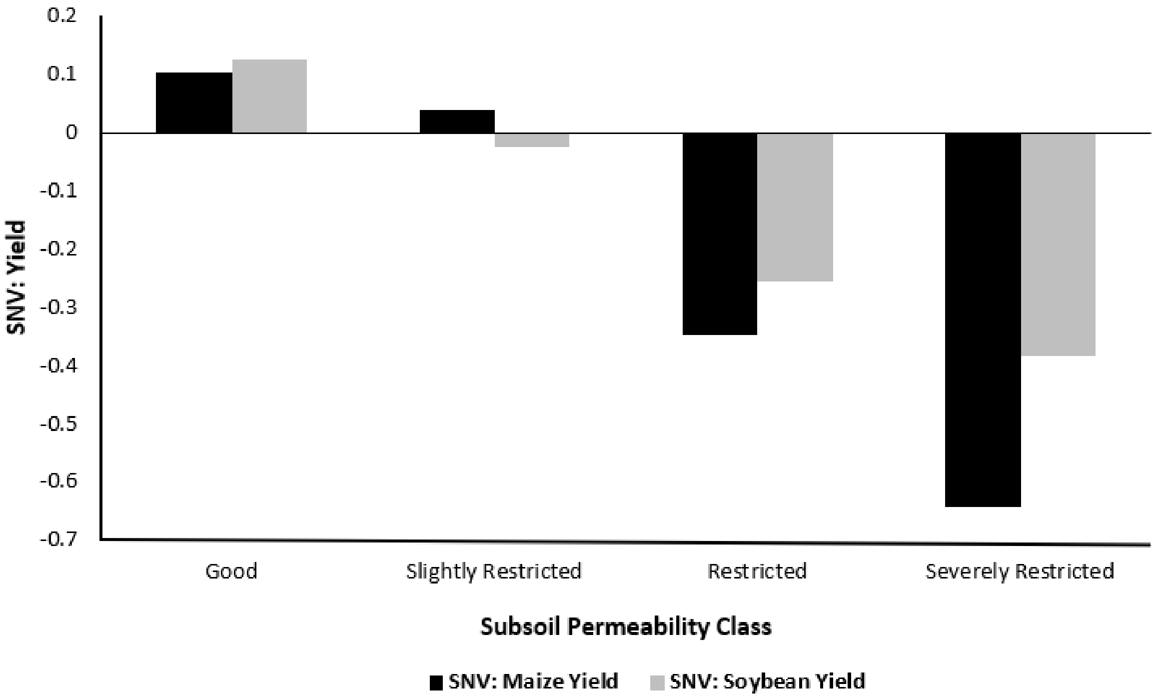

Figure 4-15 SNV Yield and KZN LC Subsoil permeability class for Maize and Soybean

Overall, yields decreased as subsoil permeability rates slowed, as water movement became more restricted by more impermeable subsoil horizons (Figure 4-15). A permeability rate of between 4 to 8 seconds is considered "Good" and produced the highest yields for both maize (+10%) and soybean (+13%). Although both crops follow the same decreasing yield trend with slower permeability rates, only maize showed statistically significant yield variation between the "Good" and the "Restricted" ($p=0.002$) as well as "Severely Restricted" ($p<0.001$) permeability classes. Further, maize yields were more suppressed in the restricted (-35%) and severely restricted (-64%), when compared to soybean, -26% and -38% respectively. This mirrors the results found in the functional soil group and textural comparisons, where duplex soils similarly depressed maize yield compared to that of soybean (cf Sections 4.3.4.3 and 4.3.4.4)

4.3.3.7 Summary

This results section focused on individual land assessment attributes, which were found to significantly impact maize and soybean yields. The results indicate that maize and soybean crop respond differently to individual land assessment attributes and these differences should be taken cognisance of in land evaluation methodologies. Further, the analyses also allowed the impact of key production drivers to be quantified using normalised yield values, across five growing seasons. This method of quantification has potential to be used in attribute importance rating and weighting. Additionally, methodological issues such as compound and holistic attributes, ease of attribute measurement, class break significance, attribute reliability and the potential use of newer technologies such as terrain analysis, were introduced with the of view incorporating these findings into new production-based land evaluation approaches.

4.4 Conclusions

This chapter focused on the verification of land assessment methodologies using a representative yield buffer around each soil observation point. Experimental variograms over five growing seasons were used to determine that an 8 m circular buffer, around each observation point, was suitable for representative yield extraction for both maize and soybean. Yields across the five growing seasons were normalised to create an intuitive classification for both maize and soybean. These representative SNV for maize and soybean yields were used to assess the performance for the five-year analysis period and following conclusion drawn:
1) The analysis reiterates the danger of utilising non-crop specific methodologies, as results differ significantly between maize and soybean. 2) Maize yields had stronger relationship to land assessment methods, compared to that of soybean 3) The highest maize yields generally corresponded to the best land evaluation class or class with highest cropping potential. 4) The application of a yield buffer around a soil observation point improved yield differentials across classes and reduced yield generalisation associated with land assessment polygons. 5) Although a significant difference between soybean and land assessment classes was not observed yield normalisation provided an improved insight into crop performance. 6) Importantly no method could statistically ($\alpha = 0.05$) separate yields across all assessment classes. Further, no land assessment classification could adequately account for soybean yield variation.

Pertinent individual factors used in land assessment were selected and compared to maize and soybean performance across the five growing seasons. Significant yield variation across individual factor classes was more common for maize, compared to that of soybean. A strong relationship between slope and crop performance could be obtained using polynomial-based regression, for both maize and soybean. Maize, however, exhibits a stronger relationship within moderate slopes, where SNV reached 5% above average yield compared to only 2% for soybean. Whereas soybean yields suffered more than twice the relative yield loss, when grown on steeper gradients. The results for terrain attributes also found that by combining interrelated factors, novel trends between a particular crop and land evaluation factors can be established. Further, slope and terrain related factors can be used to not only determine soil conservation requirements but also crop production potential.

A number of soil attributes were found to impact crop performance *inter alia* effective soil depth, soil texture, soil functional group, soil wetness indicators and subsoil permeability. A positive linear relationship was observed for both maize and soybean against the midpoint of VSA and KZN Ecotope soil depth classes. For maize the fitted linear trendline estimates that effective soil depths greater than 750 mm will result in an above average maize yield. SNV for soybean approaches above average values (SNV>0) at effective soil depths of 500 mm but then decreases between 500 and 700 mm and finally records above yields at depths greater than 800 mm. The results for continuous soil depth and yield provided a far poorer linear correlation with significant variation for all recorded soil depths for both maize and soybean. Soil textures with balanced water and air regime produced the highest yields, while heavier textured topsoil significantly suppressed maize yields. In terms of functional soil groups, both maize and soybean yield performance is similar within Groups B (well and moderately drained soils), D (mottled and moderately drained soils), E (Mottled and poorly drained soils) and H (young soils. More notable differences were observed in soil groups, which are generally considered extremely poor from an agronomic perspective, namely Group I (Gleyed Soils) and Group J (Duplex Soils). Maize yields in particular were, substantially more depressed when grown in these soil groups, compared to that of soybean. In terms of soil wetness (internal drainage limitations) both crops show similar trends, where highest yields observed where no mottles were present within 1.5 m of the soil surface. Crop performance was negatively affected as drainage limitations increased, with maize being more sensitive to wetness indicators close to the surface. Similar yield suppression trends were also evident in the subsoil permeability results, where maize yields were more suppressed in the restricted and severely restricted when compared to soybean.

The results of the individual factor analysis indicate that maize and soybean crops respond differently to individual land assessment attributes and these differences should be taken cognisance of in crop-specific land evaluation methodologies. Finally, methodological issues such as compound and holistic attributes, ease of attribute measurement, class break significance, attribute reliability and the potential use of newer technologies such as terrain analysis, were introduced with the of view incorporating these findings into new production-based land evaluation approaches.

5. DEVELOPMENT AND EVALUATION OF NEW PRODUCTIVITY-BASED APPROACHES FOR AGRICULTURAL EVALUATION IN KWAZULU-NATAL, SOUTH AFRICA

5.1 Introduction

The primary objective of land evaluation is to ensure the sustainable management of land, its properties and potential for the direct benefit of society (FAO, 2007). As the human population continues to rise there will be increasing pressure on the agricultural sector to produce more food on an ever-shrinking land base. Thus, it is imperative that land with sufficient production potential is accurately identified, effectively utilised and protected for agricultural use (Collett, 2009). Furthermore, land evaluation needs to shift away from basic classification to approaches that are able to quantify and a predict land productivity performance (Rossiter, 1996).

The first step in effectively evaluating and predicting agricultural productivity performance is to identify and assess the factors that influence it (Viana et al., 2021). A parametric approach to land evaluation can achieve this objective by identifying, scoring and weighting pertinent land characteristics (Sys, 1985). Parametric evaluation systems can also account for interactions, between the selected factors, through addition or multiplication of single-factor indexes (de la Rosa & van Diepen, 2002). Parametric approaches have successfully been used in land evaluation for decades (e.g. Riquier et al., 1970; Storie, 1978) and continue to play an important role to this day, where they are used in soil productivity ratings (e.g. USDA, 2011), cf Section 2.6.2. Importantly parametric methods allow for the integration of both qualitative and quantitative approaches to form hybrid land evaluation systems (Mugiyo et al., 2021). However, the limitations of parametric approaches are well documented with Van Diepen et al. (1991) indicating that many parametric indices are developed with no verification other than expert judgement. Further, Dalal-Clayton & Dent (1993) found that parametric indices are subjective and their scoring logic is difficult is trace. To overcome these limitations yet retain the benefits of an accessible and user-friendly scoring system, the biophysical relationship between crop yield and evaluation attributes were used to drive the selection, scoring and weighting of land assessment attributes. This hybrid approach, hereafter referred to as the *Biophysical Scoring Classification* (BSC), lies between a pure parametric approach and more complex biophysical land evaluation models, as defined by Rossiter (2003) and Bouma (1999).

Traditionally, expert knowledge, empirical models and conventional statistical methods have been applied to improve our understanding of the factors influencing agricultural production, however these methods may not fully incorporate non-linear behaviour or complex interactions (Viana et al., 2021). To address these limitations Machine Learning (ML), a branch of artificial intelligence has recently emerged as an alternative to more conventional methods (Rodriguez-galiano & Chica-rivas, 2014). In the past decade ML based methodologies have been applied across all scientific disciplines, but specifically from a soil science perspective ML is primarily data-driven to develop soil-environmental associations using training samples (Zhang et al., 2021). Ließ et al. (2021) further acknowledges that ML based algorithms are suitable for use in pedometric applications as they are able to derive knowledge and extract soil-landscape relation from complex datasets. Ultimately, ML techniques enable computers the ability to learn without the need for explicit programming and is a particularly effective technique in the fields of data analytics, system modelling and prediction (Taluja & Thakur, 2018).

In Chapter 3, existing systems such as the South African Land Capability System and the regionalised KwaZulu-Natal Land Capability System were found to be effective in delineating potentially arable areas. However, the relationship between land classification and actual productivity, at both a polygon or buffered point scale, has been identified a notable limitation (cf Chapters 3 and 4). These limitations coupled with demand to update assessment methodologies, in order to incorporate newer technologies, has ultimately created a need to develop new, crop specific productivity-based systems to compliment arability assessments. It is therefore envisioned that the new approaches will be used to improve our understanding and classification of arable land classes.

The focus of this Chapter is the development of new productivity-based assessment approaches using a more traditional scoring approach and contemporary ML technologies. These approaches aim to assist professional agricultural assessors, in a particular production environment, to make more informed decisions. With the aim to assist in the decision relating to whether a farm should remain under the protective auspices of Act 70 of 1970 (Republic of South Africa, 1970) or be released to a non-productive land use. Consequently, these approaches need to be geared towards the practical needs of a typical agricultural assessor. As such the proposed methods should consist of soil, terrain and climate attributes that can be rapidly determined during a pedological survey or in the case of terrain layers or climatic data, made available for easy access via digital channels. Assessment attributes that require extensive precision sampling, laboratory analysis or secondary modelling should be avoided.

Consequently, the objectives of the chapter are therefore to:

1. Compile a suite of land evaluation attributes from pertinent methodologies and introduce new attributes based on terrain modelling, soil colour spectrophotometry and infield pedological assessments.
2. Link land evaluation attributes to maize and soybean yields in a specific commercial production environment, across multiple growing seasons.
3. Develop novel, productivity-based land evaluation approaches to broadly predict land performance for maize and soybean.
4. Test the applicability and robustness of these approaches in different locales.
5. Demonstrate the utility of these approaches in broader agricultural land release applications.

5.2 Materials and Methods

5.2.1 Study area for model building and initial testing

The study area used in the model building process incorporates FCL Farming, the same farming area used in Chapters 3 and 4, but also includes additional farm portions owned and managed by Zunckel Farming (Figure 5-1). Both FCL and Zunckel commercial farming operations are located in the same BRU (Wxc5), which indicates that factors such as soil type, climate, altitude, terrain form and vegetation are relatively homogenous across the delineated unit (Camp, 1999). Thus, the broad climatic, geological, terrain and soil patterns as provided in Chapter 3.2.1 for FCL farming are also applicable for study area used in the overarching model building process.

The surveyed area extends from 28º 36’ 08.94” S; 29º 15’ 27.24” E to 28º 42’ 30.12” S; 29º 10’ 46.45” E and covers some 3 638 ha. Both farming enterprises are typical of the region combining dryland and irrigated cultivation of maize and soya, on a three-year rotation, as well as grazing of livestock on both natural veld and improved pastures. Both are intensively managed and considered to be highly productive operations. The reason for including multiple farming enterprises in the model building process was twofold. First, was to reduce the impact of farm specific properties or management factors. Second, it allows for the inclusion of additional soil observations, increasing sample counts particularly in more spatially constrained landscape units and soil types, with hope of creating more robust relationships between yields and land evaluation attributes.

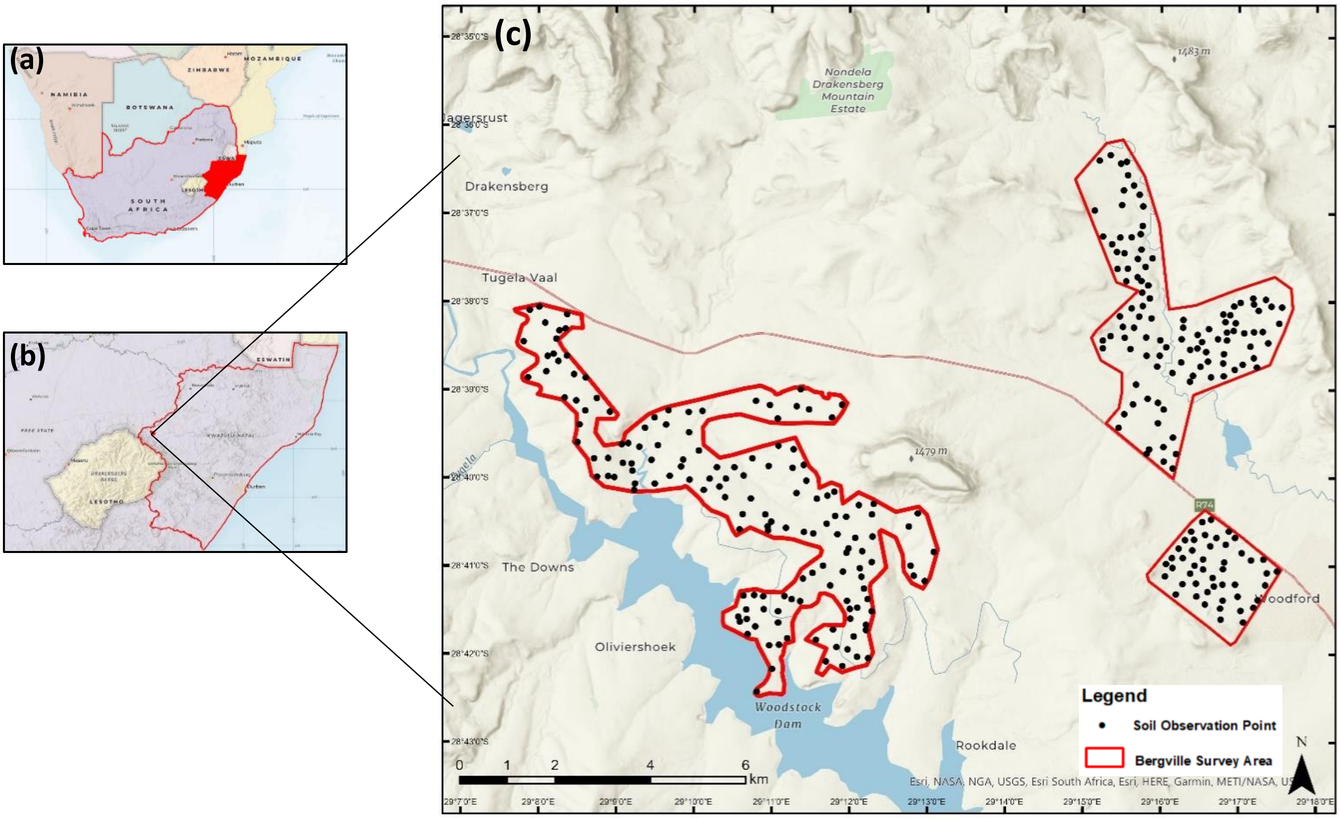

Figure 5-1: Location of the model building study area – (a) Location within southern Africa (b) Location within the Province of KwaZulu-Natal (c) Regional locality map and soil observation points for the Bergville survey area (Background Layers provided by ESRI, 2021)

5.2.2 Study areas: model verification

Three spatially explicit verification areas (Figure 5-2), hereafter referred to as Newcastle, Blood River and Luneburg, were selected to assess the performance of the new land assessment approaches, developed in the Bergville area.

The selection of suitable verification sites was restricted to where chief commercial dryland maize and soybean growing areas in KwaZulu-Natal, directly coincide with the existence of suitable precision yield data. Ultimately the availability of suitable precision yield data in terms of record length, extent and reliability, was identified as the primary limitation in terms of model verification area selection.

The three verification areas are similar to the Bergville farms, in that they are all mixed commercial farming operations, associated with high levels of land management and are considered to be regional benchmark farms in terms actual crop production. However, according to the overarching BRU reports areas (Camp, 1999) each verification area varies in terms of climatic, topographic, broad soil patterns and expected yield (Tables 5-1 and 5-2). This heterogeneity will test the robustness and spatial transferability of the classification and prediction models derived from the Bergville data.

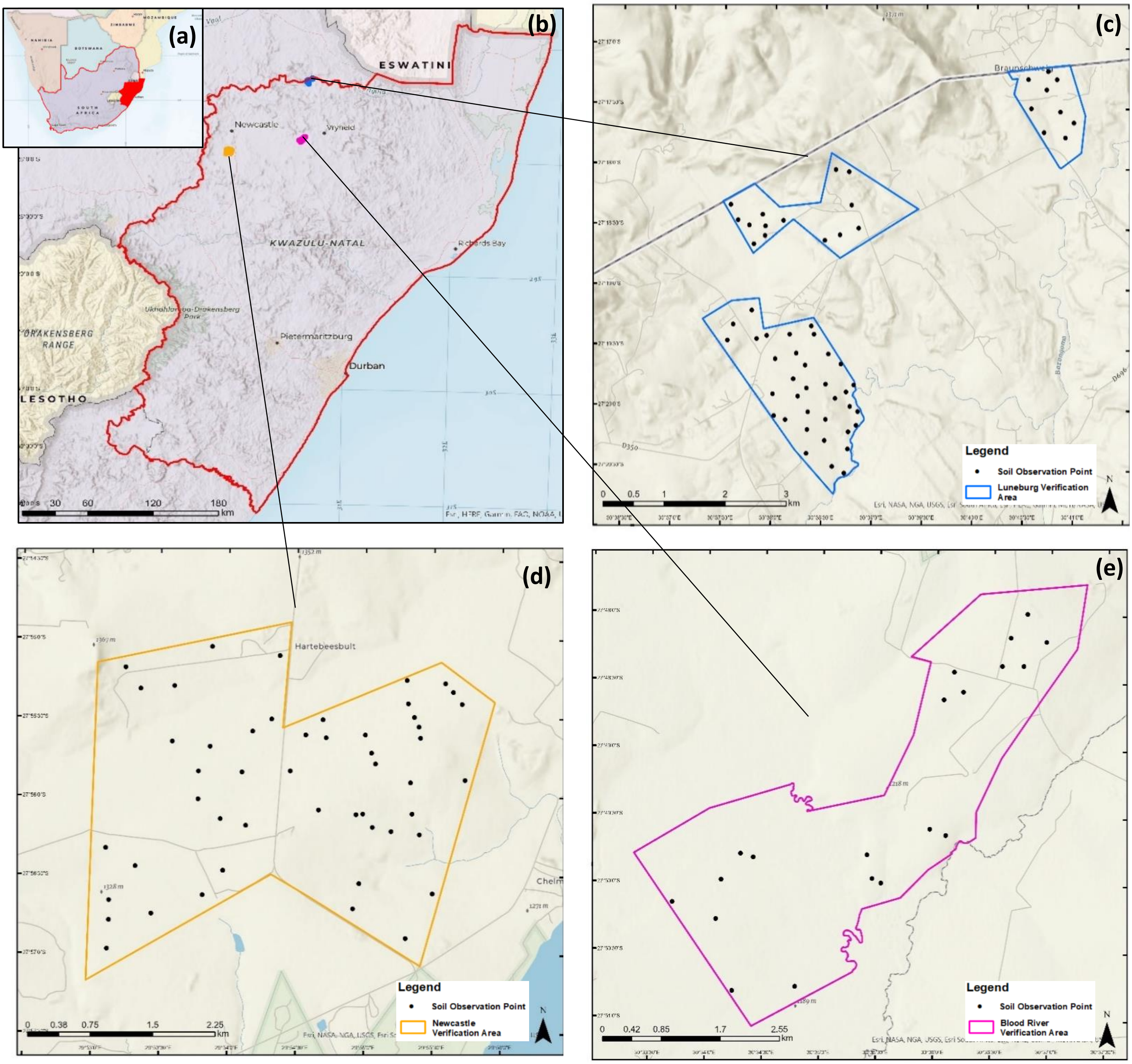

Figure 5-2: Location of the verification areas – (a) Location within southern Africa (b) Location within the Province of KwaZulu-Natal (c) Regional locality map and soil observation points for the Luneburg survey area (d) Regional locality map and soil observation points for the Newcastle survey area (e) Regional locality map and soil observation points for the Blood River survey area (Background Layers provided by ESRI, 2022)

A climatic summary table (Table 5-1) compares rainfall, temperature, and evaporation across the various study areas. The Luneberg verification area (Figure 5-2c) encompasses three non-contiguous farm portions near the KwaZulu-Natal and Mpumalanga Provincial boundary. Compared to the other study areas, the Luneberg verification area is considered to be both cooler and wetter, on average experiencing 1.5 °C less daily average temperatures and receiving nearly 100 mm more mean annual rainfall compared to that of Bergville area (Table 5-1). On average, the Luneberg verification area also receives the lowest heat units of all the study sites. A summary of broad soil, terrain and cropping potential, extracted from the overarching BRU report, is provided in Table 5-2. The Luneberg area has the largest elevation range of all the study areas and is predominately characterised by rolling terrain with slopes of between 5-12%. Camp (1999) considers the Luneberg area "good" for general crop production and predicts highest average crop yields across the four study areas, 6.0 t.ha^{-1} for dryland maize and 3.1 t.ha^{-1} for dryland soybean. Ultimately the Luneberg study area was ranked first in terms of broad production potential of the four study areas (Table 5-2).

Table 5-1 Climate summary for the Bergville model building area and three verification sites near Newcastle, Blood River and Luneburg (Camp, 1999 and Schulze, 1997)

	Unit	Area	Annual	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Median Average Rainfall	**mm**	**Bergville**	696	131	123	100	36	11	3	2	11	23	61	84	111
		Newcastle	698	141	117	85	36	10	1	3	8	27	66	88	116
		Blood River	630	114	87	70	35	12	4	3	8	28	68	91	110
		Luneburg	774	136	109	96	43	13	2	2	7	30	80	116	140
Mean Average Rainfall	**mm**	**Bergville**	824	162	139	98	40	13	9	7	23	49	68	92	124
		Newcastle	843	160	115	82	47	15	11	8	26	39	88	110	142
		Blood River	793	131	106	83	47	21	13	15	17	44	82	105	129
		Luneburg	918	147	123	88	57	23	12	14	18	52	106	132	146
Mean Average Daily Temperature	**°C**	**Bergville**	18.4	22.7	23.1	21.5	18.3	14.8	12.3	12.5	14.6	17.8	19.6	21.3	22.5
		Newcastle	17.5	21.7	21.5	20.3	17.5	14.4	11.7	11.8	14.2	17.4	18.4	20.1	21.3
		Blood River	17.1	20.9	20.7	19.9	17.2	14.4	12.0	12.0	14.0	16.9	17.8	19.1	20.3
		Luneburg	16.6	20.5	20.3	19.5	16.7	14.0	11.4	11.6	13.5	16.1	17.5	18.6	20.0
Heat Units (Base 13)		**Bergville**	167	301	283	264	159	56	0	0	50	144	205	249	295
		Newcastle	143	270	238	226	135	43	0	0	37	132	167	213	257
		Blood River	129	245	216	214	126	43	0	0	31	117	149	183	226
		Luneburg	118	233	204	202	111	31	0	0	16	93	140	168	217
Mean Daily Solar Radiation	**MJ/m^2/day**	**Bergville**	21.0	26.4	25.0	22.6	19.1	15.8	14.1	14.7	16.7	19.9	22.3	25.0	29.8
		Newcastle	18.8	22.9	21.3	20.2	17.5	14.8	13.3	14.0	15.6	18.1	19.5	21.5	26.6
		Blood River	18.4	22.6	21.3	19.9	17.0	14.4	13.1	13.7	15.4	17.8	18.8	20.6	26.2
		Luneburg	18.5	22.4	21.3	19.9	17.2	14.8	13.4	14.0	15.8	18.0	19.0	20.4	26.1
A-Pan Evaporation	**mm**	**Bergville**	1900	216	179	163	125	105	92	103	138	171	190	196	222
		Newcastle	1864	204	171	160	126	107	94	105	141	169	186	189	212
		Blood River	1827	197	166	159	129	109	95	106	142	164	177	179	204
		Luneburg	1862	201	169	160	128	109	96	107	145	169	186	184	208

The Newcastle verification area (Figure 5-2d) is located near Chelmsford dam within the WXc1 BRU. "WX" in the Bioresource codes denotes mean annual rainfall of between 801 and 900 mm, while the “c” denotes an upland area with an altitude of between 901 and 1400 m (Table 5-1). This BioResource unit (BRU) is analogous to the broad rainfall and altitude ranges found in the Bergville area. However, on average the Newcastle verification area is associated with lower daily average temperatures, fewer sunshine hours and heat units compared to the Bergville area (Table 5-1). The broad soil patterns for the overarching BRU indicates that the Newcastle area has the highest proportion of high potential and annual cropping soils of all the study areas and a high potential for crop production (Camp, 1999).

The Blood River verification area (Figure 5-2 e) is located near the town of Vryheid and falls within BioResource Unit Vc4, indicating an upland site within an average mean annual precipitation of between 750 – 800 mm. When compared to the other study areas, the Blood River verification area receives the lowest annual rainfall. The Blood River verification site also experiences, on average, lower daily temperatures and heat units as well as fewer sunshine hours compared to the model building Bergville area. The broad soil patterns also indicate that the Blood River area has the lowest proportion of both high potential and annual cropping soils (Camp, 1999). Consequently, the Blood River study area is ranked last (4) in terms of production potential of the four study areas (Table 5-2).

Rainfall data for the three verification farms was obtained from representative rain gauges, which area maintained and operated by the Agricultural Research Council (ARC). Seasonal rainfall patterns, between 2017 and 2020 across the three verification areas (Figure 5-3) were similar to those observed in Bergville area (Figure 3-10).

Table 5-2 Broad terrain, soil and production potential summary for the Bergville model building area and three verification sites near Newcastle, Blood River and Luneburg (Camp, 1999)

	Bergville	**Newcastle**	**Blood River**	**Luneburg**
BRU	WXc5	WXc1	Vc4	Yc2
Altitude Range (m)	1107-1473	1192-1418	989-1562	961-1738
Terrain	Rolling	Rolling	Rolling	Rolling
Dominant Slope (%)	< 5	< 5	5-12	5-12
High Potential Soils (%)	22.7	48.9	22.2	47.6
Annual Cropping Soils (%)	55.0	62.3	38.6	53.4
Avg. Maize Yield ($t.ha^{-1}$)	5.0	6.0	4.8	6.7
Avg. Soybean Yield ($t.ha^{-1}$)	2.2	2.7	2.2	3.1
BRU Production Potential	Good	High	Good	High
Production Potential Rank	3	2	4	1

Where the 2018 season received above rainfall, 2019 season received below average rainfall while the 2020 was average season in terms of rainfall received. This consistent pattern of seasonal rainfall variation across the model building and verification areas improves the comparability of crop performance across the various regions. This pattern is also taken cognisance of in the annual yield normalisation process, which provides a level playing field for soil and land assessment attributes comparisons across muliple growing seasons (cf Chapter 4.2.2).

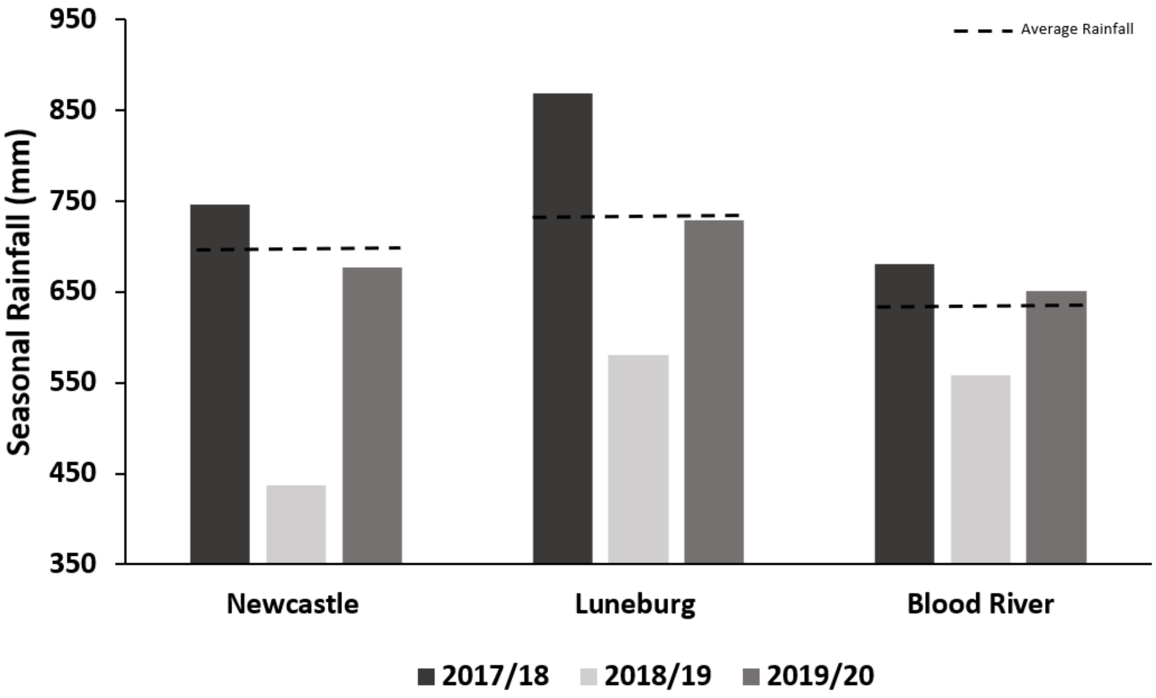

Figure 5-3 Seasonal rainfall (October – March) and average long-term rainfall for the three verification areas across three growing seasons (Schulze, 1997)

5.2.3 Soil and land assessment surveys

For the model building area near Bergville, land assessment surveys and associated soil sampling took place predominantly over the winter of 2016 and 2020 (Figure 5-1). Smaller, supplementary surveys were conducted between 2017 and 2019. The verification areas (Figure 5-2) were surveyed during the winter of 2021 Observation points were located using a purposive sampling approach using expert knowledge, current land use, slope and topographic positions.

Most observations were conducted using soil pits dug to at least 1.5 meters or refusal, while a small proportion were collected by confirmatory auger holes. All observation points were classified using the *Soil Classification, A Taxonomic System for South Africa* (Soil Classification Working Group, 1991).

At each sampling point the following information was collected:

- **General:** Spatial position, land use, crop type, geology, soil sample reference numbers and general comments;
- **Infield Terrain Characteristics:** Terrain position, unit and slope class (via Abney level);
- **Soil Classification:** Horizon name, horizon thickness and colour, soil form and family;
- **Land Assessment Attribute Data:** Total soil depth, effective depth, topsoil clay content, permeability of B1 Horizon, soil structure type and grade, wetness classification, soil crusting, rockiness, soil erosion type and severity as well as method specific land assessment attributes; and
- **Topsoil Samples:** A topsoil sample was extracted for colour spectrophotometry analysis

This information was collected using handheld *Trimble GeoXT* GPS Units with on-board Terrasync Software (www.trimble.com), to record positions and the associated attribute data using customised data dictionaries. All recorded positions were downloaded, differentially corrected, and exported as shapefiles using Trimble GPS Pathfinder Office software.

5.2.4 Precision yield data collection and processing

Seasonal precision yield data was collected from continuous precision yield monitors for both FCL and Zunckel Farming operations, near Bergville. For FCL Farming five years of yield data, from 2015 through to 2020, for both dryland maize and soyabean, were collected, cleaned and processed to provide dry yield mass per hectare. For Zunckel Farming three years of precision yield data for growing seasons 2017-18, 2018-19 and 2019-20 was available and yields for these years were similarly collected and processed. Obtained yields for both operations were aggregated to provide a single model building yield file representing the Bergville production area for the five growing seasons.

As with the model building area in Bergville, seasonal precision yield data was collected for each of the three verification farms located near Bloodriver, Luneberg and Newcastle. For all verification areas three years of precision yield data for growing seasons 2017-18, 2018-19 and 2019-20 was extracted, cleaned and processed.

All cleaned and processed yield data, for both the model building and verification areas, were then extracted using the 8 m circular yield buffer around each soil observation point, as previously determined in Chapter 4 (cf Chapter 4.2.2 and 4.3.1). The annual median yield value for each buffer was calculated and normalised to a provide a Standard Normalised Value (SNV) as outlined in Chapter 4.3.2 and 4.3.2. Finally, a binary yield classification was undertaken where SNV>0 for maize and soybean were classified as above average yields, while SNV<0 were classified as below average.

5.2.5 Assessment attributes

A total of 78 land assessment attributes were collated to analyse their relationship and influence on soybean and maize productivity. A schematic diagram (Figure 5-4) provides an overview of the major attribute subdivisions and how they relate to soil, terrain and climatic factors. Many of the assessment attributes come from the five land assessment methods utilised in Chapters 3 and 4. Other major attribute sources include DEM derived products from the ALOS 30 m model, soil colour spectrometry as well as other infield observations such as soil structure and soil classification information.

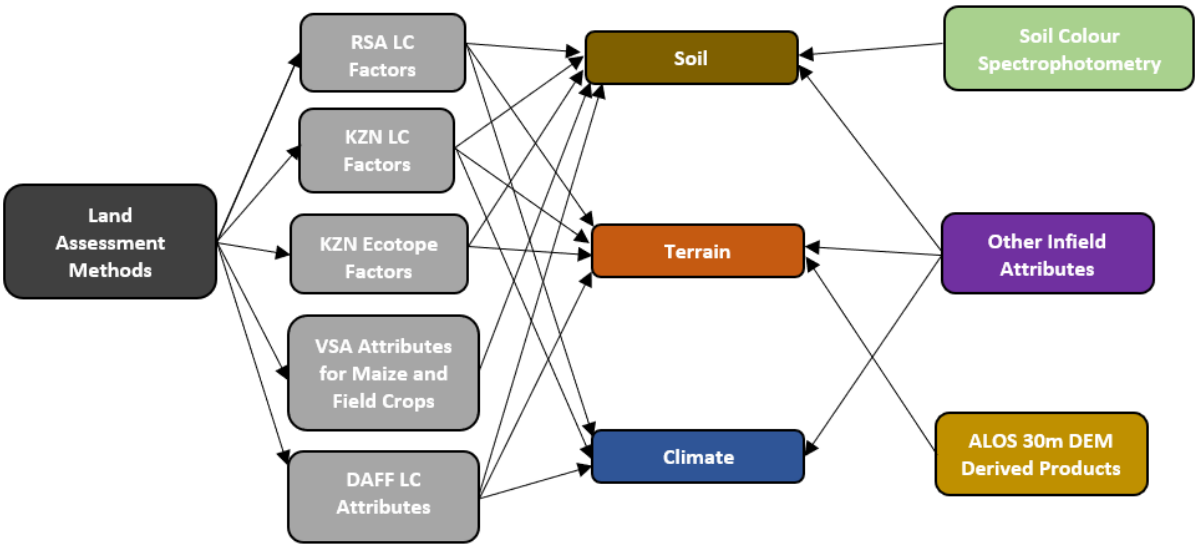

Figure 5-4 Assessment attribute schematic highlighting the various data sources and their relationship to soil, terrain and climate factors

The inclusion of new assessment attributes was based on the findings in Chapter 4.3.4, which compared crop productivity to individual land assessment attributes. For example attributes emanating from digital terrain classification and analysis were introduced to potentially find new or stronger correlations between productivity and terrain attributes.

5.2.5.1 Attributes and factors derived from land assessment methods

A total of 48 attributes were extracted from the Visual Soil Assessment (VSA) methods, the National Department of Agriculture Forestry and Fisheries National Capability Digital Product (DAFF LC), South African Land Capability system (RSA LC), Kwazulu-Natal Land Capability (KZN LC) and KwaZulu-Natal Ecotope Classification (Ecotope). Almost all the attributes extracted from these assessment methods is categorical data and stored in the form of individual scores, tallies or text (Table 5-3). The data collection methodology for the various assessment methods is provided in Chapter 3.2.3.

Table 5-3 Assessment attribute and factors derived from land assessment methods

Method / Dataset	Attribute / Factor	Type	Unit	Reference
Visual Soil Assessment for Maize and Field Crops	Soil Texture	Score	Integer	Shepherd et al. (2008); Shepherd (2010)
	Soil Structure	Score	Integer	
	Soil Porosity	Score	Integer	
	Soil Mottling	Score	Integer	
	Soil Colour	Score	Integer	
	Earthworm Presence	Score	Integer	
	Soil Smell	Score	Integer	
	Soil Rooting Depth	Score	Integer	
	Hardpan Presence	Score	Integer	
	Soil Ponding	Score	Integer	
	Soil Crusting	Score	Integer	
	Soil Erosion	Score	Integer	
	Soil Quality Score	Tally	Integer	
	Soil Quality Description	Class	Text	
DAFF Digital Land Capability	Land Capability	Score	Integer	DAFF (2018a)
	Soil Capability	Score	Integer	
	Terrain Capability	Score	Integer	
	Climate Capability	Score	Integer	
	Land Capability Description	Class	Text	
RSA Land Capability	Erosion Hazard	Class	Integer	Scotney et al. (1991)
	Flood Hazard	Class	Integer	
	Effective Soil Depth	Class	Integer	
	Soil Texture	Class	Integer	
	Soil Erodibility Group	Class	Integer	
	Internal Drainage	Class	Integer	
	Mechanical Limitations	Class	Integer	
	Other Limitations	Class	Integer	
	Climatic Factors	Class	Integer	
	Soil Capability Class	Class	Integer	
	Land Capability Class	Class	Integer	
KZN Land Capability	Slope Class	Class	Text	Camp et al. (1995); Smith (2006)
	Topsoil Texture	Class	Range	
	Effective Rooting Depth	Class	Range	
	Upper Soil Permeability	Class	Integer	
	Wetness Limitations	Class	Text	
	Soil Crusting	Class	Integer	
	Rockiness	Class	Integer	

Method / Dataset	Attribute / Factor	Type	Unit	Reference
KZN Land Capability	Land Class	Class	Integer	Camp et al. (1995); Smith (2006)
	Land Capability	Class	Integer	
	Climate Class	Class	Integer	
	Land Potential	Class	Integer	
KZN Ecotope	Soil Group	Class	Text	Camp et al. (1998); Smith (2006)
	Topsoil Clay Content	Class	Integer	
	Effective Soil Depth	Class	Integer	
	Slope	Class	Text	
	Rockiness	Class	Integer	
	Crop Ecotope	Class	Text	
	Full Ecotope	Class	Text	

5.2.5.2 Soil colour spectrophotometry

Pedology is a branch of science in which the colour is an essential element and assists to both classify and distinguish between specific soil properties (Pegalajar et al., 2020). During field surveys it was noted that topsoil bleaching was often an indicator of subsoil wetness. Similar observations were made by Clarke et al. (2020), where topsoil bleaching was found to be hydrologically driven in summer rainfall areas, such as this one. In the previous Chapter (cf Chapter 4.3.4) it was confirmed that subsoil wetness has influence on dryland yields, consequently it was conjectured that topsoil colour and bleaching status could be an indicator of yield potential.

To determine topsoil soil colour Munsell soil colour charts (Munsell Colour Company, 1975) were utilised by matching soil samples with standardised colour chips. However, this traditional method is commonly associated with inaccuracies (Marqués-mateu et al., 2018). To reduce user error, samples were extracted for spectrophotometry analysis in order to more consistently classify soil colour and determine its impact of maize and soybean yields. Topsoil samples were extracted to determine Munsell Soil Colour, bleaching status and spectral reflectance (Table 5-4) using a *Konica Minolta CM-600d spectrophotometer (Minolta, Osaka, Japan)* and associated methodology as detailed in Clarke et al. (2020). The bleaching status of a topsoil was determined using the E horizon colour specifications, as outlined in Taxonomic Soil Classification System for South Africa (SCWG, 1991), where a soil considered bleached if it has “grey” matrix colours within particular colour range. A lightness index (L_D65_) was also included in the comparative yield analysis.

Table 5-4 A summary of soil colour spectrophotometry attributes

Method / Dataset	Attribute / Factor	Type	Unit	Reference
	Munsell Soil Colour	Class	Text	Munsell Colour Company (1975)
Soil Colour Spectrometry	Bleaching Status	Class	Text	Soil Classification Working Group (1991)
	Soil Colour Lightness (L_D65_)	Index	Integer	Clarke et al. (2020)

5.3.5.3 The digital elevation model and terrain covariates

A Digital Elevation Model (DEM), with a 30 m resolution, was sourced from the Advanced Land Observing Satellite (ALOS) global surface model (Japan Aerospace Exploration Agency, 2021). The ALOS 30 m DEM was selected as it is freely available and presents lower error rates than other respective digital surface models (Nikolakopoulos, 2020). The DEM was used as the base raster to create the various terrain covariates through the use of the System for Automated Geoscientific Analyses (SAGA) software programme (Conrad et al., 2015). The list of selected terrain attributes, their unit of measurement and brief attribute description is provided in Table 5-5. The terrain attributes were selected based on their potential to either directly influence crop performance or pertinent soil properties.

Table 5-5 A summary of terrain attributes derived from the 30 m ALOS DEM

Terrain Attribute	Unit	Description	Reference
Elevation	[m]	Elevation above sea level as set by the ALOS 30 m DEM.	Japan Aerospace Exploration Agency (2021)
Aspect	[deg]	Degrees clockwise direction from North that a slope faces.	Moore et al. (1993)
Aspect	[rads]	Radians clockwise direction from North that a slope faces.	Conrad et al. (2015)
Slope	[%]	The percentage of inclination of the surface relative to the horizontal.	Tesfa et al. (2009)
Slope	[deg]	The angle of inclination of the surface relative to the horizontal.	Conrad et al. (2015)
Topographic Wetness Index	[index]	Steady state wetness index as a function of slope and upstream contributing area per unit width orthogonal to the flow of direction.	Green et al. (2007); Guo et al. (2019)
Relative Slope Position	[index]	The position of an observation point relative to the ridge and valley of a slope, with a value of 0 for the bottom of the valley and 1 for the top of the ridge.	Guo et al. (2019)
Terrain Ruggedness Index	[index]	A quantitative measure of topographic heterogeneity.	Riley et al. (1999)
Terrain Position Index	[index]	Algorithm derived index used to measure topographic slope positions and to automate landform classifications.	De Reu et al. (2013)
Profile Curvature	[m^{-1}]	The curvature of the surface in the direction of maximum slope.	Moore et al. (1993)
Plan Curvature	[m^{-1}]	The curvature of the surface perpendicular to the direction of the maximum slope.	Moore et al. (1993)

Slope Length Factor	[dimensionless]	A dimensionless factor providing an indication of slope length and steepness.	Böhner and Selige (2006)
Geomorphon Unit	[dimensionless]	Terrain classification using a pattern recognition algorithm based which delineates the landscape into the 10 most common landform elements (Figure 5-5).	Conrad et al. (2015); Jasiewicz and Stepinski (2013)
Convergence Index	[index]	An index which represents the agreement of the aspect direction of surrounding cells with the theoretical matrix direction.	Kiss, (2004)
Flow Accumulation Index	[index]	A wetness index as a function of slope and upstream contributing cells.	Marques da Silva and Silva (2008)

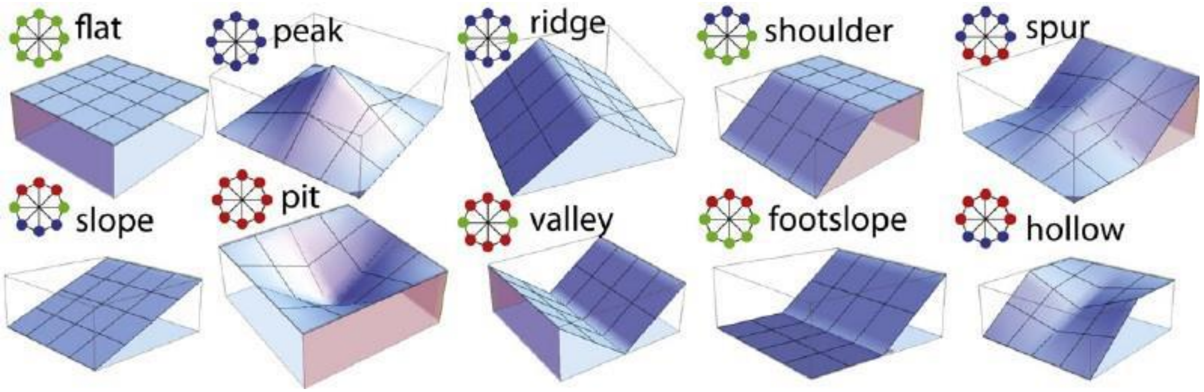

Figure 5-5 The ten most common landform elements classified by geomorphons (Jasiewicz and Stepinski, 2013)

5.2.5.4 Other Infield Attributes

Other infield attributes commonly recorded at each observation point were also included in the list of attributes (Table 5-6). This includes important attributes like unclassed effective and pedological soil depths, soil classification information, soil structural attributes and infield terrain unit classification based on the five-unit slope model (Camp et al., 1995).

Table 5-6 A summary of other infield attributes

Method / Dataset	Attribute / Factor	Type	Unit	Reference
	Seasonal Rainfall	Depth	mm	
	Terrain Unit	Class	Text	
	Topsoil Depth	Depth	mm	
	Soil Form	Class	Text	Soil Classification
	Soil Family	Class	Integer	Working Group
	Pedological Depth	Depth	mm	(1991)
Other Infield	Effective Depth	Depth	mm	
	Topsoil Clay	Range	%	
	Soil Structure	Grade	Text	
	Soil Structure	Size	Text	
	Soil Structure	Type	Text	
	Soil Form and Family	Class	Text	

5.2.5.5 The issue of climate

The influence of climate was taken into account by using climate capability ratings from RSA LC, the broad BioResource Group (BRG) from the KZN LC as well as seasonal rainfall (Figures 3-10 and 5-3). However, due to the nature of the climatic data all observation points, within a particular region, are allocated the same climatic ratings for all years, while rainfall depth only varies across growing seasons. Consequently, climatic factors cannot be used to determine intra-seasonal yield variation. The use of climate as a determinate in land assessment is a contentious issue, with Scotney et al. (1991) indicating a consensus to use climate as a criterion in the RSA LC system was never reached by either the Task Team nor the wider range of specialists during system development and testing.

For this research, the approach of Camp et al. (1998) and Smith, (1997) was adhered to, where climatic factors were not directly incorporated into the assessment methodologies but is rather seen as an overarching factor, which can be used as secondary criterion of overall land potential. Further, these production-based approaches are envisioned to be used as supplementary assessment method to typical land capability assessment where climate capability has already been included to determine broad arability. Consequently, climatic factors would be considered as part of the overarching land evaluation process.

5.2.6 Crop productivity classification and prediction models

This section focuses on the development of three new crop productivity-based approaches for use in land evaluation.

5.2.6.1 Machine learning models

Two ML methods were selected to develop novel maize and soybean models for the Bergville area, namely Random Forest (RF) and Support Vector Machines (SVM). These approaches were selected based on the data available for the model building process as well as the successful application of these methods in soil related research (e.g. Hengl, 2009; Kovačević et al., 2010). The selection is further supported by Padarian et al. (2019), who through analysis of published articles, found that between 2015 and 2020 RF and SVM were the most dominant ML methods used in soil science applications.

The RF approach (Breiman, 2001) is a supervised, ensemble ML technique, which uses decision trees to provide effective predictions for a variety of a applications involving soil properties or classification. RF is a collection of decision trees working together. As a predictive tool it is able to identify empirical relationships between randomly selected covariates and the target variable using internal cross-validation through the use of these multiple decision trees (Makungwe et al., 2021).

The selected SVM approach (Vapnik, 1995) is also a supervised ML methodology commonly used in binary classification and prediction by optimally separating classes by a line, plane or hyperplane (Est´evez et al., 2022). SVM is an algorithm for maximizing a particular mathematical function based on a given set of data (Noble, 2006). This ML technique has used widely in soil science applications *viz.* soil health (Wilhelm et al., 2022), soil quality (Liu et al., 2016), soil classes (Kovačević et al., 2010) and soil properties (Bayat et al., 2020).

The contribution of Mr Stephan van der Westhuizen, from the University of Stellenbosch, is acknowledged for his role in the development and testing of the two ML models. The R Statistical Package (R Core Team, 2013) was used to develop both the RF and SVM models using binary classification, with the goal of predicting above average Standardised Normal Values (SNV) for maize and soybean yields. Of the total observations for Bergville, 592 for maize and 352 for soybean, 80% were used to train the ML models while 20% were used for local testing. Both the RF and SVM models were trained using a 10-fold cross-validation with five repeats each and the developed models were assessed again using a 10-fold cross-validation procedure.

To assess ML models the classification accuracy, specificity, sensitivity and the area under the receiver operating characteristic (ROC) curve (AUC) were used as performance metrics. A ROC is a probability curve with sensitivity (true positive rate) on the y-axis and 1- specificity (false positive rate) on the x-axis, while AUC represents the degree separability from the ROC and is an indication of how well the model is able distinguish between classes (Narkhede, 2018). AUC is a single measure, between 0 and 1, that provides an overarching metric of ML model performance, which exhibits a number of desirable properties and should be used in preference to overall accuracy for the evaluation of ML algorithms (Bradley, 1997). The higher the AUC, the better the model is at distinguishing between observations with above and below average SNV of maize and soybean yield. Generally, the model with the best performance has a curve with the largest AUC value, of between 0.5 and 1.0 (Chen et al., 2018; Hong et al., 2018). An AUC value of less than 0.5 indicates the model is performing worse than a random classifier, while an AUC value of 1 represents a perfect prediction (Yesilnacar & Topal, 2005). Classification accuracy, specificity and sensitivity have been defined in Chapter 3.2.4, as part of the binary land use classification.

5.2.6.2 The biophysical scoring classification

To generate two crop specific classifications, influential land assessment factors were determined, selected and weighted for both maize and soybean. First, each of the 78 individual attributes were assessed using a combination of statistical analyses, to determine their relationship to crop productivity. For categorical data, which forms the majority of the data, the significance of yield variation was assessed using a one-way ANOVA or where applicable, the non-parametric Kruskal-Wallis test, as outlined in Chapter 3.4.4, while the significance of continuous data was assessed using the Pearson Correlation Coefficient. Where duplicate attributes, common to multiple methods (e.g. texture and effective rooting depth) were identified, factors *inter alia* ease of measurement, class break significance, and attribute reliability, as outlined in Chapter 4.3.4, were used to guide final attribute selection. This process is advocated by Sys (1985) where the number contributing factors should be minimised to avoid the duplication of similar factors.

Yield differences between classes, within each selected attribute, were then used to determine its score and subsequent weighting. Classes within significant attributes were generally deemed to impact productivity if the SNV values were approximately 10%, where the SNV value was either positive or negative. A negative SNV indicates that the attribute or particular class is causing a reduction in yields. An example of this is wetness class, where a higher class is an indication of higher probability of waterlogging and reduced yields. Conversely, a positive SNV indicates that a particular attribute is associated with above average yields, for example deeper effective soils depths. Positive and negative scores were then assigned to the selected attributes. While in some instances a zero, or neutral score was assigned to particular classes within a land assessment factor, which in the class itself was not significantly impacting yields. In most cases the classes, within the selected attribute were weighted, wherein multiple positive and negative scores were assigned to classes that considerably impacted yield performance. For example severely restrictive subsoils scored a -2 as they greatly suppressed yields. As a general rule positive and negative SNV values of around 20%, where assigned a double weighting while SNV values of 30% were assigned a triple weighting. Where applicable individual classes, with similar SNV, were merged to create a single scoring class.

An example of the scoring and weighting processes is provided in Figure 5-6. In this example soil crusting was found to be a significant determinate of soybean yield. The classes within the soil crusting attribute were then compared to normalised soybean yields to determine their scoring (+/-) and weighting. In this instance crusting classes t0 (no crusting) and t1 (moderate crusting) were combined and assigned a score of +1 (SNV = +9.3), while t2 (extreme crusting), a negative influence on yield was assigned a score of -2 (SNV = -19.1).

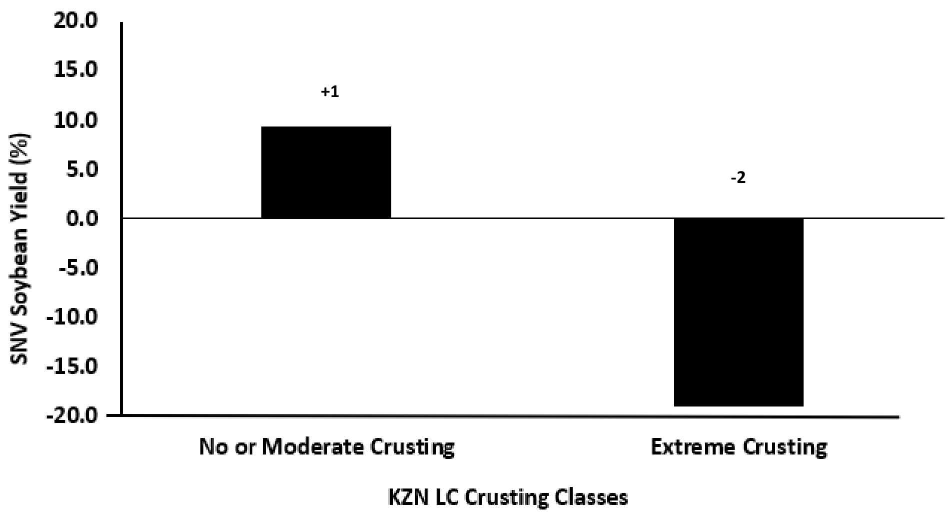

Figure 5-6: Example of scoring methodology using SNV soybean yield and KZN LC Crusting Classes. The numbers above the bars represent the score assigned to each crusting class.

To produce the final scoring classification the scores were tallied for the selected crop specific factors. For both maize and soybean, scores below 1 were classed as negative, below average for crop production, while scores greater than 1 were assigned a positive classification. This crop specific binary classification of either positive or negative scoring performance was compared to above and below SNV yield values for both maize and soybean. To assess the performance of these binary classifiers, typical performance metrics were determined, including classification accuracy, specificity, sensitivity and the Matthews Correlation Coefficient as defined in Chapter 3.2.4.

5.2.7 Statistical Software and Methods

A combination of Microsoft Excel 365, Microsoft Access 365, IBM SPSS (IBM, 2020), Statisica (TIBCO, 2018) and R (R Core Team, 2013) were used to manage and statistically analyse the large precision yield datasets and method performance.

5.3 Results and Discussion

5.3.1 Machine learning results for the Bergville model building area

The 78 land assessment attributes were used as inputs in the RF and SVM ML, with aim of predicting above and below average SNV for maize and soybean yields in the Bergville model building area. A total of 80% of maize and soybean observations were used to develop the models while the remaining 20% were retained to test the model performance. The result of this model testing process, for the various performance metrics, is provided in Table 5-7, which are averaged results of the 10-fold cross-validation for both RF and SVM. Individual ROC curves were created for both models and every fold for both maize and soybean, creating a total of 40 ROC curves from which average ROC and AUC was calculated. Figure 5-7 provides a sample of 5 of the ROC curves used in this process for both RF and SVM.

Table 5-7 Machine learning results for the Bergville area for maize and soybean using Random Forest (RF) and Support Vector Machines (SNV)

Crop	Model	Accuracy	Sensitivity	Specificity	AUC
Maize	RF	57.2	73.6	39.9	0.67
Maize	SVM	56.7	71.3	41.1	0.64
Soybean	RF	65.3	69.0	61.2	0.76
Soybean	SM	63.0	64.4	62.0	0.74

The two different ML approaches provided differing results for maize and soybean yields (Table 5-8). Both ML approaches were able to predict soybean yields more accurately than maize. RF obtained the highest classification accuracy of 65.3% for soybean, compared to 57.3% for maize. Both ML methods consistently scored higher in specificity than sensitivity, indicating that the models are better predicting above average yields for maize and soybean. AUC, which provides general predictive ability of the model, indicates that the RF model is superior to SVM for both crops. The RF model achieved an AUC of 0.67 maize and 0.76 soybean, indicating soybean yield performance was easier to predict than maize.

One of the outputs from the RF model is a list of the top features selected in the model building process in the Bergville area. For maize (Table 5-8) the flood hazard factor from the RSA LC method was found to be the most important attribute. Of the top ten features, four were sourced from ALOS DEM covariates including elevation, slope, terrain position and flow accumulation. The selection of these terrain attributes along with the flood hazard factor indicates that low-lying areas with significant water accumulation influence yield performance.

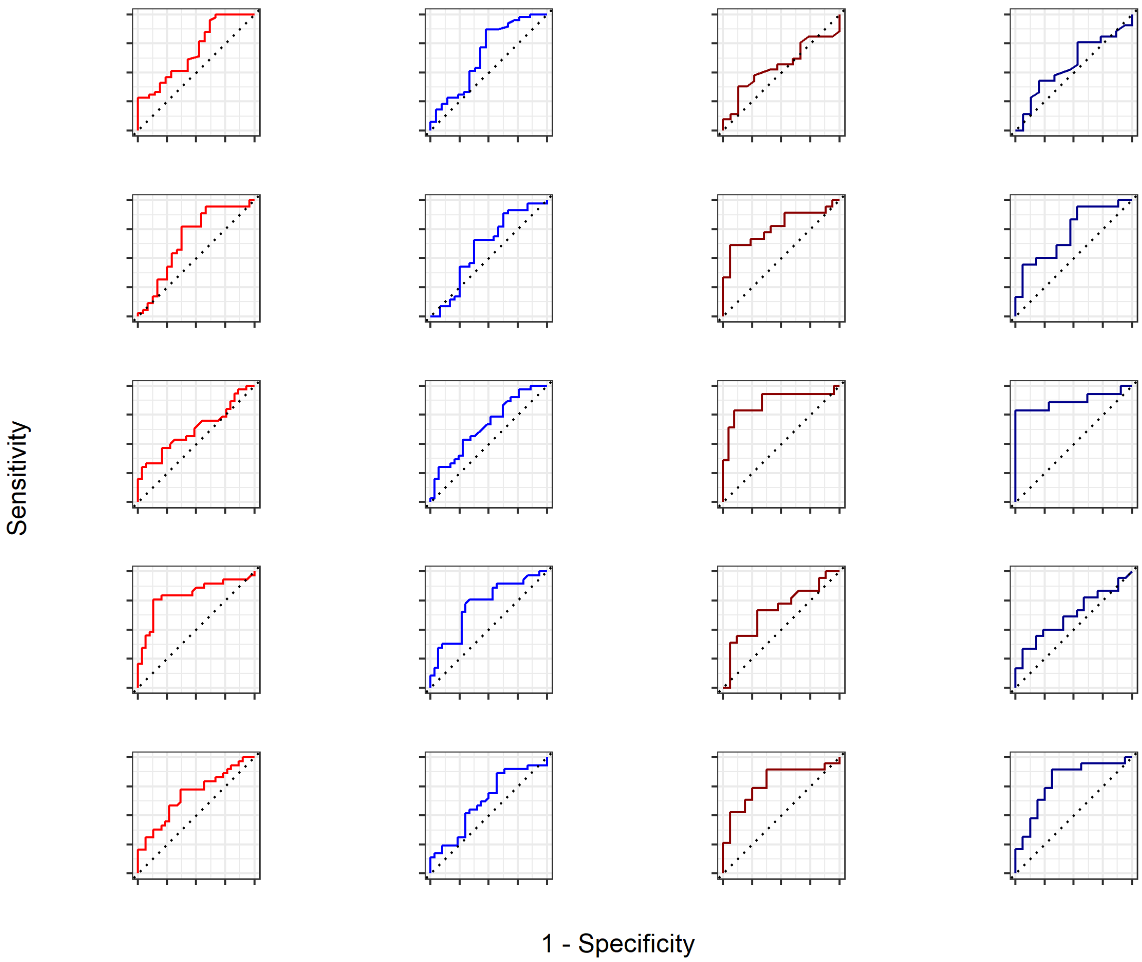

Figure 5-7 A sample of five individual ROC curves from the cross-validation process for (left to right) RF Maize (Red), SVM Maize (Blue), RF Soybean (Dark Red) and SVM Soybean (dark blue)

The top features for soybean RF model (Table 5-9) differ significantly from the selected maize features. Only two attributes are shared between maize and soybean namely elevation and unclassed effective rooting depth, indicating that crop specific yield drivers differ and evaluation methods using universal attributes for all crops will tend to be less accurate (e.g. Camp et al., 1998).

Table 5-8 Top features selected by the RF model for Maize production in the Bergville area

Rank	Attribute / Factor	Unit / Type	Source
1	Flood Hazard	Class	RSA LC
2	Elevation	m	ALOS DEM
3	Effective Rooting Depth	mm	Survey (unclassed)
4	Slope	Degree	ALOS DEM
5	Terrain Position Index	Index	ALOS DEM
6	Soil Permeability	Class	KZN LC
7	Visual Soil Assessment Score	Tally	VSA Maize
8	Flow Accumulation	Integer	ALOS DEM
9	Rockiness	Class	Ecotope
10	Sterkspruit 2100	Soil Form / Family	Soil Classification

Table 5-9 Top features selected by the RF model for Soybean production in the Bergville area

Rank	Attribute / Factor	Unit	Source
1	Aspect	Radians	ALOS DEM
2	Relative Slope Position	Index	ALOS DEM
3	Soil Colour Reflectance	Integer	Spectrophotometer
4	Soil Smell	Class	VSA Soybean
5	Elevation	m	ALOS DEM
6	Effective Depth	mm	Survey (unclassed)
7	Profile Curvature	Integer	ALOS DEM
8	Grey Soil Colours	Class	Spectrometer
9	Crop Ecotope B22	Class	KZN Ecotope
10	Convergence Index	Index	ALOS DEM

5.3.2 Biophysical scoring classification development and testing for the Bergville model building area

5.3.2.1 Development of a maize productivity classification scoring system using a biophysical scoring classification

Three terrain and four soil factors, including one compound soil factor, were selected to create the maize BSC (Table 5-10). These six factors were then scored and weighted using yield differential within each attribute (Table 5-11).

Table 5-10 Selected attributes for the maize biophysical scoring classification

Attribute / Factor	Attribute Group	p	Source
Geomorphon Unit	Terrain	< 0.001*	ALOS DEM
Slope	Terrain	0.041*	Infield Survey
Primary Aspect	Terrain	0.006*	ALOS DEM
Surface Rocks	Soil	0.022*	KZN LC
Soil Mottling	Soil	0.010*	VSA: Maize
Crop Ecotope	Soil	< 0.001*	KZN Ecotope
Soil Permeability	Soil	0.001*	KZN LC

In terms of terrain, geomorphon unit, infield slope and cardinal aspect were found to significantly impact maize yields. Of the ten geomorphons, six specific units, generated from the ALOS 30m DEM, were found to consistently influence maize yields within the 80% classification dataset. Depressions, valleys and summits were associated with reduced yields and were assigned a negative score. Depressions and valley units are associated with wetness limitations while Summit Units are often shallower in term of effective rooting depths. Conversely footslopes, shoulders and flat geomorphons units were assigned a positive score due to their consistently favourable production output. The remaining geomorphon units were given a neutral score of 0, as these units had a low impact on normalised maize yields across the five growing seasons.

Actual infield slope measurements were used in the scoring system to score the impact of slope gradient on maize yields. Gently sloping land at elevated landscape positions tended to produce above yields, however this was negated by flat areas adjacent to major drainage areas, thus slopes of between 0 and 3% were given a neutral score of 0. Slopes of between 3 and 8% within the assessment area are generally associated with deeper, well drained soils and thus a score of +1 was given to this slope class. Finally, steeper areas of over 8%, associated with increased runoff and erosion were given a negative score of -1. Aspect (degrees) derived from the ALOS DEM was converted to the four cardinal directions. North facing slopes were found to have consistently higher yields in this particular landscape and thus a +2 score was assigned to all north facing slopes.

Table 5-11 Maize biophysical scoring classification

Terrain Factors	
Geomorphon (ALOS 30m DEM)	**Score**
Depression	-2
Valley	-2
Footslope	+2
Hollow	0
Slope	0
Spur	0
Shoulder	+2
Ridge	0
Summit	-2
Flat	+2
Infield Slope (%)	**Score**
0 - 3	0
4 - 8	+1
> 8	-1
Primary Aspect (ALOS 30 Dem)	**Score**
North	+2
South	0
East	0
West	0
Soil Factors	
Surface Rocks (KZN LC)	**Score**
R0 - No Rockiness	0
R1 - 2 - 10% Rockiness	-2
R2 - 10 -20% Rockiness	-4
Soil Mottling (VSA: Maize)	**Score**
No Mottles	+1
Mottles Common	0
Mottles Profuse	-2

Soil Factors Continued			
Crop Ecotope (KZN Ecotope)			
Crop Ecotope	**Score**	**Crop Ecotope**	**Score**
B11	+3	E22	-2
B12	0	E23	0
B13	+1	E24	0
B21	+2	H13	-2
B22	-1	H22	-2
B23	0	I13	-2
D11	+2	I24	-2
D12	0	J12	-2
D13	-2	J13	-2
D21	0	J12	-2
D22	0	J13	0
D23	+1	J14	0
E11	+2	J22	-2
E12	-2	J23	0
E13	+2	J24	-2
Soil Permeability Class (KZN LC)			**Score**
(1) Impermeable			-3
(2) Severely Restricted			-2
(3) Restricted			0
(4) Slightly Restricted			0
(5) Good			+1
(6) Rapid			-2
(7) Extremely Rapid			-2

For maize Relative Slope Position (RSP) and Terrain Position Index (TPI) were also found to be significantly correlated to maize yields using the Pearson Correlation Coefficient, where yields decreased with increasing index values. For RSP maize yields decreased towards the valley bottom, where RSP values neared 0. However, this yield trend was more accurately represented by the geomorphon classification, where low lying areas were automatically identified and delineated, without the need to create subjective class breaks for a continuous terrain variable. Consequently, neither RSP nor TPI were included in the final maize BSC. However, for ease of reference correlation values between crop yields and all continuous terrain attributes, is provided in Appendix C.

In terms of soil factors surface rocks, soil wetness class, crop ecotope and subsoil permeability were selected for maize BSC for the Bergville model area. The presence of surface rocks, periodic wetness and either overly restrictive or overly rapid permeabilities were consistently detrimental to production and were scored negatively, while well drained soils and soils with favourable permeability characteristics were scored positively.

Crop ecotope, consisting of soil functional group, topsoil clay and effective rooting depths were found to be important contributing factors to maize production. To reduce scoring complexity this compound attribute was given preference to individual attributes. Generally, crops grown in well, moderately drained and poorly drained soils, with sufficient rooting depths were associated with higher yields. While young, gleyed and duplex soils performed poorly across the five growing seasons and were given negative scores. Potential overfitting, in terms of localised scoring did occur in some instances where standardised yields increased in what would be considered traditionally “poorer” ecotopes and shallower depth classes and this could lead to inconsistent translation in other production areas. No significant relationships could be found between soil colour, bleaching and crop yields using the methods associated with the BSC.

5.3.2.2 Testing the maize biophysical scoring classification

Of the 591 maize observations obtained in the Bergville area, 20% or 118 observations were randomly selected across both farming enterprises and removed from the model building and classification process. These 118 observations were subsequently used to test the accuracy of the locally generated BSC.

Of the 118 observations 55% had a positive or above average SNV for dryland maize, while 45% had a below average standardised yield. The BSC developed for maize production in the Bergville area (Table 5.11), indicates that 78% of the observations had a combination of terrain and soil attributes that should provide an above average yield, a maize classification score of greater than 0, indicating that the resources in the Bergville area are advantageous for dryland maize production. This is supported by the fact that annual maize production across the two farming enterprises is consistently higher than those achieved at a Provincial and National level. Annual crop quality reports produced by The Southern African Grain Laboratory (SAGL) between 2016 and 2020 indicates the average maize production for KwaZulu-Natal was 5.68 $t.ha^{-1}$ and 4.73 $t.ha^{-1}$ for South Africa (SAGL, 2016a – 2020a), compared to 9.5 $t.ha^{-1}$ across the Bergville assessment area for the same 5-year period. Ultimately, FCL and Zunckel

Farming enterprises are producing 167% more than the Provincial average and 202% more than the National average.

The high-performance resource base, from which these yields are achieved, is mirrored in resulting dryland maize classification scores (Figure 5-8). The graph summarises the average SNV of dryland maize yield per scoring class across the 118 test observations using the maize scoring criteria in Table 5-10. The results show that the average classification score, achieved across the observations was +2.8, again suggesting above average growing conditions.

The maize scoring classification (Figure 5-8) produced a total of four “negative” classes with classification scores ranging from -3 to 0. It is anticipated that these classes will produce below average yields while the ten positive scoring classes, with scores ranging from 1 - 10, are expected to produce above average yields. The results achieved almost match these expectations, with all negative classes producing below average yields and all but one positive class producing above average yields. Class 1, the poorest positive class obtained an average SNV of -5%, this result indicates that areas with relatively average terrain and soil resources are difficult to predict in terms of actual production performance with small yield variations negatively influencing classification accuracy.

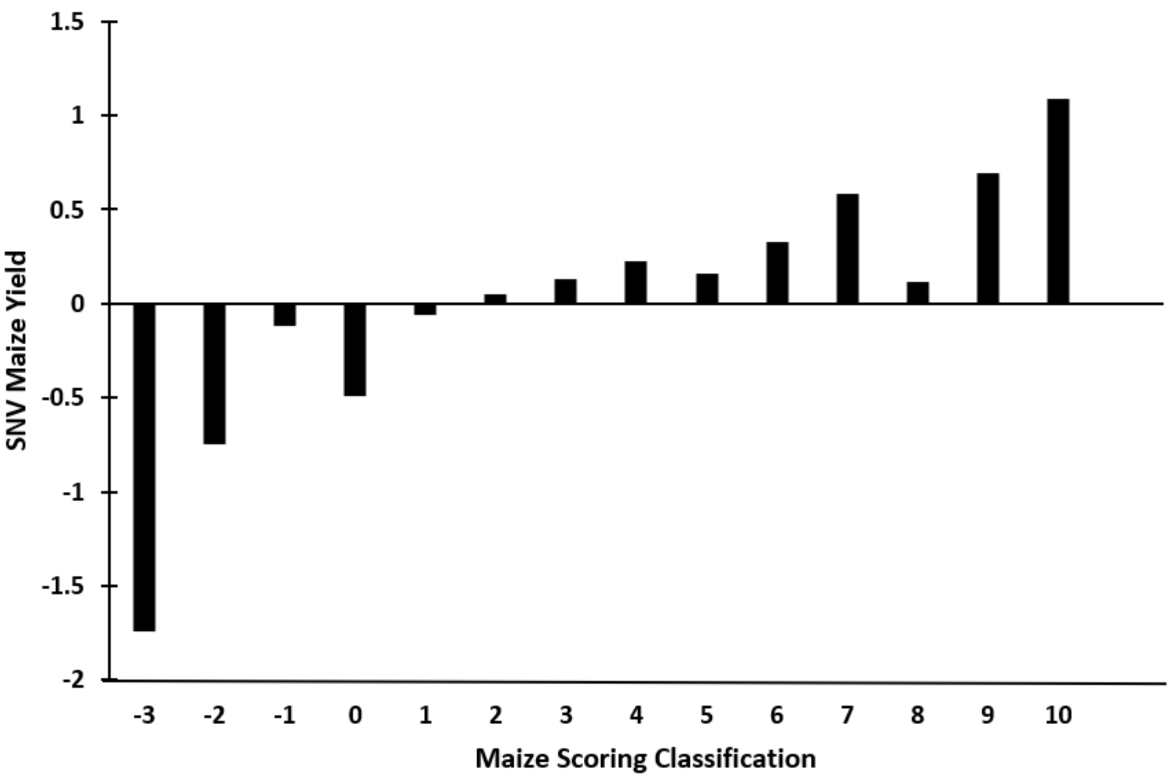

Figure 5-8 Average SNV maize yield achieved in each maize biophysical classification scoring class for the Bergville area

If the classification is working adequately then standardised yield values should increase with maize scoring class, meaning lowest yields should be obtained in Class -3, while highest yields should be obtained in Class 10. The results shown in Figure 5-8 illustrates this expectation, with the lowest average SNV indeed occurring in the lowest scoring class with the highest SNV occurring in the highest maize scoring class. This indicates that the biophysical classification system, produced for the Bergville area, can highlight extreme terrain and soil factors that cause both very low and very high maize high yields. Finally, the results in Figure 5-8 illustrates that obtained SNV yields follow a S-Shape, typical of a cubic polynomial function with SNV scores being close to 0 for the maize scoring classes between -1 and 2, this shape is caused by the normally distributed maize yield data, with most of the yield data falling within one standard deviation. Consequently, the yield performance within these classes will generally be the most difficult to predict. A correlation analysis, using Pearson 2-tailed correlation found that both average median yield and average SNV were significantly correlated, to the biophysical scoring classification, at a 0.001 significance level. Average median yield produced a 0.302 correlation with maize scoring class, while average normalised produced a 0.404 correlation. These results indicate that the classification is performing adequately.

Figure 5-9 summarises the binary classification accuracy in each maize scoring class. Overall, the classification accuracy for 118 test observations was 65%. A correct prediction occurs either when a positive maize class (1 - 10) corresponds to an above average SNV yield value (SNV>0), or when a negative maize class (-3 - 0) corresponds to a negative SNV yield value (SNV<0). As suggested the BSC system was most accurate at either end of the maize scoring classification, with a 78% classification accuracy for classes -3 and -2 and classes between 6 and above. While binary accuracy was reduced in moderate scoring classes with a total 54% of the misclassifications occurring in 4 classes between -1 and 2. This again highlights the difficultly of yield performance predication in areas associated with average terrain and soil factors. Lower classification accuracies obtained in Classes 4 and 5 could be a reflection of non-physical soil properties impacting crop performance or a poor model fit for these types of areas.

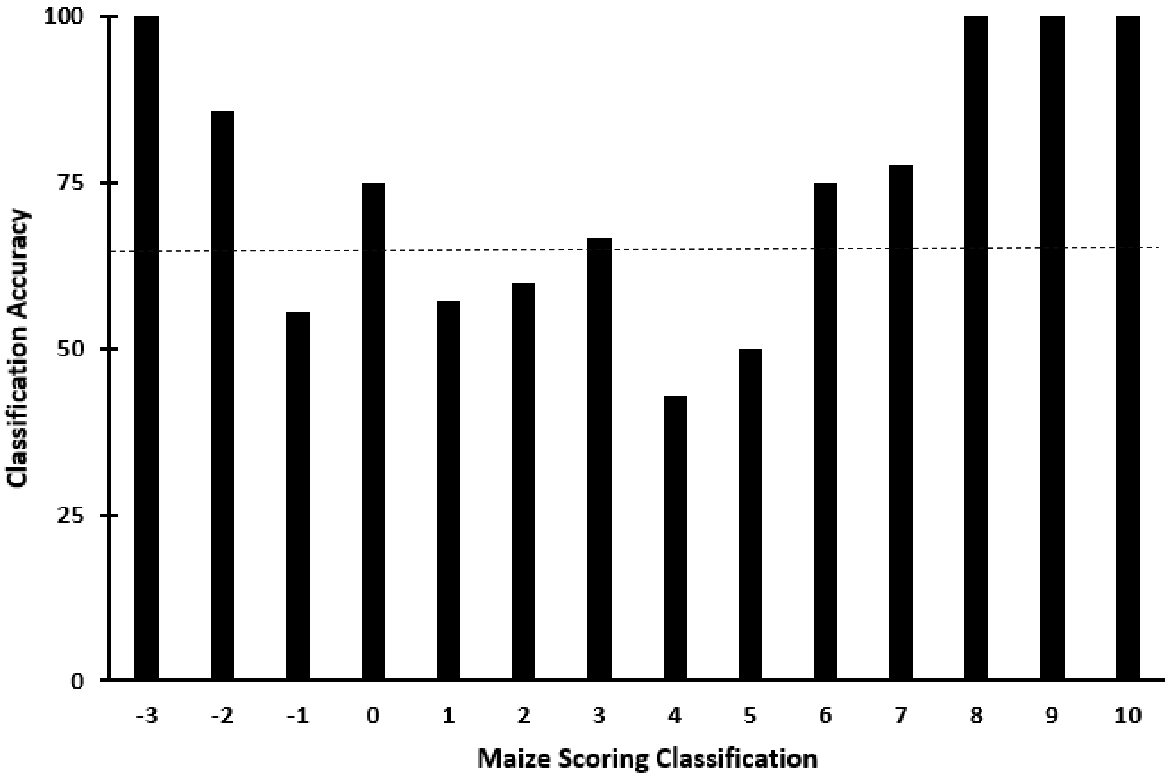

Figure 5-9 Classification accuracy achieved in each biophysical maize classification scoring class for the Bergville area. Average classification accuracy is indicted by the dotted line.

The 14 individual maize scoring classes were combined into a binary classification. Classes -3 – 0 were combined into a single class, named "Negative Maize Class" while classes 1 – 10 where combined into the "Positive Maize Class". The average dryland yield performance of maize within this binary classification is provided in Table 5-12. The average median maize yield was significantly different ($p = 0.003$) across the two classes with observations with a positive maize scoring classification of producing 1.84 t.ha^{-1} more maize. Correspondingly, observations with a positive maize scoring classification had a significantly higher average SNV ($p < 0.001$) across the five growing seasons. Observations obtaining a negative maize score also obtained a negative average SNV for maize, averaging 52% lower. These results indicate that the maize BSC is able predict below and above average yield performance in this production environment.

Table 5-12 Average median maize yield, standard deviations and average SNV per maize prediction class using the biophysical scoring classification

Maize Classification Prediction	n	Avg. Median Yield (t.ha^{-1})	SD of Avg. Median Yield	Avg. SNV (%)
Positive Maize Score	92	10.49^{a}	1.30	0.21^{a}
Negative Maize Score	26	8.65^{b}	1.10	-0.52^{b}

The confusion matrix for dryland maize BSC system is provided in Table 5-13, and its associated performance metrics are provided in Table 5-14. The system has a classification accuracy rate of 65%, with most correct classifications occurring where a positive maize score corresponded with an above average yield. Of the 41 individual misclassifications 34 occurred where the terrain and soil factors combined to predict an above average SNV yield result, where in fact a below average SNV was recorded. Only 7 of 118 (6%) observations were classified as a False Negative, where negative maize scores lead to above average yield performance. In this case a False Negative is soil whose production performance is underestimate by the BSC. This low occurrence should be viewed as an improvement on more conservative classification methods, such as RSA where land is often downgraded to lower capability classes greatly underestimating crop performance.

Table 5-13 Confusion matrix results for the maize biophysical scoring classification at Bergville

Bergville Maize	Actual Positive SNV	Actual Negative SNV	Total
Positive Maize Score	58	34	92
Negative Maize Score	7	19	26
Total	65	53	118

Classification Precision is the ratio of positive predicted values to actual true values (Table 5-14). The BSC resulted in a Precision value of 63%, meaning that only 63% of the observations classed as a having positive maize score did in fact produce above average yields. The method scored well in terms of Classification Recall (89.2%), which is the probability that the maize scoring system correctly predicts an actual true value. This again suggests that the scoring system can distinguish between low and high yielding areas. Finally, the maize BSC achieved an MCC of 0.30, which indicates a moderate model performance in terms of prediction (Mukaka, 2012).

Table 5-14 Classification performance metrics for the maize biophysical scoring classification at Bergville

Metric	Result
CA (%)	65.3
MR (%)	34.7
Precision (%)	63.0
Sensitivity (%)	89.2
Specificity (%)	33.9
MCC	0.30

5.3.2.3 Development of a soybean productivity classification scoring system using a biophysical scoring classification

As with maize, soybean crop performance was assessed using a BSC approach. However, for soybean the soil and terrain factors driving production differ (Table 5-15), as well as their respective scoring and weights (Table 5-16).

Table 5-15 Selected attributes for the soybean biophysical scoring classification

Attribute / Factor	Attribute Group	p	Source
Basic Terrain Unit	Terrain	0.048*	Infield Survey
Soil Functional Group	Soil	0.044*	KZN Ecotope
Effective Rooting Depth	Soil	0.002*	Reclassed from Infield Survey
Soil Texture	Soil	0.007*	RSA LC
Soil Crusting	Soil	0.046*	KZN LC
Soil Permeability	Soil	0.004*	KZN LC

In terms of terrain attributes the five-unit terrain model, recorded infield, along with RSP, Flow Accumulation and Aspect (radians) were all found to be significant in determining soybean yield. When converted to cardinal directions the aspect results were found to be insignificant. Of the three remaining significant terrain attributes, the basic five-unit terrain model was preferred in the final soybean BSC as it is a more holistic terrain attribute with predefined class breaks. Again, for reference, correlation values between crop yields and all continuous terrain attributes, is provided in Appendix C.

The basic terrain unit model which segments the landscape into five discrete units was found to significantly impact yields in four of this units, rests, scarps, footslopes and valley bottoms. The valley bottom unit, associated with significant water accumulation and flooding risk achieved an average SNV of -37% and were consequently assigned a score of -3. Footslopes, associated with moderate water accumulation, consistently produced above average yields and were assigned of +1. Generally, units higher in the landscape produced below average yields and were scored accordingly, with the crest and scarp unit significantly depressing yields and were assigned a score of -2.

In terms of significant soil attributes crusting and permeability along with soil functional group, soil texture and effective rooting depth, were selected. However, unlike maize these factors were scored separately to more accurately assess soybean production (Table 5-16). Higher yielding soil groups including well drained, alluvial and moderately drained soils, were assigned a positive scoring classification.

Table 5-16 Soybean biophysical scoring classification

Terrain Factors	
Basic Terrain Unit (Infield)	**Score**
Crest	-2
Scarp	-2
Midslope	0
Footslope	+1
Valley Bottom	-3
Soil Factors	
Soil Texture Group (RSA LC)	**Score**
Group T1	+1
Group T2	+1
Group T3	-2
Effective Rooting Depth (Field Survey)	**Score**
0 - 450 mm	-2
450 - 750 mm	-1
> 800 mm	+1
Soil Crusting Class (KZN LC)	
(t0) No crusting	+1
(t1) Moderate crusting	+1
(t2) Exterme crusting	-2

Soil Factors Continued	
Soil Functional Group (KZN Ecotope)	
Soil Group	**Score**
(B) Well and moderately drained	+1
(C) Alluvial	+1
(D) Mottled and moderatly drained	+1
(E) Mottled and poorly drained	0
(H) Young	-2
(I) Gleyed	0
(J) Duplex	-2
Soil Permeability Class (KZN LC)	**Score**
(1) Impermeable	-3
(2) Severely Restricted	-3
(3) Restricted	-3
(4) Slightly Restricted	-1
(5) Good	+2
(6) Rapid	0
(7) Extremely Rapid	0

Young and duplex soils, all associated with severe physical soil limitations, were assigned a score of -2. Of the observations containing theses soils groups, 75% produced below average yields across the five growing seasons. Three classes of effective rooting depths were created from the unclassed rooting depth attribute. Only depths of more than 750 mm were assigned a positive score while moderate depths, 450 - 750 mm and effective depths of less than 450 mm were given a score of -1 and -2 respectively. Finally, severe physical characteristics linked to soil textural group 3, restrictive permeabilities as well as extreme surface crusting were assigned negative scoring classifications. While soil textural classes 1 and 2 topsoils, “good” subsoil permeabilities and non-severe crusting were assigned positive scores.

To produce the final soybean BSC the scores were tallied across the six individual factors. Like maize and in terms of the binary classification, scores below 1 were classed as negative for soybean production, while scores greater than 1 were assigned a positive classification.

5.3.2.4 Testing the biophysical soybean productivity classification scoring system

Of the 351 soybean observations obtained in the Bergville area, 20% were randomly selected and excluded from the model building and classification process, in order to test the accuracy of the locally generated classification. Of the 70 testing observations 46% produced a negative SNV, while 54% obtained positive SNV over the five-year assessment period. Conversely, the BSC predicted that of the 70 test observations, 70% should produce above average yields.

While only 30% of the test observations were predicted to produce below average yields. This result is 8% lower than that of the maize BSC, but still suggests the resource base in Bergville is weighted towards higher yielding crop production. The yields achieved in Bergville, when compared to Provincial and National averages, again support this observation. With the soybean yields achieved at Bergville being 128% more than the Provincial dryland average of 2.49 $t.ha^{-1}$ and 189% more than National average 1.69, as per SAGL Reports across the five-year period (SAGL, 2017b-2020b).

The results of the soybean BSC (Table 5-16) versus SNV soybean yield is presented in Figure 5-10. Across the testing observations a total of fourteen different scores were recorded, seven negative (-14 - 0) and seven positive (1 - 7). Of the seven negative classes all seven produced below average SNV for soybean, the yields decreased relatively constantly from class -14 to 0. The three lowest classes (-14, -10 and -4) also produced the high classification accuracies, in terms of binary classification with a perfect 100% accuracy (Figure 5-11). Of the seven positive classes (>0), six produced above yields with classes 5, 6 and 7, being the top yield performers. These classes, associated with conducive soil and terrain properties for soybean production all produced high CA, of over 75%.

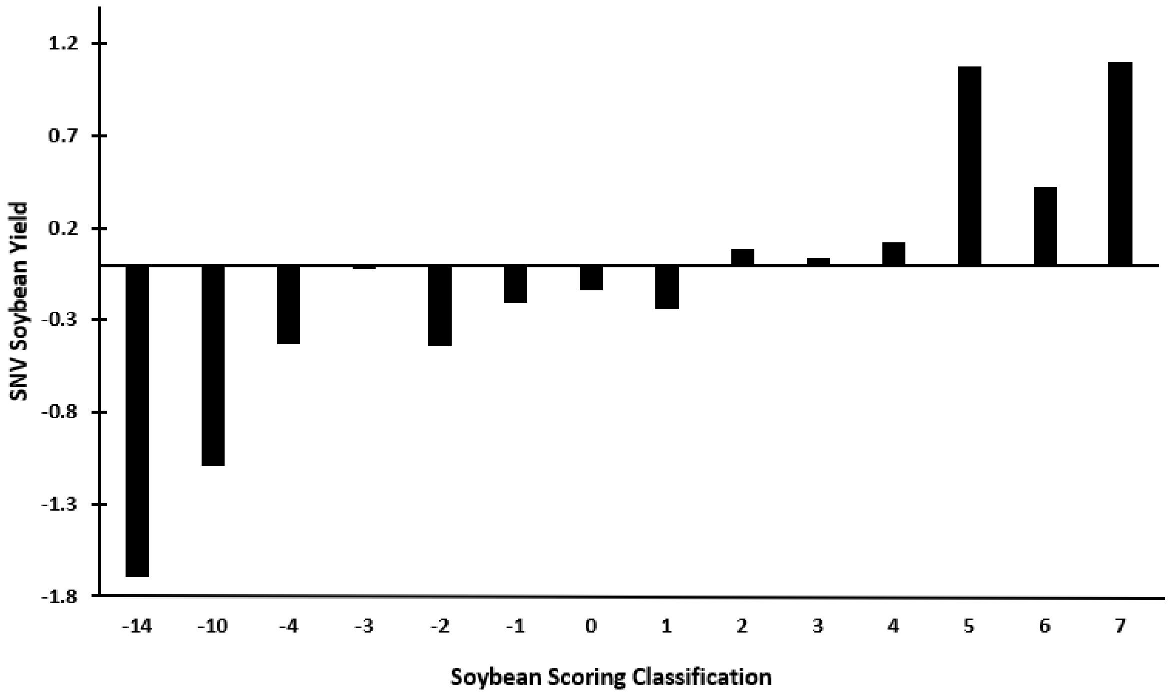

Figure 5-10 Average SNV soybean yield achieved in each soybean classification scoring class for the Bergville area

Like maize, biophysical soybean classification binary accuracy was reduced in moderate scoring of between -2 and 3 (Figure 5-11). Again, highlighting the difficultly of yield performance prediction in areas associated with average terrain and soil factors. However, on the whole SNV for soybean production followed the expected trend where yields were below average in the negative scoring classes and above average in the positive scoring classes. This was confirmed by a correlation analysis using Pearson 2-tailed correlation found that both average median yield (0.370) and average SNV (0.504) were significantly correlated, to the soybean BSC, at a 0.001 significance level. These results should also be viewed context, given the difficulty of traditional land assessment methods (Chapter 4), to predict soybean performance in this production environment.

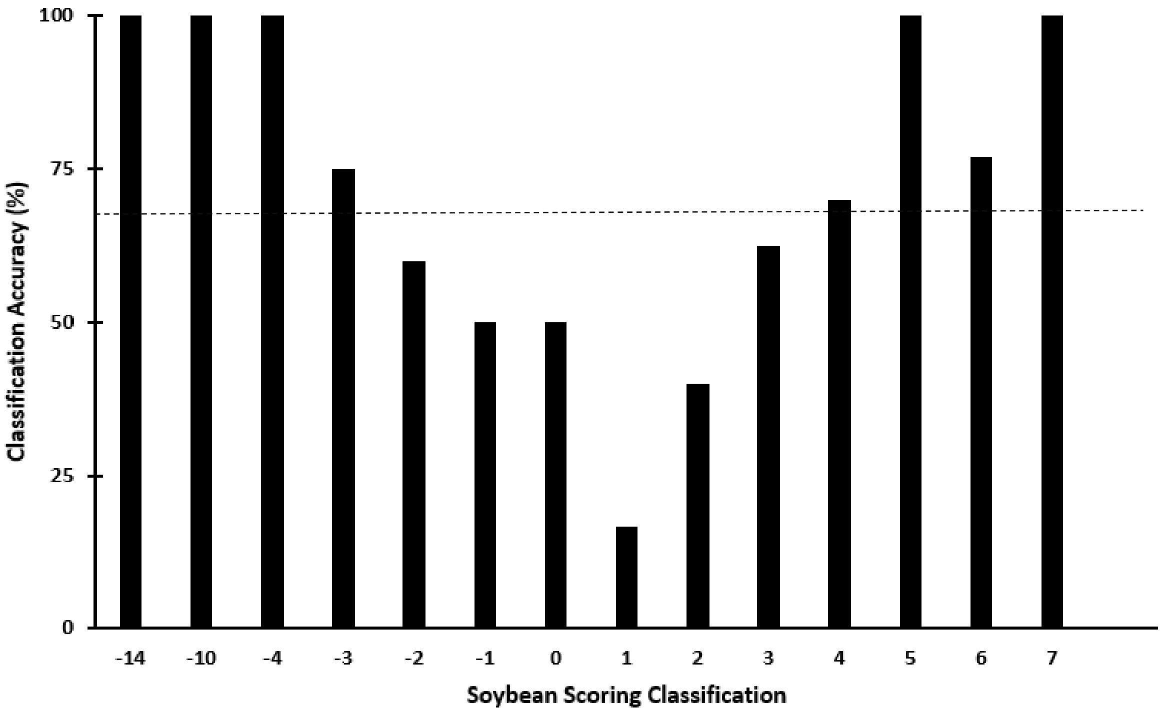

Figure 5-11 Classification accuracy achieved in each soybean classification scoring class for the Bergville area. Average classification accuracy is indicted by the dotted line.

The 14 individual soybean scoring classes, based on the scoring of the selected soil terrain factors, were combined into a binary classification, labelled “Negative Soybean Class” and “Positive Soybean Classes”. The average dryland yield performance of soybean within this binary classification is provided in Table 5-17. The average median soybean yield was significantly different ($p = 0.022$) across the two classes, with observations with a positive soybean scoring classification producing 0.5 $t.ha^{-1}$ more than it negative counterparts. Correspondingly, observations with a positive soybean BSC had a significantly higher average SNV ($p < 0.003$) across the five growing seasons, combining to produces yields 28% above average. While observations obtaining a negative soybean score averaged some 37% lower, over the five growing seasons. These results indicate that the soybean BSC is able predict below and above average yield performance in this production environment. These significant yield differences achieved across these classes should be viewed as a significant improvement on traditional approaches (cf Chapter 4.3.3).

Table 5-17 Average median soybean yield, standard deviations and average SNV per soybean prediction class using the biophysical scoring classification

Soybean Classification Prediction	n	Avg. Median Yield ($t.ha^{-1}$)	SD of Avg. Median Yield	Avg. SNV (%)
Positive SNV	49	3.20^{a}	0.91	27.93^{a}
Negative SNV	21	2.70^{b}	0.77	-37.37^{b}

The confusion matrix for dryland soybean BSC is provided in Table 5-18 and its associated performance metrics are provided in Table 5-19. The system has a classification accuracy rate of 65.7%, with most correct classifications occurring where a positive soybean score corresponded with an above average yield. Of the 24 individual misclassifications 72% occurred where the terrain and soil factors combined to predict an above average SNV yield result, where in fact a below average SNV was recorded. In these instances, the crop yield performance is below the predicted resource potential.

Only 10% of observations were classified as a False Negative, where negative soybean scores led to above average yield performance. In this case a False Negative is soil whose production performance is underestimated by the BSC approach. As with the maize classification, this low occurrence should be viewed as an improvement on more conservative classification methods, such as the RSA LC.

Table 5-18 Confusion matrix results for the biophysical dryland soybean classification at Bergville

Bergville Soybean	Actual Positive SNV	Actual Negative SNV	Total
Predicted Positive SNV	32	17	49
Predicted Negative SNV	7	14	21
Total	39	31	70

In terms of Precision, 65.3% of the observations classed as a having positive soybean score did in fact produce above average yields (Table 5-19). The method scored well in terms of Classification Sensitivity (82.1%), which is the probability that the soybean scoring system correctly predicts an actual true value. Like the parametric maize classification this high Sensitivity suggests that soybean scoring system can distinguish between low and high yielding areas. Finally, the parametric soybean system achieved an MCC of 0.29, which is considered to be moderate in terms model performance (Mukaka, 2012).

Table 5-19 Classification performance metrics for the biophysical dryland soybean classification at Bergville

Metric	Result
CA (%)	65.7
MR (%)	34.3
Precision (%)	65.3
Sensitivity (%)	82.1
Specificity (%)	45.2
MCC	0.29

5.3.2.5 Contextualising maize and soybean classification performance using biophysical scoring classification

In agricultural land assessment there is a need to contextualise the findings of the classification system or model. An assessed portion of land should not only be viewed in isolation but also compared to land and crop performance at a regional, Provincial and/or National scale. Unfortunately, reliable regional yields were not available and thus Provincial and National yields from SAGL (2016-2020), were used to judge crop and farm performance from a Provincial and National perspective (Tables 5-20 and 5-21).

As previously stated, the FCL and Zunckel Farming enterprises which comprise the model building area, produce on average,167% more maize than the Provincial average and 202% more than the National average. Similarly, soybean yields achieved in the model building area produced 128% more than the Provincial dryland average and 189% more than the National average over the five growing seasons. In order to contextualise land and crop performance the Provincial and National yield averages were used to benchmark the BSC for maize and soyabean.

Confusion matrixes (Table 5-20) were compiled using the annual Provincial and National averages between 2016 and 2020 and compared to the binary BSC for maize and soybean. The lower yields achieved at both a Provincial and National level greatly increases the count of above average yield observations in the Bergville area. When benchmarked against Provincial yields 111 of 118 (94%) of Bergville maize observations are above the Provincial average while 98% are above the National average. A similar increase occurs in soybean benchmarking with all but one observation being below the National soyabean average. This benchmarking skews both True Positive and False Negative counts, while almost removing False Positive and True Negative counts (Table 5-21). This ultimately results in a higher classification accuracy as both maize and soybean BSC approaches assigned above average scores to the local terrain and soil resources.

When benchmarked against Provincial and National crop averages the classification accuracy improves from approximately 65% at farm level to between 69 and 78% for both maize and soybean (Table 5-21). Along with classification accuracy, precision is also improved using Provincial and National yield benchmarking, compared to locally achieved yield averages and highlight the significant production potential of the Bergville farms. Unfortunately, the benchmarking process also skews the number of True Positive and False Negative occurrences resulting in decreases in both Sensitivity and MCC values. A regionally determined yield benchmark for both maize and soybean, could see improved classification accuracies without comprising the overall model performance. The importance of benchmarking and spatial contextualisation is further discussed for the three verification areas.

Table 5-20 Set of confusion matrixes using Provincial and National Yield Averages. Bracketed figures show changes in counts from farm level results

KZN Maize	Above Average Yield	Below Average Yield	Total
Positive Maize Score	88 (+30)	4 (-30)	92
Negative Maize Score	23 (+16)	3 (-16)	26
Total	111	7	118
RSA Maize	Above Average Yield	Below Average Yield	Total
Positive Maize Score	91 (+33)	1 (-33)	92
Negative Maize Score	25 (+18)	1 (-18)	26
Total	116	2	118
KZN Soybean	Above Average Yield	Below Average Yield	Total
Positive Maize Score	46 (+14)	6 (-11)	52
Negative Maize Score	14 (+7)	4 (-10)	6
Total	60	10	70
RSA Soybean	Above Average Yield	Below Average Yield	Total
Positive Maize Score	47 (+15)	2 (-15)	49
Negative Maize Score	20 (+13)	1 (-13)	21
Total	67	3	70

Table 5-21 Performance metrics for various confusion matrixes for maize and soybean yields at farm, Provincial and National levels

	Maize Farm	Maize KZN	Maize RSA	Soybean Farm	Soybean KZN	Soybean RSA
CA (%)	65.3	77.1	78.0	65.7	71.4	68.6
Precision (%)	63.0	95.7	98.9	65.3	88.5	95.9
Sensitivity (%)	89.2	79.3	78.4	82.1	76.7	70.1
Specificity (%)	33.9	42.9	50.00	45.2	40.0	33.3
MCC	0.30	0.13	0.09	0.29	0.13	0.02

5.3.3 Performance assessment of biophysical scoring classification and machine learning models for dryland maize in three different verification areas

A total of 234 observation points, across the three verification areas, were used to verify the performance of the BSC as well as the RF and SVM machine learning models. The three selected verification farms all produced above average maize yields, when compared to the Provincial and National records (Table 5-22). Newcastle produced the highest yields, followed by Bloodriver and Luneburg, which averaged less than 8 $t.ha^{-1}$ across the three verification seasons. Based on the broad climatic and resource potential extracted from the BRU reports (Camp, 1999), Luneburg should be highest performer of the three verification areas, however local soil and terrain conditions may be limiting production potential.

Table 5-22 Summary of Provincial and National dryland maize yield performance compared to the three verification areas at Bloodriver, Luneberg and Newcastle

Year	Bloodriver Maize Yields ($t.ha^{-1}$)	Luneburg Maize Yields ($t.ha^{-1}$)	Newcastle Maize Yields ($t.ha^{-1}$)	Provincial Maize Yields ($t.ha^{-1}$)	National Maize Yields ($t.ha^{-1}$)
2018	9.04	8.79	9.03	5.82	4.80
2019	7.05	7.82	8.04	5.63	4.34
2020	8.12	6.18	9.37	5.79	5.46
Average	8.07	7.60	8.81	5.75	4.87

The verification process again utilises binary classification, using farm specific yields, to assess model performance. For the BSC approach, a correct prediction occurs either when a positive maize classification corresponds to an above average SNV yield value (SNV>0), or when a negative classification corresponds to a negative SNV yield value (SNV<0). For the machine learning models its performance is based on the accuracy of the model to predict an above average yield event as well as to classify the probability of both the above below average event. A summary of the performance of the three approaches, across the three verification areas, is provided in Tables 5-23- 5-25.

Overall, the Random Forest model achieved the highest accuracy across the three verification areas. The prediction accuracy of 56.8% is almost identical to the accuracy achieved in the Bergville model building area of 57.2%. The BSC was second best approach in terms of accuracy and achieved an overall accuracy of 53%, some 12% lower than the accuracy achieved in Bergville. The SVM model was not only the poorest overall performer for maize binary yield prediction but also for each individual verification area (Table 5-25). It also produced the lowest classification performance metrics of the three approaches. Thus, the remainder of the results section for maize will focus on the performance of the RF Model and maize BSC.

Both the RF model and BSC performed well at Bloodriver with prediction accuracies of 66% and 71%, respectively. Overarching performance indicators for both approaches, MCC for the biophysical classification and AUC for the RF model, were also the highest of the three verification areas. Both approaches also had high specificity values, which is the proportion of true negatives that are correctly predicted by the model, in this case below average maize yields. Most of the misclassifications occurred when the models incorrectly predicted a below average yield. This results in lower sensitivity values of 52% for RF and 56% for the maize BSC and were the lowest achieved across the verification farms. This error produces False Negative which is a Type II error, whereby the model or classification is underestimating the actual land performance. In terms of land assessment this type of error should be avoided as you are undervaluing land potential. A longer assessment period, incorporating more seasonal variation, as well as addition samples could determine if this error is consistently occurring, which may require local calibration of the method.

Table 5-23 Summary of performance metrics for the maize biophysical scoring classification for the three verification areas

	Overall	Bloodriver	Luneburg	Newcastle
n	234	44	68	122
CA (%)	52.6	70.5	57.4	43.4
Precision (%)	54.6	87.5	56.5	47.8
Sensitivity (%)	66.4	56.0	74.3	66.2
Specificity (%)	36.7	89.5	39.4	17.5
MCC	0.03	0.47	0.15	-0.18

Table 5-24 Summary of performance metric for the Maize Random Forest Model for the three verification areas

	Overall	Bloodriver	Luneburg	Newcastle
n	234	44	68	122
CA (%)	56.8	65.9	54.4	54.9
Sensitivity (%)	79.2	52.0	77.1	90.7
Specificity (%)	31.2	84.2	30.3	14.0
AUC	0.47	0.73	0.53	0.41

Table 5-25 Summary of performance metric for the Maize Support Vector Machine Model for the three verification areas

	Overall	Bloodriver	Luneburg	Newcastle
n	234	44	68	122
CA (%)	46.1	56.8	52.9	38.5
Sensitivity (%)	66.1	28.0	31.4	27.3
Specificity (%)	28.8	94.7	75.8	50.8
AUC	0.49	0.63	0.59	0.43

More insight into the performance of the biophysical maize scoring classification across the three verification areas is provided in Figures 5-12a-c. and 5-13a-c. The Bloodriver verification area (Figure 5-12a) shows that all negative scoring classes produced below average yields, indicating that the biophysical maize classification can accurately determine limiting soil and terrain factors, which consistently produce negative SNV maize yields. The scoring distribution for Bloodriver ranged from -4 and 3. While 40% of observations were classified as negative in terms the maize scoring classification. These results indicate that soil and terrain factors are poorer in Bloodriver than at Bergville and the other two verification areas. These poorer baseline resources are also highlighted in the BRU overviews (Table 5.2). In terms of classification accuracy, the BSC was generally lowest on either end of the scoring classification, where scoring classes -4 and 3 achieved accuracies of near 50%. The low accuracies associated with the -4 scoring class, consists of sandy and poorly drained soils, whose combination was not found extensively in the model area building. To improve prediction accuracies in the Bloodriver area these specific ecotopes may require further calibration with localised yield data.

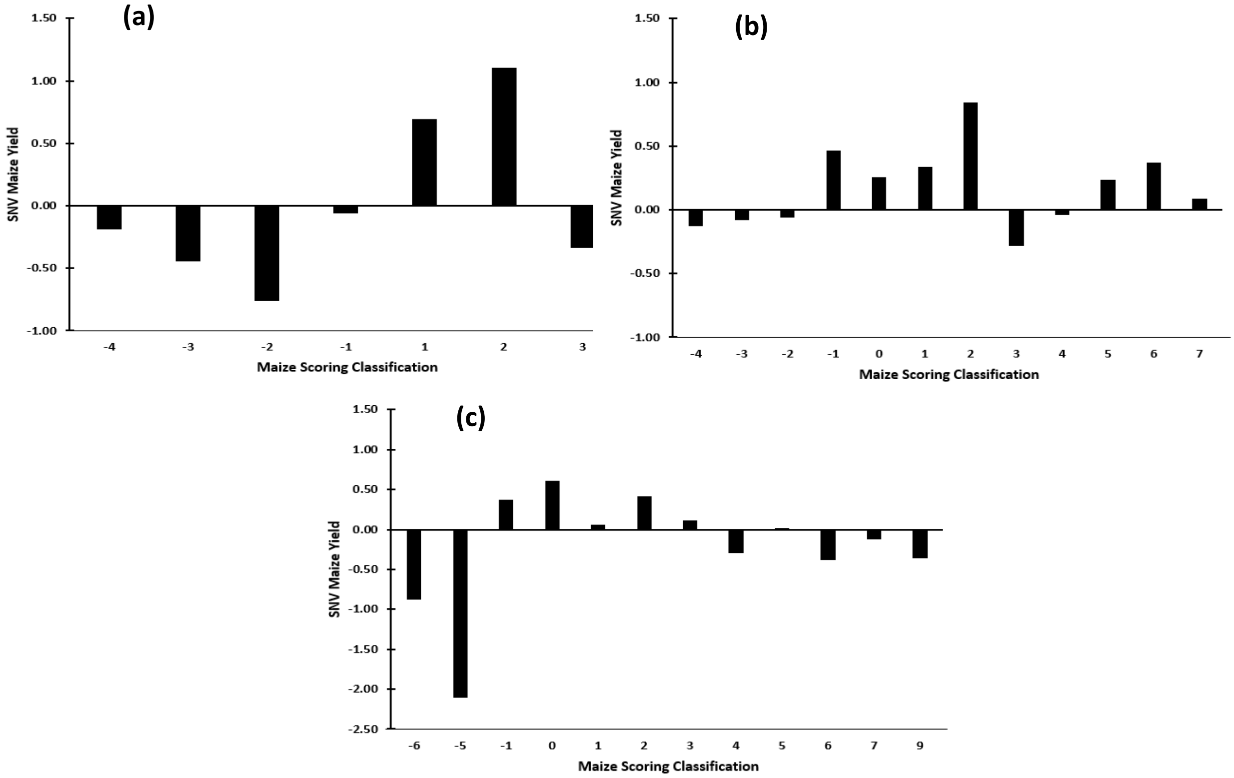

Figure 5-12 Average SNV maize yield achieved in each maize classification scoring class for - (a) Bloodriver (b) Luneburg and (c) Newcastle

The Luneburg area produced moderate prediction accuracies of 54% for RF model and 57% for the BSC (Tables 5-23 and 5-24). Both approaches produced low specificity values for Luneberg, where less than 40% of true negative were correctly predicted by either model. The MCC for biophysical maize classification drops from 0.30 for model building Bergville area to 0.15 for Luneburg. While AUC similarly drops for the RF model from 0.67 in Bergville to 0.53 for Luneberg, reducing the confidence in transferability of both approaches. Figures 5-12b and 5-13b indicate that average SNV maize yield across the biophysical scoring classes and associated classification accuracies are erratic, further reducing the successful transferability of the Bergville models to Luneburg. Only three of the five negative scoring classes contain below average yields (Figure 5-12b), while accuracies achieved in the high scoring classes (3-7) are below 60% (Figure 5-13b). Expected trends such as increasing SNV with scoring class and high prediction accuracies at both ends of the scoring spectrum are not observed with the Luneburg data. This indicates that the physical terrain and soil drivers as well as the associated scoring system are not performing adequately and significant local calibration may be required for this production area.

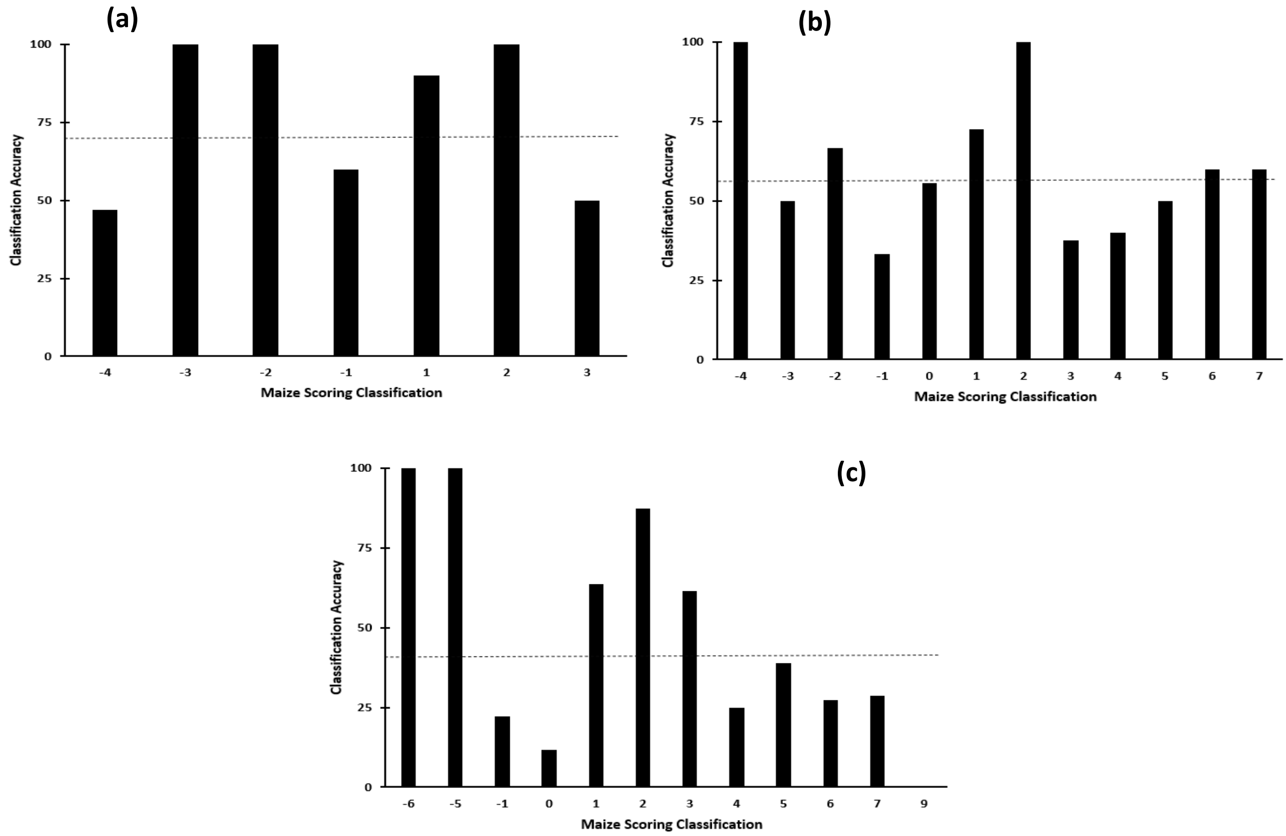

Figure 5-13 Average classification accuracy achieved in each maize classification scoring class for - (a) Bloodriver (b) Luneburg and (c) Newcastle. The dotted line represents average classification accuracy across all scoring classes

One of the potential reasons for poor model performance at Luneburg is that from an overarching climatic and potential perspective the Luneberg area is significantly different to both Bergville and the other two verification areas. The Luneburg verification area is both cooler and wetter and on average receives the lowest heat units of all the study sites (Table 5-1). For the three verification seasons between 2017 and 2020 the Luneberg area consistently received more rainfall than either Bergville, Bloodriver and Newcastle. The comparison between yield in high and low rainfall years (Section 3.3.3.4) indicates that physical properties generally, on which both these approaches are based, became more pronounced during drier cycles. While higher rainfall appeared to mask these physical soil limitations, which may in all likelihood be occurring in Luneberg. Guo et al. (2012) reported similar results for cotton yields where yield and soil properties had a weaker correlation in wetter years. Due to higher rainfall many of the physical soil and terrain drivers and scoring criteria will differ in the Luneberg area. Ultimately, to improve model performance and better assess farm level yield variation a regionalised model or modified BSC may need to be developed in the Luneberg area.

The Newcastle area appears to be a challenging area for the both the BSC and RF model, resulting in low prediction accuracies and poor performance metrics (Tables 5.23 and 5.24). The BSC produced an accuracy of less than 50% and -0.18 MCC, indicating the scoring was negatively correlated to yield (Table 5-23). The AUC for the RF model was 0.41, the lowest of the three verification areas and indicating the models are performing worse than a random classification. The metrics extracted from the confusion matrixes indicate that Newcastle produced specificity values of less than 20%, meaning both models could not predict True Negative events. For the biophysical maize classification nearly 70% of the misclassifications were False Positives, meaning the model predicted above average yields but a below average SNV was observed.

Even though the overarching performance metrics, MCC and AUC, for both approaches are poor in Newcastle, the BSC approach is performing as expected in certain aspects. For example, when maize scoring classes is compared to average normalised yield (Figure 5-12c) extreme negative scoring classes, -6 and -5 produce, negative yields at high classification accuracies (Figure 5-13c). While moderate scoring classes -1, 0 and 1, associated with average soil and terrain scores are erratic and the most difficult to predict. These are expected trends based on the Bergville model building results. Further, the maize BSC produces eight positive maize scoring classes and only 4 negative classes and predicts that 90 of 122 observations will produce above average yield. A similar ratio to the Bergville area, which is associated terrain and soil properties suited to high yielding maize production.

The source of the poor classification performance is in the high maize scoring classes (>4), where terrain and soil properties should produce significantly above average yields (Figure 5-12c). These observations are located on moderate slopes and are dominated by deep, well and moderately drained apedal soils with favourable water holding characteristics, yet across these five classes the average SNV for maize production is -23%. These soils do not have any physical limitations and are traditionally given the highest potential ratings in terms of maize production (Camp, 1999). Essentially crop performance is well below expected land potential.

To investigate the source of this disjuncture soil samples were extracted at 0-10 cm, 30 cm and 50 cm depths for the affected soils. Of the 35 observations, within these high scoring maize classes, 28 contained at least one sample depth where acid saturations exceeded 20%, the highest Permissible Acid Saturation level in maize (Manson et al., 2017). When acid saturations exceed 20%, crop performance is diminished and yields reduced. The average acid saturation in the affected soils was 36%, ranging from a maximum of 55% to a minimum of 29% with the most common depth exceeding PAS being 30 cm. Based on this analysis the maize BSC is performing adequately in Newcastle, but whose accuracy is compromised by poor soil fertility management.

The classification and models for Bergville and verified at Bloodriver, Luneburg and Newcastle were developed under assumption that these commercial operations are performing at high level where fertility and management are optimised and yields maximised. This assumption allows for physical properties to be the major determinant in yield variation. Yet, for Newcastle acid saturation, which can easily be ameliorated through by liming is the major yield determinant across many high potential soils. The necessary absence of chemical and management factors is a definite drawback to this approach but does provide an unexpected application of these physically based models and classifications in commercial environments. Whereby high scoring soils, with consistently low SNV, can easily be identified for management and chemical interventions.

5.3.4 Performance assessment of biophysical scoring classification and machine learning models for dryland soybean in three different verification areas

A total of 72 observation points, across the three verification areas, were used to assess the performance of the biophysical classification system as well as the RF and SVM machine learning models. Table 5-26 provides an annual yield summary for the three verification areas as well as Provincial and National soybean yields achieved during the assessment period. Over the year assessment period the Bloodriver and Luneburg farms produced well above 3 $t.ha^{-1}$ well above the Provincial soybean average of 2.58 $t.ha^{-1}$. While Newcastle produced a mixed set of production results, when compared to Provincial averages. In 2018 Newcastle produced below average yields, did not plant soybean in 2019 and produced well above average yields in 2020. The three selected verification farms all produced above average soybean yields, when compared to the Nationally recorded yields (Table 5-26).

Table 5-26 Summary of Provincial and National dryland soybean yield performance compared to the three verification areas at Bloodriver, Luneberg and Newcastle

Year	Bloodriver Soybean Yields ($t.ha^{-1}$)	Luneburg Soybean Yields ($t.ha^{-1}$)	Newcastle Soybean Yields ($t.ha^{-1}$)	Provincial Soybean Yields ($t.ha^{-1}$)	National Soybean Yields ($t.ha^{-1}$)
2018	2.99	3.80	1.84	2.85	1.83
2019	3.32	3.42	n/a	2.35	1.45
2020	3.90	2.65	3.40	2.55	1.63
Average	3.40	3.29	2.62	2.58	1.64

As with maize, soybean performance was assessed using the BSC along with the machine learning RF and SVM models, developed at the Bergville model building area. A summary of the performance of the three approaches, across the three verification areas, is provided in Tables 5-27 - 5-29. Overall, the soybean BSC approach produced higher accuracies than the two machine learning models (Table 5-27). A performance comparison between the two machine learning models, using prediction accuracy and AUC, indicates that generally the SVM model (Table 5-29) was superior across the verification areas, compared to the RF model (Table 5-28). However, in certain locations and in particular assessment metrics the RF Model did outperform the SVM. For example, the RF model in Newcastle recorded a higher accuracy and specificity. This highlights the fact that the performance of machine learning models may be both crop specific and location specific, when used within land assessment applications.

Table 5-27 Summary of performance metrics for the soybean biophysical scoring classification for the three verification areas

Soybean Farm	Overall	Bloodriver	Luneburg	Newcastle
n	72	15	43	14
CA (%)	58.3	73.3	48.8	71.4
Precision (%)	58.2	75.0	51.2	60.0
Sensitivity (%)	82.1	90.0	73.9	100
Specificity (%)	30.3	40.0	20.0	50.0
MCC	0.14	0.35	-0.07	0.55

Table 5-28 Summary of performance metric for the Soybean Random Forest Model for the three verification areas

Soybean Farm RF	Overall	Bloodriver	Luneburg	Newcastle
n	72	15	43	14
CA (%)	44.4	26.7	46.5	57.1
Sensitivity (%)	53.8	30.0	65.2	50.0
Specificity (%)	33.3	20.0	25.0	62.5
AUC	0.46	0.59	0.44	0.46

Table 5-29 Summary of performance metric for the Soybean Support Vector Machine Model for the three verification areas

Soybean Farm SNV	Overall	Bloodriver	Luneburg	Newcastle
n	72	15	43	14
CA (%)	52.8	40.0	60.4	42.9
Sensitivity (%)	82.1	60.0	95.7	66.7
Specificity (%)	18.2	0.0	20.0	25.0
AUC	0.52	0.68	0.53	0.46

The performance of the three methods in the Bloodriver verification area produced mixed results (Tables 5-27 – 5-29). The BSC for soybean performed well, correctly predicting 11 of the 15 observations, resulting in a classification accuracy of 73.3%, a higher accuracy than produced in Bergville model building area. The method also scored well, in terms of precision, sensitivity and overarching MCC (Table 5-27). The biophysical soybean classification produced two negative classes and of these two only one produced below average yields (Figure 5-14a).

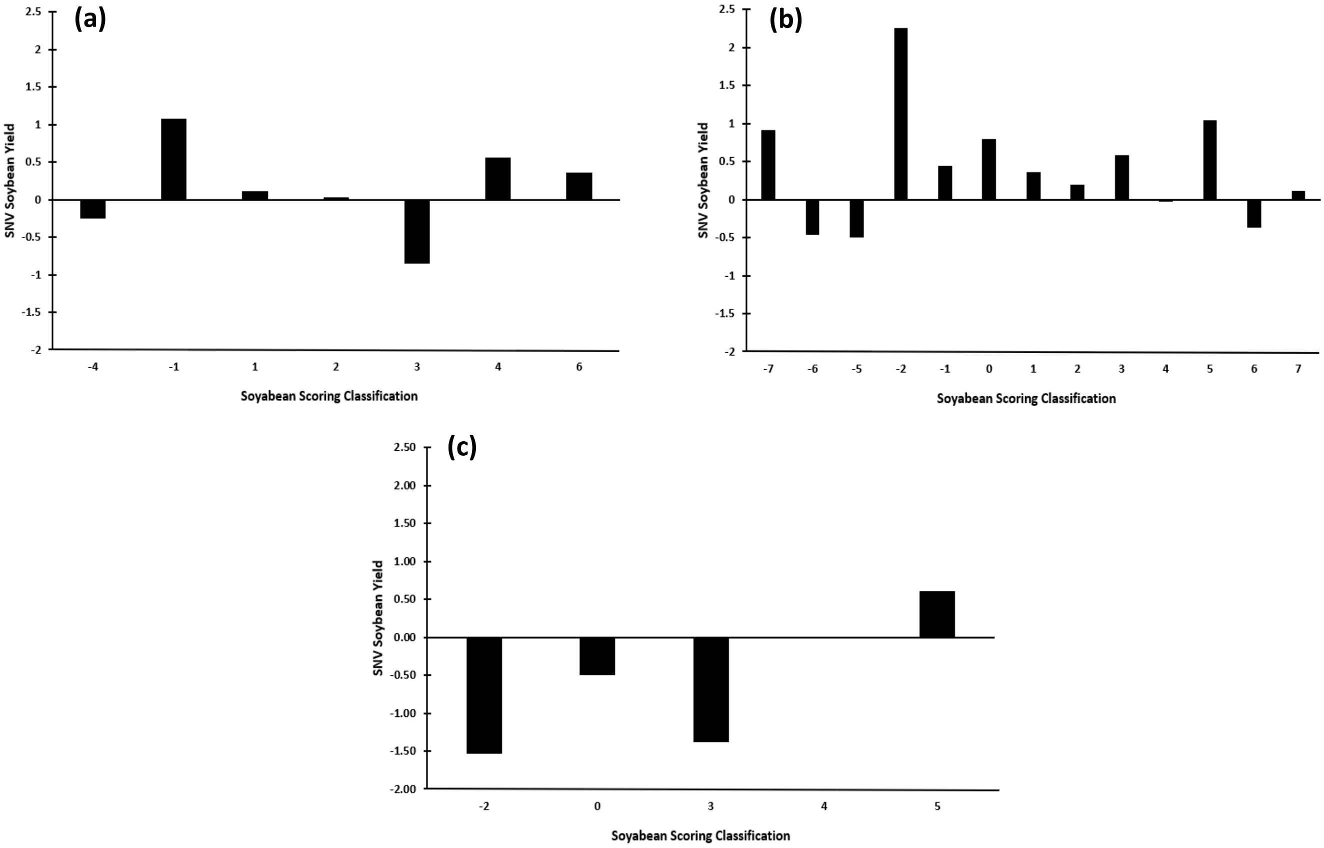

Figure 5-14 Average SNV soybean yield achieved in each biophysical scoring class for - (a) Bloodriver (b) Luneburg and (c) Newcastle

Classification accuracies in Bloodriver were high in the extreme scoring classes (-4, 4 and 6) an indication the BSC can accurately identify both extremely poor and extremely advantageous soil and terrain conditions (Figure 5-15a). Only one of the four misclassifications were a result of a False Negative event, where yield performance was underestimated by the BSC, resulting in a sensitivity value of 90%. This misclassification occurred on a mottled, sandy soil with effective rooting depths limited by the occurrence of hydromorphic features. Bloodriver is the driest verification area, consequently this poorly drained soils may not be as limiting as in wetter parts of the Province. Thus, a regional calibration of these soils will be required to improve classification accuracies. The Machine Learning models did not perform well in this environment with prediction accuracies below 50% and very low specificity results, an indication that both models could not accurately classify negative events. Ultimately, the results indicate that the soybean machine learning models are not directly transferable to this environment without additional calibration.

For the Luneberg area the SVM model was the best performer, but only achieved a moderate accuracy of 60% and low AUC of 0.53 (Table 5-29). The SVM model produced extremely high sensitivity values of 96%, indicating the model was able to detect above average soybean yields to a high degree of very accuracy. However, the SVM model as well as the BSC (Table 5-27) and RF model (Table 5.28), all scored poorly in terms of classification specificity, which in this case, is the ability of the method to correctly identify observations with below average SNV for soybean yields. This poor model performance linked to low specificity is further illustrated in Figure 5-15b which summarises yield performance for each biophysical classification scoring class, where four of the six negative scoring classes produce above average yield values. This leads to very lower classification accuracies for observations with below average yields, Figure 5-15b.

The BSC approach produced an MCC of -0.1 which indicates there is very little agreement between the soybean scoring and actual yield performance (Table 5-27). As with maize, the performance metrics for biophysical soybean classification suggest that a difference in overarching climatic variables, associated with the Luneburg area, are not sufficiently taken cognisance of in the physical soil and terrain drivers and scoring criteria. This severely reduces the transferability of the Bergville biophysical classification to the cooler and wetter Luneburg area. Ultimately, to improve model performance and better assess farm level yield variation a regionalised model or modified BSC may need to be developed in the Luneberg area.

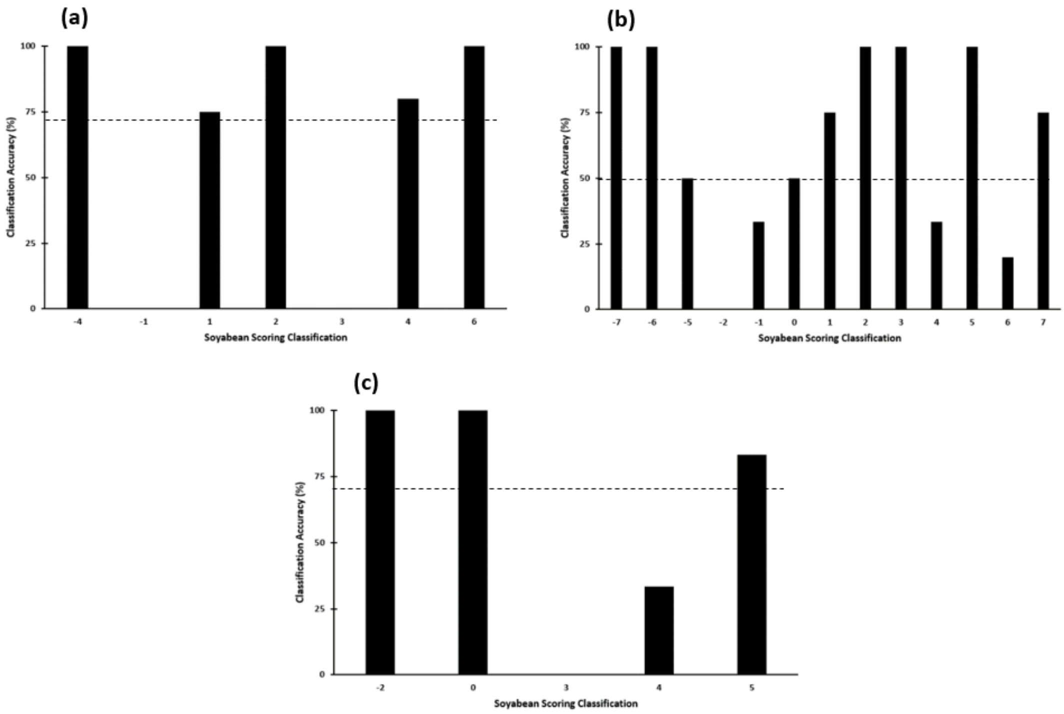

Figure 5-15 Average classification accuracy achieved in each soybean classification scoring class for - (a) Bloodriver (b) Luneburg and (c) Newcastle. The dotted line represents average classification accuracy across all scoring classes

For the Newcastle area the biophysical classification was the best performer, producing a classification accuracy of 71% and a MCC of 0.55 (Table 5-27), which is considered as a strong positive relationship between soybean scoring classification and normalised soybean yield (Mukaka, 2012). The biophysical classification was able to successfully classify all below average yield observations (Figure 5-15c). All misclassifications were caused by above average terrain and soil factors producing below average soybean yields. With biophysical soybean scoring classes 3 and 4, producing below average yields (Figure 5-14c). Low classification accuracies in high scoring classes are an indication that yields may have been reduced by non-physically related factors such as poor soil fertility management, incorrect variety selection or incorrect planting dates.

In terms of the machine learning methods RF was the marginally better performer, producing a classification accuracy of 57.2% but a low AUC of 0.46. A major source of model error is related to the 50% sensitivity metric for the RF model, which indicates that the model can only predict an above average yield event on every second occasion (Table 5-28). Based on the AUC values the ML models developed at Bergville are not transferrable to Newcastle without additional calibration.

5.3.5 Machine learning verification: contextualising maize and soybean classification performance

As previously undertaken for Bergville (Section 5.3.2.5), Provincial and National yield averages were used to benchmark and contextualise model performance for three verification areas. This was achieved by adjusting the above and below average binary yield classification by subtracting the seasonal Provincial and National yield averages (SAGL, 2018- 2020), from the observed yield point rather than the seasonal farm average. In many instances this benchmarking can substantially increase above average yield counts and equally improve prediction accuracy and other assessment metrics. To illustrate the impact of this benchmarking the RF model for maize and the SVM model for soybean were assessed using the Provincial and National Yield averages (Tables 5-30 – 5-33).

The aim of both ML models is to predict an above average yield event using a suite of attributes at a particular location. At a farm level, where average seasonal farm yields are used, this a far more difficult task. Essentially the model predicts whether a particular observation will produce an above average farm yield. At this level of predication, the RF model only obtains an average accuracy of 56.8% across the three verification farms. If the prediction goal posts are shifted to predict whether the same observation will produce above average yields from a Provincial perspective, then the model has much greater chance of obtaining a true prediction, as most actual yield observations are above average Provincial yields (Table 5-30).

Table 5-30 Summary of performance metrics for the Maize Random Forest Model for the three verification areas using KwaZulu-Natal Yield Benchmarks. Changes in values compared to farm level yields are bracketed.

	Overall	Bloodriver	Luneburg	Newcastle
n	234	44	68	122
CA	74.3 (+17.5)	45.5 (-20.4)	67.6 (+13.2)	88.5 (+33.6)
Sensitivity	77.4 (-1.8)	40.0 (-12.0)	74.1 (-3.0)	92.7 (+2.0)
Specificity	50.0 (+18.8)	100.0 (+15.8)	30.0 (-0.3)	50.0 (+36.0)
AUC	0.91 (+0.44)	0.96 (+0.23)	0.87 (+0.34)	0.92 (+0.51)

The chance of a successful prediction is even higher if the prediction benchmark is set to a National yield average (Table 5.31). For example, for the 234 total maize yield observations, across the three verification farms, only 16 observations were below average. Essentially the ML models will predict, to a very high degree of accuracy, that a particular observation will produce a yield above the National average. This can be clearly observed in the Newcastle area when the prediction level is set using a National average (Table 5.31). Based on the areas advantageous resources in terms a soil, terrain and climatic properties, for maize production, the RF model can, at a 91.8% accuracy, predict that an observation will produce an above average National yield. This prediction accuracy however, even when using National yields, is far lower in more marginal areas such as Bloodriver.

Table 5-31 Summary of performance metrics for the Maize Random Forest Model for the three verification areas using the National Yield Benchmark. Changes in values compared to farm level yields are bracketed.

	Overall	Bloodriver	Luneburg	Newcastle
n	234	44	68	122
CA	76.9 (+20.1)	40.9 (-25.0)	73.5 (+19.1)	91.8 (+36.9)
Sensitivity	77.5 (-1.7)	38.2 (-13.9)	75.8 (-1.3)	92.9 (+2.2)
Specificity	68.8 (+37.6)	100.0 (+15.8)	50.0 (-19.7)	75 (+61.0)
AUC	0.97 (+0.5)	0.97 (+0.24)	0.93 (+0.40)	0.99 (+0.58)

Similar prediction trends occur in the soybean results for the SVM ML. Where prediction accuracies are significantly increased at a Provincial and National yield benchmarks (Tables 5-32 and 5-33) for Bloodriver and Luneburg, high yielding soybean farms but reduced in Newcastle where yields were erratic. The results indicate that RF and SVM ML models produced at Bergville can be used to assess other high potential resource areas, with the aim of determining whether they will produce above average Provincial and National averages.

Table 5-32 Summary of performance metrics for the Soybean Support Vector Machine Model for the three verification areas using KwaZulu-Natal Yield Benchmarks. Changes in values compared to farm level yields are bracketed.

Soybean KZN SNV	Overall	Bloodriver	Luneburg	Newcastle
n	72	15	43	14
Accuracy	66.7 (+13.9)	66.7 (+26.7)	81.4 (+21.0)	21.4 (-21.5)
Sensitivity	83.0 (-0.9)	71.4 (+11.4)	91.7 (-4.0)	33.3 (-33.4)
Specificity	21.2 (+3.0)	0 (0)	28.6 (+8.6)	18.2 (-6.8)
AUC	0.78 (+0.26)	0.97 (+0.29)	0.87 (+0.34)	0.13 (+0.33)

Table 5-33 Summary of performance metrics for the Soybean Support Vector Machine Model for the three verification areas using National Yield Benchmarks. Changes in values compared to farm level yields are bracketed.

Soybean RSA SNV	Overall	Bloodriver	Luneburg	Newcastle
n	72	15	43	14
Accuracy	80.6 (+27.8)	73.3 (+33.3)	93.0 (+32.6)	50.0 (+7.1)
Sensitivity	84.6 (+2.5)	73.3 (+13.3)	92.7 (-3.0)	66.7 (0)
Specificity	42.9 (+24.7)	NA	100 (+80)	20 (-5.0)
AUC	0.93 (+0.41)	1.0 (+0.32)	1.0 (+0.47)	20.0 (+0.11)

5.3.6 Biophysical scoring classification versus machine learning

Based on the various performance metrics the BSC approaches generally outperformed the ML models for both maize and soybean, particularly where farm yields were used to benchmark model performance. The BSC approach also allow for greater insight into crop performance and their inter-relationship with soil and terrain factors. Conversely the ML models are more of a black box approach, which is based purely on input data, not whether it makes “practical” sense e.g. aspect measured in radians was selected as the top attribute for soybean production, while flood hazard rating was the top attribute for maize production. This view of ML models in soil-landscape relations is shared by Rossiter (2018) who asserts that pure ML models often ignore pedological knowledge and can produce results that are difficult to interpret, misleading and wrong.

Both crop specific biophysical approaches used thirteen land assessment attributes (Tables 5.10 and 5.14), with soil group, effective depth, texture and permeability class being common to both methods. Out of the selected attributes only the geomorphons and aspect layers will need to be produced digitally, the remainder of the attributes can be determined infield, as part of a traditional land assessment survey. Whereas, if the SVM models were to be practically implemented all the attributes used in the model development would need to be collected during each survey, which is not feasible for a typical land assessor.

Overfitting of data from model building area for both the BSC approach and ML was identified as problem. Overfitting attributes to region-specific conditions reduces transferability of the models to other locales.

5.3.7 The application of the new biophysical productivity methods as part of a more holistic approach to agricultural land assessment

The previous sections focused on BSC and ML model development and their performance at a farm level. From a utility perspective, the new approaches need to have not only farm level relevance, in terms of yield performance but should also supplement and compliment land assessment analysis, ultimately informing release applications in terms of Act 70 of 1970. The BSC approach for maize and soybean was selected to demonstrate how these new approaches can be used to assess arable land in release applications.

Agricultural professionals who assess farms for possible release are often faced with a scenario where the current land use either is not actively cultivated or does not correlate to its actual potential, such low intensity grazing on arable land. In these situations, the agricultural significance of the farm is based solely on the results of the selected land assessment methodology as the intensive land use is not being applied and production performance is not available. In South Africa the RSA LC (Scotney et al., 1991) remains the standard methodological approach to assess land at a farm scale (cf Chapter 2.7.2).

A summary of RSA LC classification for all soil observations located within arable classes, across both the model building and verification areas is provided Table 5-34. The arable land classified across the various farms is a mix of Class III and IV land. According to the RSA LC manual (Scotney et al., 1991) land in capability class III "*has severe permanent limitations that restrict the choice of alternative uses and the intensity of crop production and is of moderate potential*". While land in Class IV is defined as having "*very severe permanent limitations*", its use is for cultivated crops is "*largely restricted*" and is only suitable for "*occasional cultivation*". Based on yield results presented in the previous result sections the land capability classifications and associated descriptions do not convey the true message of the high production potential of these farms but rather focuses on their negative land characteristics.

Table 5-34 Soil observations per South African land capability class

RSA LC Class	Bergville		Bloodriver		Luneburg		Newcastle	
	n	%	n	%	n	%	n	%
III	234	71	0	0	46	81	45	96
IV	96	29	20	100	11	19	2	4

If these farms were virgin areas with no history of cultivation or yield records could the land assessor fully justify keeping these arable areas for agricultural production given the land capability classification results (Table 5-34) and associated definitions? Similarly, given the land capability results and definitions can the decision-making authority (DALRRD) justify the preservation of these agricultural lands if the proposed development is socially beneficial, for example rural housing for a previously disadvantaged community or an economically significant development, such as a new industrial park?

Land release statistics (Table 5-35), as published in the Draft Policy Document for the Preservation and Development of Agricultural Bill (DAFF, 2015), suggests not. The statistics indicate that up until 2011, the highest proportion, a total of nearly 2 million hectares, or 54% of all agricultural land released to non-productive uses comes from the exact land capability classes identified in the survey areas, Classes III and IV. Further illustrating that these classes are the most at risk to land use change and as an assessment method land capability alone, cannot adequately protect agricultural land.

Table 5-35 Summary of agricultural land permanently converted to non-agricultural land uses per Land Capability Class up to the year 2011, as replicated from DAFF (2015)

Land Capability Class	Total (ha)	Permanently Converted	Remainder
I	2 733	99	2 634
II	1 878 597	158 091	1 720 506
III	14 003 339	1 031 922	12 971 417
IV	16 447 446	788 505	15 658 941
V	13 609 335	254 809	13 354 526
VI	18 114 793	538 692	17 576 101
VII	45 343 216	281 774	45 061 442
VIII	12 279 370	85 398	12 193 972
Water	246 052	-	-
TOTAL	121 924 881	3 385 343	118 539 538

To better answer these questions both the land assessor and the decision-making authority require supplementary approaches such as the new BSC, to better reflect production and ultimately make more informed decisions regarding the release of agricultural land.

In order to convert the BSC from a farm specific method to a broader land assessment method a yield based cut off score was determined from the model building area in Bergville. The annual KZN average maize yield between 2016 and 2020 was selected as an aggressive benchmark for farm viability in the Bergville area. In spite of this high benchmark only 7% of the observations, in cultivated areas, produced maize yields below this benchmark. For soybean this increases to 20% for all point observations when using the average soybean yields as a benchmark. These percentages where then used to guide the selection of specific biophysical scores for the model building area.

For maize a score equal to and less than -2 incorporated the lowest 7% of the observations, while a -3 score for soybean incorporated the lowest 20% soybean of the observations. In this instance observations obtaining a score equal or less than these values were considered to be marginal for commercial dryland maize or soybean crop production. For comparability these scores were then used to classify all soil observation points surveyed in dryland areas, considered arable by the RSA LC classification.

The results for maize (Table 5-36) provide a far more realistic picture of agricultural potential for the study areas. For Bergville, Luneberg and Newcastle the vast majority (> 84%) of soil observations scored above -2 in terms of the maize scoring classification. Indicating that over 84% of arable soils are considered viable for commercial maize production using the selected Provincial benchmark. Bloodriver, which has far poorer soil resources for maize production still indicates that 60% of arable areas are considered commercially viable. For soybean, these ratios actually increase, with over 80% of all soil observation across the three verification farms being commercially viable for soybean production (Table 5-37). Importantly, these values would all increase if National yield averages or an economically determined break-even benchmark was implemented. This example however demonstrates how a production-based approach assessment method can more realistically convey actual agricultural value. It also demonstrates that the new BSC approaches are adaptable to various assessment scenarios.

Table 5-36 Summary of selected biophysical maize scores for all arable, dryland soil observations

Biophysical	Bergville		Bloodriver		Luneburg		Newcastle	
Maize Score	n	%	n	%	n	%	n	%
≤ -2	30	9	8	40	9	16	5	11
> -2	300	91	12	60	48	84	42	89

Table 5-37 Summary of selected biophysical soybean scores for all arable, dryland soil observations

Biophysical	Bergville		Bloodriver		Luneburg		Newcastle	
Soybean Score	n	%	n	%	n	%	n	%
≤ -3	79	24	3	15	10	18	1	2
> -3	251	76	17	85	47	82	46	98

When placed within the context of an agricultural assessment for a land release application, the BSC should be used as supplementary assessment to better reflect production potential. The proposed assessment methodology and workflow diagram is provided in Figure 5-16. This stepwise workflow diagram indicates that production-based approaches, such as these, should be applied after broad suitability and arability has been established using climate, soils and terrain. If the proposed workflow process is applied at FCL Farming in Bergville the assessment report would state the following:

"*Of the total assessment area, 86% is considered arable using soil, terrain and climate capability as outlined in the RSA LC system. Of these arable area 91% is considered viable for commercial maize production, while 76% is considered viable for commercial soybean production*."

Compared to the earlier class descriptions associated with the RSA LC this new description, linked to the results of the BSC approaches, assists both the land assessor to better classify the production potential of the land, as well as the decision-making authority to justify preserving more land for agricultural purposes in these threatened capability classes.

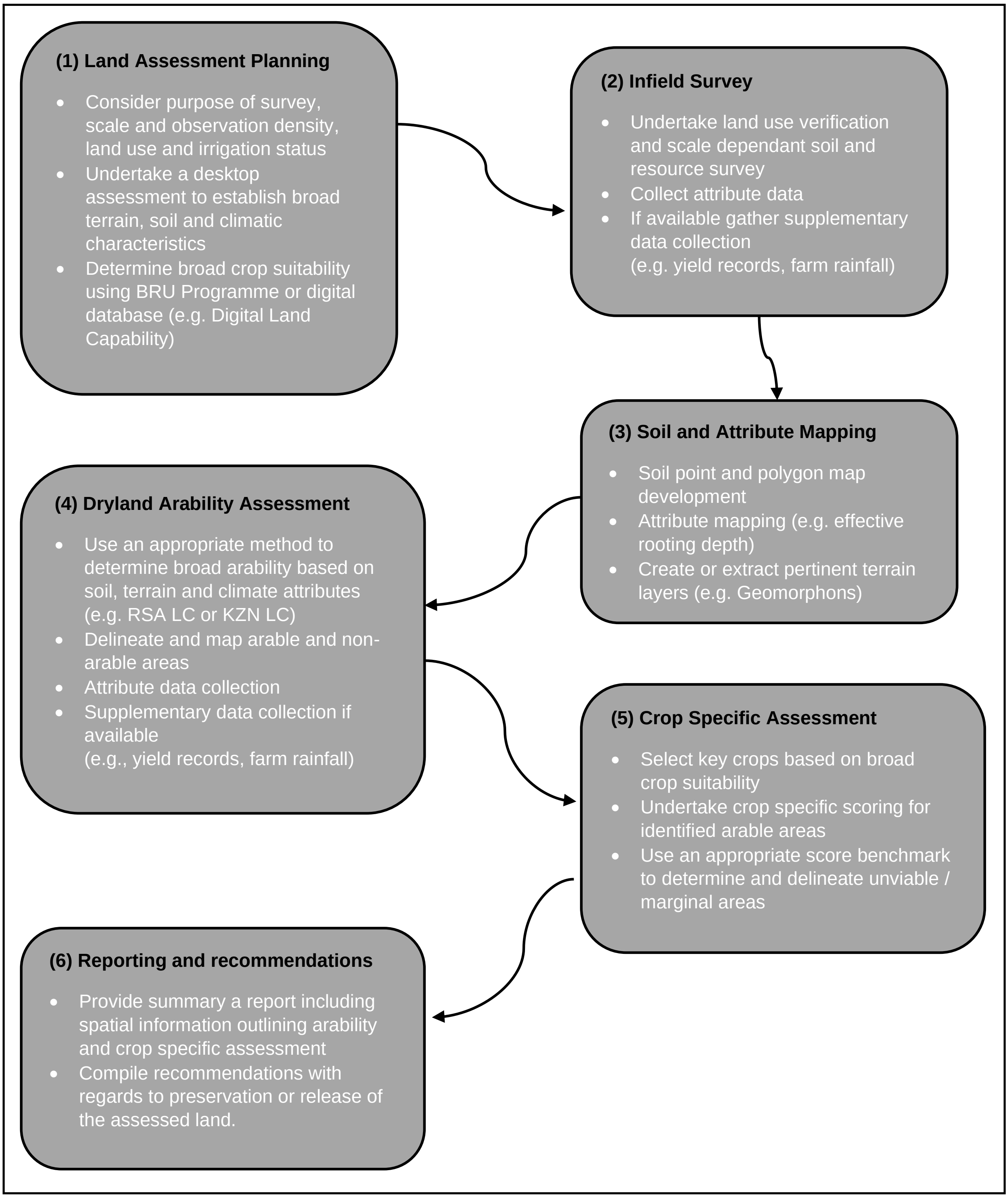

Figure 5-16 Proposed agricultural assessment process diagram

5.3.8 Additional applications: Addressing the yield gap and removing low potential areas from production

The primary aim of this Chapter was to develop methods to enhance our understanding of land performance and crop specific production drivers, ultimately improving our decision-making processes with regards to the release or protection of agricultural land. However, the methods developed for this purpose have also had useful supplementary applications in the commercial farming environment. First, a well distributed landscape soil survey, using a sample density of approximately 1:10 ha, can provide an accurate representation of farm performance, reducing the need of expensive precision scale surveys. Second, the combination of normalised yield values and crop-specific, physically based evaluation systems, such as the BSC, can be used to identify underperforming areas and yield gaps. A yield gap is defined as the difference between maximum land potential, using the best available crop genetics and technologies, and actual productivity (Godfray et al., 2010). In a given climatic area crop performance is influenced by a variety of factors including soil fertility, physical soil and land attributes as well as overarching land management (Nkurunziza et al., 2020).

Soil fertility and land management factors, such as planting dates and variety selection are reasonably dynamic with Omer et al., (2018) finding significant seasonal difference in nitrate-nitrogen, extractable potassium and extractable phosphorus contents. However, physical soil and land attributes, such as slope, soil type, soil texture and effective rooting depth are relatively static. It is these static physical attributes that provide a baseline for production potential, which actual productivity can be measured against, the intension of highlighting soil fertility and land management factors that potentially could be depressing crop yields. Section 5.3.3 highlights the applicability of yield gap analysis in Newcastle survey area, where high potential areas identified by the physically based maize scoring system was able to ring fence areas which consistently underproduced. In this case unaddressed subsoil acidity was found to be the primary cause of the yield gap, resulting in loss of maize yields of approximately 4 t/ha^{-1}, across 80 ha of affect field, over the three seasons. The estimated average maize grain price across the three seasons was R 2 400.00 per tonne (SAGL, 2018a-2020a), equating to a total loss of income of R 768 000.00 (+/- USD $ 46 000) across the 80 ha. The application of the newly developed scoring system, with a comparatively rapid and inexpensive survey, can certainly assist commercial farmers to identify underperforming areas and close the yield gap.

Another application of the maize and soybean BSC approaches is to highlight areas where extreme physical limitations reduce SNV to such an extent that they should be removed from production. This has already been successfully applied to FCL Farming in Bergville, where a seasonal break-even yield was established and used as an input into the SNV calculations. All consistently negative SNV observations, representing areas which fail to meet the break-even yield were highlighted and where practical to do so, were removed from production.

These two examples show how these newly developed BSC approaches are applicable in real world scenarios and are already benefiting commercial farmers.

5.4 Conclusions

A suite of 78 land evaluation attributes, collated from various sources, were linked to maize and soybean yields in the Bergville area, across multiple growing seasons. Three new, productivity-based land evaluation approaches using a BSC, RF and SVM were developed and tested on three verification farms, located across northern KZN.

The following conclusions were drawn with regards the model development and their performance at a farm level: 1) Attribute selection differed between both the approaches and the selected crops. 2) The BSC generally outperformed ML models. 3) The performance of the ML models varied between regions and crops and neither ML model was consistently the best performer. 4) The BSC approach was able to identify observations associated either extremely poor or extremely advantageous soil and terrain attributes, these conditions were associated with high CA rates. 5) The transferability of the models to other regions with different resources produced mixed results, highlighting the need for wider calibration in some instances. 7) Poor soil fertility and overarching land management, which was assumed to optimised at this production level, can override physical soil and terrain attributes in terms of being the primary yield determinate, detracting from the utility of these approaches 8) Farm specific yield performance can be contextualised by using provincial and national benchmarks, which can increase specific model performance metrics such CA.

The new productivity-based approaches can also have useful applications in the commercial farm management, where the crop specific biophysical scoring approaches can identify underperforming areas and yields gaps, which can be ringfenced for appropriate interventions.

The application of the new productivity-based biophysical approaches can be used to supplement and compliment agricultural assessments as part of potential land release applications, in terms of Act 70 of 1970. The use of BSC for maize and soyabean demonstrated that defined scoring benchmarks better reflect production potential across the survey areas. Ultimately, the application of these production-based approached can assist the land assessor to better classify the production potential of the land, as well as the decision-making authority to justify preserving more land for agricultural purposes.

6. AN INTEGRATED DISCUSSION OF KEY RESULTS

This chapter provides an integrated discussion, where the major findings of this research are contextualised within broader international literature and research. For the purposes of discussion three key result areas have been expanded, namely the verification of land assessment methods, the analysis of individual land assessment attributes in a production environment and the development of the biophysical and machine learning models.

6.1 Verification of land assessment methods

The lack of recent verification studies investigating the performance of land evaluation methodologies, both locally and internationally, was identified as a major knowledge gap during the literature review. This research selected five land evaluation methods and compared their resulting classifications to land use and productivity at a farm scale. From a local perspective this study is the first of its kind. Internationally, comparative studies of this nature are also rare, with many having taking place decades ago (e.g. Anderson, 1987), with no studies being done at both a polygon and point scale nor at this level of detail, making direct comparisons difficult. It is far more common for a single method, such as VSA to be compared to actual field measurements or against a single alternative technique (e.g. Leeuwen et al., 2018; Emmet-Booth et al., 2019).

Model evaluation via verification and validation is universally acknowledged as being a critical process to provide a technically defensible basis and ultimately support the decision making process, in this case the release of agricultural land (Thacker et al., 2004). Therefore, this verification study should be viewed as a crucial first step in an overarching review of traditional land assessment methodologies in South Africa.

6.2 Analysis of individual land assessment attributes in a production environment

In Chapter 4, pertinent individual land assessment attributes such as slope and soil depth, were compared to maize and soybean production performance across five growing seasons. The results indicate that maize and soybean crops respond differently to individual land assessment attributes and that generally maize was more sensitive to poorer growing conditions than soybean.

From a physiological perspective maize differs significantly from soybean and thus one should expect differing crop specific responses to stress. Soybean is a nitrogen fixing legume which uses C3 photosynthetic pathways, while maize is an annual grass which uses C4 photosynthetic pathways, C4 plants have approximately 50% higher photosynthesis efficiency than those of C3 plants (Wang et al., 2012). Maize is also taller and if root growth is unrestricted will extend approximately 1.5 m laterally and downwards to 2.0 m or deeper (du Plessis, 2003). Soybean rooting system combines both shallow lateral roots as well as a long tap root which can extend to 1.5 m, this rooting systems improves the plant's resilience when planted in heavier textured soils as well as during dry spells, and is ultimately able to utilise water at deeper soil depths compared to maize (DAFF, 2010).

This study found that seasonal rainfall and total crop yield were highly correlated (cf Chapter 3.3.3.2), thus any individual factor influencing plant available water would likely influence crop yield e.g. soil texture and effective rooting depth (cf Chapter 4.3.4). Water related stress, is caused by extended dry spells and is further exacerbated by certain soil properties, such as shallow effective soil depths or high clay content soils. This combination of factors ultimately reduces plant available water, with international literature finding that soybean is resilient in these water scarce conditions. For example Wang et al. (2020), who reviewed long term climate and yield studies between 1961 and 2017, across China, found that during severe drought years yield losses were double for maize compared to that of soybean (cf Chapter 3.3.3.2). Comparative work done by Antonio et al. (2013) also found that maize genotypes were more affected by water related stress than soybean genotypes. Both these studies correlate to the findings of this research that found soybean plants more resilient to water related stress, caused by soil properties, than maize.

Waterlogging is one factor which causes plant stress due to hypoxia rather than water-based restriction. In terms of waterlogging this research found that maize is also more sensitive to waterlogging and low soil permeabilities. This result again correlates to international tolerance guidelines as presented by Ransom & Mattern (2011) who rank soybean as a more tolerant crop to waterlogging stress, when compared to maize.

Additional to water related stress there have been numerous international studies that investigate the relationship between maize and/or soybean yield and soil and terrain properties. However, many of these studies only investigate a single crop (Silva & Silva, 2008; Takoutsing et al., 2016) and virtually all comparative studies, akin to this research, are done across very small areas < 30 ha (e.g. Kaspar et al., 2003; Kaspar et al., 2004; Marques da Silva & Alexandre, 2005), rather than the hundreds of hectares covered in this project.

Importantly, this research also combines both topographic and soil attributes, where many other studies focus on either soil or terrain attributes.

In terms of terrain, slope was found to significantly impact crop yields. The results obtained in this study indicate that steeper gradients (>8%) showed a negative correlation to production across the five growing seasons for both maize and soybean. This observation corresponds to the work of Marques da Silva and Silva (2008) who also observed negative correlations between maize yield and slope, while Leuthold et al. (2022) also found that that soybean yield was similarly negatively correlated to slope gradient. Other pertinent terrain attributes, such as erosion and flood hazards, which were also investigated, are locally derived compound factors, which makes direct comparison with international literature difficult.

Like terrain, there many studies comparing soil properties to maize and/or soybean yield. However, this study is unique insofar as it only focuses on physical soil properties linked to pertinent land assessment methodologies. Other studies in this research area combine both physical and chemical soil properties. For example Kaspar et al. (2004) uses a combination of A horizon depth, carbonate depth, pH, coarse sand, sand, silt, clay, organic C, N, Fe, K, P, and Zn. Importantly, this research is focussed on attributes that can be rapidly determined for use in agricultural land assessment, consequently assessment attributes that require extensive precision sampling, laboratory analysis or secondary modelling were avoided.

6.3 Development of biophysical and machine learning models

In this study three models, two ML models and one biophysical model were developed to predict crop performance using a suite of 78 land evaluation attributes. These models are novel for farm level land evaluation in South Africa and outside the discussions already provided in Chapter 5, a comparison with wider literature is difficult. However, *The Revised Storie Index for Use with Digital Soil Information* (O'geen et al., 2008), similarly uses a physical driven approach where fertility is ignored but utilises rating curves for attributes such as slope, depth and texture, rather than the productivity-based biophysical scoring classification used in this study. Critically this revised index uses a multiplicative scoring system and does not include micro-relief features, features which this research has found to significantly influence yield at a farm level.

Although direct comparisons with wider literature is difficult it remains important to determine where these models' fit within accepted international frameworks. An adapted framework presented by Bouma (1999), provides a three dimensional conceptual arrangement to categorise land evaluation methods (Figure 6-1). In this framework the first two dimensions on the horizontal plane are (1) the degree of computation, ranging from qualitative to quantitative; (2) the descriptive complexity of the model, ranging from empirical to mechanistic; while the third dimension, represented by the vertical axis, is the scale of the processes being modelled (Rossiter, 2003).

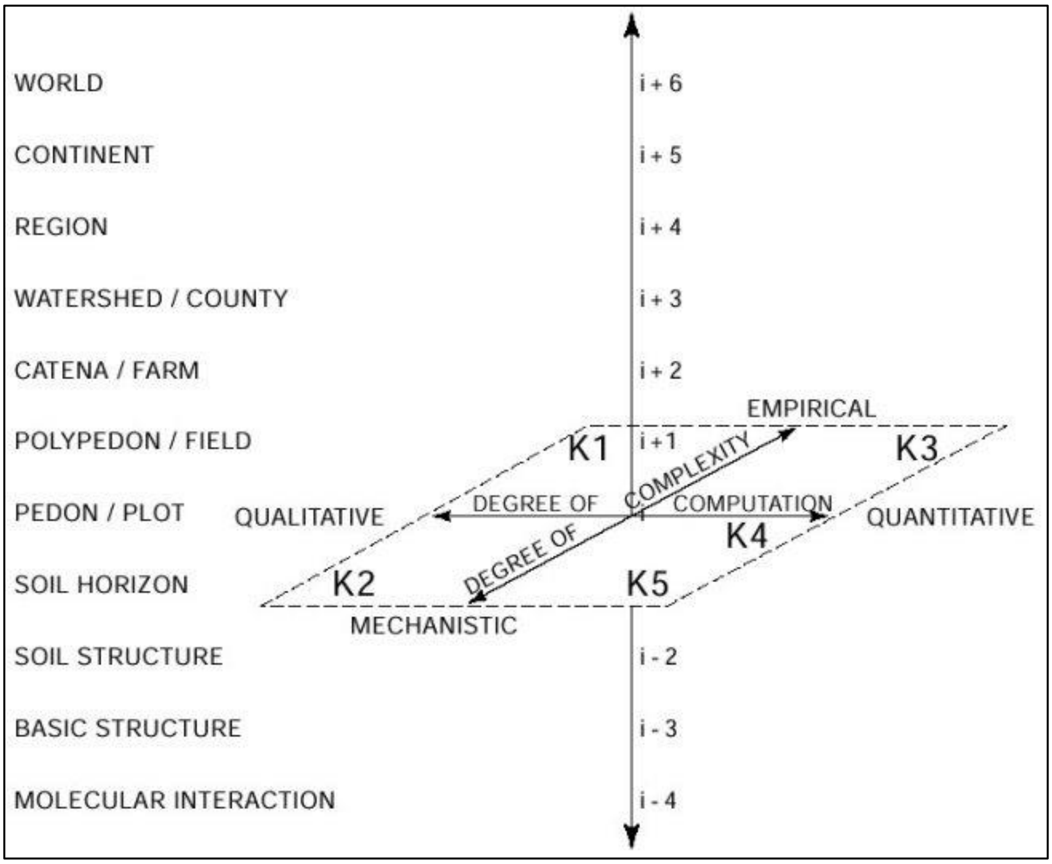

Figure 6-1 A conceptual framework to classify land evaluation models (Bouma, 1999)

When, applying this framework the models developed in this study would fall under the "K3" type models, which can be applied at a farm scale (i + 2). According to Rossiter (2003), *K3* models are empirical but quantitative where statical relations between yield and pre-determined attributes are established using large datasets and should only be used within the area of calibration. This *K3* classification and scale of application are viewed as suitable descriptions for the methods developed in this research.

The farm level applicability of these methods is an important feature of this research as it ultimately aims to supplement farm level assessments, the scale at which release applications are made in South Africa. This differs from other land evaluation studies which are generally applied at much larger scales (e.g. Gruszczyński & Gruszczyński, 2022) and are often built on national databases (e.g. Hudson & Birnie, 2000). The lack of productivity based land assessment methodologies at a farm scale is recognised by Mueller et al. (2010), whose review concluded that "*a common internationally applicable method providing field soil productivity ratings is required but does not exist*". It is hoped that this research will assist in this regard.

Although the newly developed models and their proposed sphere of application is unique to this research, all applied methodologies are founded on sound and accepted scientific research. Consequently other crop specific models, following a similar methodology could be developed, ultimately improving the spatial applicability, versatility and comprehensiveness of these new production-based approaches in South African land evaluation.

7. GENERAL CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

Agricultural land in South Africa is under pressure, to not only produce more food to meet population growth rates but is also threatened by unsustainable land use change, to non-productive uses. To ensure pertinent agricultural land is preserved for food, fibre and fuel production, assessment methods need to be geared for such purposes. Unfortunately, in recent decades the development of revised or novel land evaluation methodologies has stalled for South African farm-level assessments, the scale at which land release decisions are made. Ultimately, regular scientific validation, review and advancement is critical, to ensure the methods being utilised in practice reflect actual agricultural production levels and serve their intended purpose.

7.1 Aims and objectives revisited

The primary aims of this research, as outlined in the introductory chapter are as follows:

1. Explore pertinent literature and legislation surrounding agricultural land assessment and where applicable, highlight challenges and the need for review.
2. Assess if soil and land assessment approaches, currently being practised in industry, reflect actual land utilisation and production levels.
3. Investigate and quantify the relationship between individual land assessment attributes and productivity.
4. Develop novel, locally calibrated procedures for use in a specific commercial production environment.
5. Test the utility and robustness of these approaches in different locales and for different applications.

The achievement of these aims and objectives is provided in below.

To address the first objective, a detailed literature review was undertaken to explore the primary concepts relating to agricultural land assessment and evaluation. The review found that agricultural land evaluation is a critical process in land use management and when implemented effectively can improve decision-making, optimise land use, reduce environmental degradation and improve productivity. Further, South Africa has a number of sound legislative policies and Acts, which aim to promote the sustainable use of agricultural resources. However, significant legislative overlap and paralysis exists, which has triggered an uncoordinated and inconsistent approach to decision making across various governmental departments. These administrative and legislative failures are leading to a loss of critical agricultural land and the degradation of the resource base.

South Africa is characterised by natural resource diversity. Consequently, no single or universal method should be relied upon to evaluate all possible scenarios emanating from agriculturally based assessment and land use planning. The review of pertinent literature also found that there is a need for local verification studies, to analyse the performance of land assessment methodologies currently been practiced in industry. Additionally, local assessment methodologies, particularly at farm level, require revision to incorporate recent pedological revisions, legislative requirements and address the current challenges facing both land use planners and agricultural scientists.

To achieve objective two, five unique land evaluation methodologies were selected to assess whether they could adequality reflect actual land utilisation and production levels using land assessment polygons. By comparing land use with broad arability it was determined that the regionally calibrated KZN LC method was consistently the best performer. Along with the KZN LC, it was determined that the RSA LC, KZN ecotope and both VSA classifications could also be used in future arability assessments in this environment. However, the DAFF LC digital product severely overestimated arability and it was recommended this method should not be applied in future farm level arability assessments.

It was found that land evaluation polygons, linked to dryland precision maize and soybean yields can provide a general overview of method performance. This production-based analysis also determined that seasonal variation of rainfall influences the relationship between land classification and yield, with physical factors becoming more apparent in drier years. The analysis highlights the danger of utilising non-crop specific methodologies, as results and seasonal trends differ significantly between maize and soybean. Further, maize yields had stronger relationship to land evaluation polygons, compared to that of soybean where significant yield differences were rarely established. It was concluded that yield performance and variation, across land evaluation methods and classes, is only explicit on or near a soil observation point where measurements are taken. Thus, a point-based verification of land assessment methodologies is required to better understand the physical drivers affecting crop performance.

Experimental variograms over five growing seasons were used to determine that an 8 m circular buffer, around each soil observation point, was suitable for representative yield extraction for both maize and soybean. To account for seasonal variation, yields across the five growing seasons were also normalised to create an intuitive Standardised Normal Values for both maize and soybean. This point-based approach generally improved the relationship between land assessment classification and production. Importantly the analysis reiterated the danger of utilising non-crop specific methodologies, as results differed significantly between maize and soybean. Overall, maize yields had stronger relationship to the various land assessment classifications, compared to that of soybean with the highest maize yields generally corresponded to the best land evaluation class or class with highest cropping potential. Importantly no method could statistically separate yields across all assessment classes. Further, no method could adequately account for soybean yield variation.

To address objective three, highly influencing individual factors used in land assessment were determined and compared to maize and soybean performance across the five growing seasons. Significant yield variation across individual factor classes was more common for maize, compared to that of soybean. The results for terrain attributes found that by combining interrelated factors, novel trends between particular crops and land evaluation factors can be established. Further, slope and terrain related factors can be used to not only determine soil conservation requirements but also crop production potential. Effective soil depth, soil texture, soil functional group, soil wetness and subsoil permeability were all found to impact crop performance. The results of the individual factor analysis indicate that maize and soybean crops respond differently to individual land assessment attributes and these differences should be taken cognisance of in crop-specific land evaluation methodologies. Methodological issues such as compound and holistic attributes, ease of attribute measurement, class break significance, attribute reliability and the potential use of newer technologies such as terrain analysis, were introduced with the of view incorporating these findings into new production-based land evaluation approaches.

To address objective four a suite of 78 land evaluation attributes, collated from various sources were used to develop three new productivity-based, land evaluation approaches using Biophysical Scoring Classification (BSC), Random Forests and Support Vector Machines. The study found that attribute selection differed between the three approaches as well as the two selected crops.

To achieve objective five these newly developed approaches were tested on three verification farms, located across northern KwaZulu-Natal. The study found that the performance of the

Machine Learning (ML) models varied between regions and crops and neither ML model was consistently the best performer. The BSC generally outperformed ML models and was particularly accurate when classifying observations associated with either extremely poor or extremely advantageous soil and terrain attributes. The transferability of the models to other regions, with different resources produced mixed results, highlighting the need for wider calibration in some instances. A deficiency of the methodological approach was identified where the soil fertility and overarching land management, which was assumed to optimised at this production level, can override physical soil and terrain attributes in terms of being the primary yield determinate. The study found that the new productivity-based approaches can also have useful applications in commercial farm management, where the crop specific biophysical approaches can identify underperforming areas and yields gaps, which can be ringfenced for appropriate interventions.

Finally, the newly developed BSC was used to demonstrate the utility of these newly development approaches in broader agricultural land release applications. The study found these new approaches better reflect production potential across the various survey areas and should be used to supplement existing methodologies in land release applications. Ultimately, the application of these production-based approaches can assist the land assessor to better classify the production potential of the land, as well as the decision-making authority to justify preserving more land for agricultural purposes.

7.2 Recommendations for future research

Although this research was able address some of the issues surrounding agricultural land evaluation in South Africa many more exist.

Although it has limitations, like all land evaluation systems, the South African System Land Capability needs to be updated to take cognisance of the new Soil Classification System (SCWG, 2018). The primary purpose of the land capability is to determine broad arability and soil conservation requirements. This is still critical for both agricultural land assessment and the conservation of soil resources. Similarly, the KZN Ecotope method and associated soil functional groups need to be updated based on the new soil forms in the new Soil Classification. Appendix D includes the first attempt at classifying the new soil forms into soil functional groups. This initial classification still requires further refinement to take into account the recognition of deeper soil materials.

This research focused on maize and soybean, however other crops are also important from a food security and/or economic perspective. Scoring criteria, following a similar methodology outlined in this study for important crops *inter alia* wheat, dry beans and sugarcane, should be undertaken. These could be used to supplement agricultural assessments, in applicable production environments. Additional research, aimed at developing a contemporary land degradation index, could also be used as a supplementary assessment methodology to ensure vulnerable resources are not comprised. As the popularity of precision agriculture increases so will the spatial extent of precision yield data and its associated record length. More yield data, stretching across more growing seasons will allow scoring criteria to be refined and reduce regional overfitting.

For this research terrain is linked back to soil point observation and buffered yield. However, products derived from digital elevation model products can also be used at a landscape level, increasing the amount of yield points included in terrain analysis. Research investigating separate spatial scales for soil and terrain data could improve yield correlations.

Compared to traditional land assessment methodologies, ML and Digital Soil Mapping (DSM) is still in their relative infancy. It is recommended that additional ML and DSM research be applied in the field of both pedology and land assessment. Finally, the lack of precision yield data was identified as a significant bottleneck for model development and verification. This could be overcome by applying remote sensing technologies, such as the Sentinel-derived Products (*www.sentinel-hub.com*), to estimate yields in areas not serviced by precision harvesters.

8. REFERENCES

Akinyemi, B. E. & Mushunje, A. (2019). Land ownership and usage for agriculture: Empirical evidence from South African Living Conditions Survey. *Cogent Social Sciences*, *5*(1).

Andreev, D. V. (2020). The use of GIS technology in modern conditions. *IOP Conference Series: Earth and Environmental Science*, *421*(4).

American Farm Land Trust (2006). *Land Evaluation and Site Assessment; Fact Sheet* (p. 2). American Farmland Trust.

Anderson, L.T. (1987). *Seven methods for calculating land capability/suitability*. Planning Advisory Service (PAS), Report No.402, American Planning Association, Chicago, USA.

Anikwe, M. A. N. (2000). *Amelioration of a heavy clay loam soil with rice husk dust and its effect on soil physical properties and maize yield*. *Bioresource Technology*,*74* (2), 169–173.

Antonio, H. Joris, W. & Caires, E. F. (2013). *Effects of soil acidity and water stress on corn and soybean performance under a no-till system Effects of soil acidity and water stress on corn and soybean performance under a no-till system*. *Plant Soil,* 365, 409–424.

Armstrong, R. A. (2014). *When to use the Bonferroni correction*. *Ophthalmic & Physiological Optics 34*, 502–508.

Asgarzadeh, H., Mosaddeghi, M. R., Mahboubi, A. A., Nosrati, A., Dexter, A. R. (2010). Soil water availability for plants as quantified by conventional available water, least limiting water range and integral water capacity. *Plant and Soil, 335,* 229–244*.*

Atkinson, P. M., Tate, N. J., Atkinson, P. M., Tate, N. J. (2010). *Spatial Scale Problems and Geostatistical Solutions :* A Review. *The Professional Geographer, 52 (4),* 607-623.

Bagheri Bodaghabadi, M., Martínez-Casasnovas, J. A., Khakili, P., Masihabadi, M. H., Gandomkar, A. (2015). Assessment of the FAO traditional land evaluation methods, A case study: Iranian Land Classification method. *Soil Use and Management*, *31*(3), 384–396.

Bagherzadeh, A., Ghadiri, E., Reza, A., Darban, S., Gholizadeh, A. (2016). Land suitability modeling by parametric-based neural networks and fuzzy methods for soybean production in a semi-arid region. *Modeling Earth Systems and Environment*, *2*(2), 1–11.

Bayat, H., Asghari, S., Rastgou, M., Reza, G. (2020). Estimating Proctor parameters in agricultural soils in the Ardabil plain of Iran using support vector machines, artificial neural networks and regression methods. *Catena*, *189* (1), 104467.

Beehner, L. (2006). Land Reform Revisited. *African Renaissance*, *3*(2), 22–26.

Beek, K. J. (1978). *Land evaluation for agricultural development*. ILR1 Publication 23. ILRI, Wageningen.

Bibby, J. S., Mackney, D. (1969). *Land-use capability classification*. *Technical Monograph No. 1: Soil surveys of England and Wales and Scotland*. Harpenden: Rothamsted Experimental Station, and Aberdeen: Macaulay Institute for Soil Research

Blackmore, S., Moore, M. (1999). Remedial Correction of Yield Map Data. *Precision Agriculture*, *1*(1), 53–66.

Böhner, J., & Selige, T. (2006). Spatial prediction of soil attributes using terrain analysis and climate regionalisation. In J. Böhner, K. R. McCloy, & J. Strobl (Eds.), *SAGA–Analyses Applications, and Modelling*, Vol. 115, 13–28.

Boroumand, A., Rajaee, T., Masoumi, F. (2018). Semivariance analysis and transinformation entropy for optimal redesigning of nutrients monitoring network in San Francisco bay. *Marine Pollution Bulletin*, *129*(2), 689–694.

Bouma, J. (1999). Land evaluation for landscape units. In Sumner M. E. (Ed.), *Handbook of soil science* (pp. E393–E412). CRC Press.

Bradley, A. E. (1997). The use of the area under the ROC curve in evaluation of machine learning algorithms. *Pattern Recognition, 30(7)*, 1145–1159.

Breiman, L. (2001). Random Forests. *Machine Learning, 45,* 5–32.

Brink, A. (2015). *Subdivision – end of the road? Options to purchase subdivision of agricultural land*. Retrivied from https://www.derebus.org.za/subdivision-end-of-the-road-options-to-purchase-subdivision-of-agricultural-land/

Brownlee, J. (2016). *What is a Confusion Matrix in Machine Learning*. Retrived from https://machinelearningmastery.com/confusion-matrix-machine-learning/

Bryant, C. (2017). What Can Agricultural Land Use Planning Contribute to Food Production and Food Policy? *International International Journal of Avian & Wildlife Biology*, *2(1)*, 13-20.

Buol, S. W. (1972). *Agronomic-Economic Research on Tropical Soil: Annual Report for 1971*, 45–50.

Buol, S. W. (1986). *Fertility capability classification system and its utilization. Soil Management Under Humid Conditions in Asia and Pacific—ASIALAND* (pp. 318–331). IBSRAM, Bangkok.

CADI. (1990). *Soyabeans in Natal*. Cedara Agricultural Development Institute, Cedara.

CADI. (1993). *Maize in Natal*. Cedara Agricultural Development Institute, Cedara.

Camp, K. G. T. (1999). *A bioresource classification for KwaZulu-Natal, South Africa*. *January*, 140. *Department of Agriculture, Pietermaritzburg, South Africa.*

Camp, K. G. T., Gerrard, A. D., Manson, A. D., Milborrow, D. J., Schröder, R. W., Smith, J. M. B., & Webster, R. M. (1998). *Farm Business Planning in KwaZulu-Natal*. Cedara: KwaZulu-Natal Department of Agriculture.

Camp, K. G. T., Manson, A. D., Smith, J. M. B. (1995). *Soil Potential and Land Capability.* Cedara: KwaZulu-Natal Department of Agriculture.

Chen, W., Li, H., Hou, E., Wang, S., Wang, G., Panahi, M., Li, T., Peng, T., Guo, C., Niu, C., Xiao, L., Wang, J., Xie, X., Bin, B. (2018). Science of the Total Environment GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models. *Science of the Total Environment*, (*634)*, 853–867.

Chicco, D., Jurman, G. (2020). *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics, 1(6),* 1–13.

Chinene, V. R. . (2007). Land evaluation using the FAO Framework: an example from Zambia. *Soil Use and Management*, *8(3)*, 130–138.

Cisternas, I., Velásquez, I., Caro, A., Rodríguez, A. (2020). Systematic literature review of implementations of precision agriculture. *Computers and Electronics in Agriculture*, *176*(7), 105626.

Clarke, C. E., Roux, J. L., Ellis, F., Clercq, W. P. De, Waals, J. H. Van Der. (2020). Catena Pedogenesis of bleached topsoils occurring on weakly structured , high chroma subsoils in South Africa. *Catena*, *193*(5), 104634.

CLS (2015). *Land Rights Under Ingonyama Trust*. Centre for Law and Society, University of Cape Town.

Cloete, F. (1992). Comparative lessons for land reform in South Africa. *Africa Insight*, *22(4)*, 249–257.

Collett, A. (2009). *The determination, protection and management of high potential agricultural land in south africa with special reference to gauteng*. MSc Thesis. University of Pretoria, South Africa.

Collett, A. (2019). *Refined Land Capability Evaluation for South Africa.* Natural Resource Inventories and Assessments. Directorate: Land Use and Soil Management. Department of Agricuture, South Africa.

Collett, A. (2021). *Personal Communication* (Date: 24/02/2021).

Conrad, O., Bechtel, B., Dietrich, H., Fischer, E. K. (2015). *System for Automated Geoscientific Analyses (SAGA) v. 2.1.4*.

California Department of Conservation (1997). *California Agricultural Land Evaluation and Site Assessment Model: Instruction Manual,* pp. 1–32.

Cordoba, M., Balzarini, M. (2021). *A random forest-based algorithm for data-intensive spatial interpolation in crop yield mapping*. *Computers and Electronics in Agriculture, 184(03), 106094.*

Cox, M. S., Gerard, P. D., Wardlaw, M. C., & Abshire, M. J. (2003). Variability of Selected Soil Properties and Their Relationships with Soybean Yield. *Soil Science Society of America Journal*, *67*, 1296–1302.

D'Exelle, B. (2014). Representative Sample. In: *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht.

DAFF. (2010). *Soya Beans: Production Guidelines*.Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

DAFF. (2015). *Draft Policy Document on the Preservation and Development of Agricultural Land*. pp79-86. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

DAFF. (2016a). *Draft Policy Document on the Preservation and Development of Agricultural Land*: Updated. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

DAFF (2016b). *Preservation and Development of Agricultural Land Bill, Version 2*. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa. Government Printer.

DAFF. (2018a).*Land capability evaluation and classification for South Africa*. *Version 1 : 2018*. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

DAFF. (2018b). *Agricultural specialist survey specification survey standardards*. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

Dahake, P. S. (1971). *Project report on An Approach for Land Use Capability Classification for Indian Conditions*. Unpublished.

Dalal-Clayton, B., Dent, D. (1993). Methods of land evaluation and land use planning. In: *Surveys , plans and people: A review of land resource information and its use in developing countries*. International Institute for Environment and Development.

Davidson, D. A. (2002). The assessment of land resources: Achievements and new challenges. *Australian Geographical Studies*, *40*(2), 109–128.

de la Rosa, D., van Diepen, C. A. (2002). Qualitative and Quantitative Land Evaluation. In *Encyclopedia of Life Support System*. Eolss Publishers. Oxford, UK.

de Reu, J., Bourgeois, J., Bats, M., Zwertvaegher, A., Gelorini, V., De Smedt, P., Chu, W., Antrop, M., De Maeyer, P., Finke, P., Van Meirvenne, M., Verniers, J., Crombé, P. (2013). Application of the topographic position index to heterogeneous landscapes. *Geomorphology*, *186*, 39–49.

Delgado, R., Tibau, X. (2019). Why Cohen's Kappa should be avoided as performance measure in classification*. PLoS ONE, 14(9): e0222916*

Dent, D. and Young, A. (1981). *Soil Survey and Land Evaluation.* Allen & Unwin, London, England.

Department of Agriculture and Fisheries: Natal Region. (1982). *Ecotope Norm Excercise*. Natal Region, Cedara.

Diskin, P. (1997). Agricultural productivity indicators measurement guide. *Food Security and Nutrition Monitoring (IMPACT) Project.*

du Plessis, J. (2003). *Maize Production Guidelines*. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

du Toit, H. (1923). *Final report of the Drought Investigation Commission*. Cape Times Ltd, govt. printers.

DWAF. (2008). *Updated Manual for the Identification and Delineation of Wetlands and Riparian Areas*. Department of Water Affairs and Forestry. Pretoria, South Africa.

Egli, D. (2008). Comparison of Corn and Soybean Yields in the United States: Historical Trends and Future Prospects. *Agronomy Journal*, *100*(S3), S-79-S-88.

Eliasson, Å. (2007). *Review of Land Evaluation Methods for Quantifying Natural Constraints to Agriculture*. The Institute for Environment and Sustainability. Joint Research Centre, Ispra (Italy).

Ellis, E. C. (2011). Anthropogenic transformation of the terrestrial biosphere. *Philosophical Transactions - Royal Society of London-A*, (*369)*, 101–1035.

Emmet-Booth, J. P. Forristal, P. D. Fenton, O. Bondi, G. and Holden, N. M. (2019). Visual soil evaluation – Spade vs. profile methods and the information conveyed for soil management. Soil and Tillage Research, 187, 135–143.

ENPAT. (2002). *Environmental Potential Atlas for South Africa (GIS data)*. Pretoria: Department of Environmental Affairs and Tourism in conjunction with the University of Pretoria.

ESRI. (2016). *ArcGIS 10.5*. Environmental Systems Research Institute Inc. New York Street, Redlands, CA, USA.

Est´evez, V., Beucher, A., Mattbsck, S., Boman, A., Auri, J., Bjork, K.M., Osterholm, P. (2022). *M*achine learning techniques for acid sulfate soil mapping in southeastern Finland. *Geoderma, 406* (09), 115446.

FAO-Unesco. (1974). *Soil Map of the World* (p. 65). Unesco.

FAO. (1976). *A framework for land evaluation, Soils bulletin 32*. Food and Agriculture Organization of the United Nations. Rome, Italy.

FAO. (1983). *Guidelines: Land Assessment for Rainfed Agriculture - FAO Soil Bulletin - 52*. Food and Agriculture Organization of the United Nations. Rome, Italy.

FAO. (1983). *Guidelines: Land Assessment for Rainfed Agriculture - FAO Soil Bulletin - 52*. Food and Agriculture Organization of the United Nations. Rome, Italy.

FAO. (2007). *Land evaluation: Towards a revised framework*. Food and Agriculture Organization of the United Nations. Rome, Italy.

FOA. (2008). *Visual Soil Assessment (VSA) Field Guides* (p. 504). Food and Agriculture Organization of the United Nations.

FAO. (2020a). *Storie Index*. Retrieved from: https://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1032177/

FAO. (2020b). *Visual Soil Assessment Field Guides (VSA)*. Retrived from: https://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1197602/

FAO. (2020c). *Fertility Capability Classification (FCC)*. Retrivied from https://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1028001/

FAO. (2021). *Land Evaluation and Site Assessment (LESA)*. Retrived from: http://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1029523/

Fey, M. (2010). *Soils of South Africa*. Cambridge University Press.

Florinsky, I. V. (2012). Influence of Topography on Soil Properties. In *Digital Terrain Analysis in Soil Science and Geology* (pp. 145–149). Academic Press.

Franz, T. E., Pokal, S., Gibson, J. P., Zhou, Y., Gholizadeh, H., Amor, F., Rudnick, D., Heeren, D., Mccabe, M., Ziliani, M., Jin, Z., Guan, K., Pan, M., Gates, J., Wardlow, B. (2020). Field Crops Research The role of topography , soil , and remotely sensed vegetation condition towards predicting crop yield. *Field Crops Research*, *252*(9), 107788.

Getis, A. (2008). A History of the Concept of Spatial Autocorrelation: A Geographer's Perspective. *Geographical Analysis*, *40*(3), 297–309.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., Toulmin, C. (2010). Food Security : The Challenge of Feeding 9 Billion People. *Science*, *327*(5967), 812–818.

Green, T. R., Salas, J. D., Martinez, A., Erskine, R. H. (2007). Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma, 139(1-2)*, 23–37.

Gruszczyński, S., & Gruszczyński, W. (2022). Supporting soil and land assessment with machine learning models using the Vis-NIR spectral response. Geoderma, 405.

Guo, W., Maas, S. J., Bronson, K. F. (2012). Conductivity, topography, and Landsat imagery. *Precision Agriculture, 13(6)*, 678–692.

Guo, Z., Adhikari, K., Chellasamy, M., Greve, M. B., Owens, P. R., Greve, M. H. (2019). Geoderma Selection of terrain attributes and its scale dependency on soil organic carbon prediction. *Geoderma*, *340*(9), 303–312.

Hattingh, A. M. (2018). *Spatial variability of soil properties in two maize producing areas surveyed for precision farming in Northwest Province, South Africa.* PhD Thesis. Univerity of the Free State, South Africa.

Haverkort, B. (1988). Agricultural production potentials part I: Inherent, or the result of investments in technology development? The influence of technology gaps on the assessment of production potentials in developing countries. *Agricultural Administration and Extension*, *30*(2), 127–141.

Hawkins, C. (1989). *Agricultural Capability of Land, Tasmania. A report on a suitable system of capability classification and its application to the agricultural lands of the State*.Tasmania, Australia.

Hengl, T. (2009). A Practical guide to Geostatistical Mapping. In *Scientific and Technical Research series*.

Hong, H., Tsangaratos, P., Ilia, I., Liu, J., Zhu, A., Chen, W. (2018). Science of the Total Environment Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China. *Science of the Total Environment*, *625*, 575–588.

Huang, Y., CHEN, Z. xin, YU, T., HUANG, X. zhi, GU, X. fa. (2018). Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture*, *17*(9), 1915–1931.

Hudson, G., & Birnie, R. V. (2000). A method of land evaluation including year to year weather variability. Agricultural and Forest Meteorology, 101(2–3), 203–216.

IBM. (2021). *IBM SPSS Statistics for Windows*. IBM Corp., Armonk, N.Y., USA.

Jagtap, S. T., Phasinam, K., Kassanuk, T., Jha, S. S., Ghosh, T., & Thakar, C. M. (2021). Towards application of various machine learning techniques in agriculture. *Materials Today: Proceedings, 51 (1), 793-797.*

Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons: a pattern recognition approach to classification and mapping of landforms. *Geomorphology*, *182*, 147–156.

Jenny, H. (1941). *Factors of Soil Formation: A System of Quantitative Pedology*. Dover Publications.

Jung, J., Maeda, M., Chang, A., Bhandari, M., Ashapure, A., Landivar-Bowles, J. (2021). The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Current Opinion in Biotechnology*, *70*, 15–22.

Kaspar, T. C., Colvin, T. S., Jaynes, D. B., Karlen, D. L., James, D. E., Meek, D. W., Pulido, D., Butler, H. (2003). Relationship between six years of corn yields and terrain attributes. *Precision Agriculture*, *4*(1), 87–101.

Kaspar, T., Pulido, D., Fenton, T. E., Karlen, D. (2004). Relationship of Corn and Soybean Yield to Soil and Terrain Properties. *Agronomy Journal, 96(3)*, 700-709.

Kirkham, M. B., Grecu, S. J., Kanemasu, E. T. (1998). Comparison of minirhizotrons and the soil-water-depletion method to determine maize and soybean root length and depth. *European Journal of Agronomy 8(1-2),* 117–125.

Kiss, R. (2004). Determination of drainage network in digital elevation models, utilities and limitations. *Journal of Hungarian Geomathematics*, *2*, 16–29.

Klingebiel, A. A., Montgomery, P. H. (1961). *Land Capability Classification. Agriculture Handbook No. 210. 210*, 1–18. Soil Conservation Service, United States Department of Agriculture. Washington, United States of America.

Klompenburg, T. Van, Kassahun, A., Catal, C. (2020). Crop yield prediction using machine learning : A systematic literature review. *Computers and Electronics in Agriculture*, *177(7),* 105709.

Kollias, V. J., Kalivas, D. P., Yassoglou, N. J. (1999). Mapping the soil resources of a recent alluvial plain in Greece using fuzzy sets in a GIS environment. *European Journal of Soil Science*, *50,* 261–273.

Kotze, H., Basson, M. F. (1993). Politics-Land reform in South Africa: An overview. *Africa Insight*, *23(4)*, 190–197.

Kovačević, M., Bajat, B., & Gajić, B. (2010). Soil type classification and estimation of soil properties using support vector machines. *Geoderma*, *154*, 340–347.

Kravchenko, A.,Bullock D.G. (2000). Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties. *Agronomy Journal*, *92*(1), 75–83.

KZNDARD. (2018). *KwaZulu-Natal Agricultural Masterplan*. KZNDARD, Cedara.

Laker, M. C. (2004). Advances in soil erosion, soil conservation, land suitability evaluation and land use planning research in South Africa, 1978–2003. *South African Journal of Plant and Soil*, *21*(5), 345–368.

Land type survey staff, (1976-2006). Land type Survey Database. ARC-ISCW, Pretoria.

Larson, M. G. (2008). Analysis of Variance. *Circulation*, *117*, 115–121.

Leeuwen, M. M. W. J. Van, Heuvelink, G. B. M., Wallinga, J., Boer, I. J. M. De, Dam, J. C. Van, Essen, E. A. Van, Moolenaar, S. W., Verhoeven, F. P. M., Stoorvogel, J. J., & Stoof, C. R. (2018). Soil & Tillage Research Visual soil evaluation : reproducibility and correlation with standard measurements. *Soil & Tillage Research, 178*, 167–178.

Legg, B. J., Stafford, V. J. (1998). Precision agriculture-New Technologies. *Proceedings of the Brighton Crop Protection Conference: Pests & Diseases*, 1143–1150.

Leuthold, S. J., Wendroth, Salmerón, M., Poffenbarger, H. (2022). Weather-dependent relationships between topographic variables and yield of maize and soybean. *Field Crops Research, 276.*

Liddell, T. M., Kruschke, J. K. (2018). Analyzing ordinal data with metric models : What could possibly go wrong ? *Journal of Experimental Social Psychology, 79,* 328-348.

Ließ, M., Gebauer, A., Don, A. (2021). Machine Learning With GA Optimization to Model the Agricultural Soil-Landscape of Germany: An Approach Involving Soil Functional Types With Their Multivariate Parameter Distributions Along the Depth Profile. *Frontiers in Environmental Science, (9)*.

Liu, Y., Wang, H., Zhang, H., Liber, K. (2016). A comprehensive support vector machine-based classi fi cation model for soil quality assessment. *Soil & Tillage Research*, *155*, 19–26.

Loxton, R. F. (1962). *A simplified soil survey procedure for farm planning.* Department of Agricultural Technical Services, Pretoria.

Ludorf, R. (1970). *The land capability classification of the Natal Region: Agricultural land classification.* Department of Agriculture, Natal.

Lyle, G., Bryan, B. A., Ostendorf, B. (2014). Post-processing methods to eliminate erroneous grain yield measurements: Review and directions for future development. *Precision Agriculture*, *15*(4), 377–402.

Mandal, S., Choudhury, B. U., & Satpati, L. (2020). Soil site suitability analysis using geo-statistical and visualization techniques for selected winter crops in Sagar Island, India. *Applied Geography*, *122*(8), 102249.

Manikandan, K., Kannan, P., Sankar, M., Vishnu, G. (2013). Concepts on land evaluation. *Earth Science India*, *Popular Issue VI (I)*, 20–26.

Marqués-mateu, Á., Moreno-ramón, H., Balasch, S., Ibáñez-asensio, S. (2018). Quantifying the uncertainty of soil colour measurements with Munsell charts using a modified attribute agreement analysis. *Catena*, *171*(07), 44–53.

Marques da Silva, J., & Alexandre, C. (2005). Spatial Variability of Irrigated Corn Yield in Relation to Field Topography and Soil Chemical Characteristics. Precision Agriculture 6, 453–466.

Marques da Silva, J., Silva, L. (2008). Evaluation of the Relationship betwen Maize Yield Spatial and Temporal Variability and Different Topographic Attributes. *Biosystems Engineering*, *101*, 183–190.

Moore, I. D., Gessler, P. E., Nielsen, G. A., Peterson, G. A. (1993). Soil Attribute Selection using Terrain Analysis. *Soil Science Society of America Journal*, *57*, 443–452.

Mueller, L., Schindler, U., Mirschel, W., Graham Shepherd, T., Ball, B. C., Helming, K., Rogasik, J., Eulenstein, F., Wiggering, H. (2010). Assessing the productivity function of soils. A review. *Agronomy for Sustainable Development*, *30*(3), 601–614.

Muller C. B. (2004). n Ondersoek na die verwantskap tussen mielieopbrengste en geselekteerde grondeienskappe met behulp van opbrengskartering. Magister Technologiae: Landbou. Departement Landboubestuur. Fakulteit landbou, Tuinbou en Natuurbewaring, Tshwane Universiteit vir Tegnologie.

Mugiyo, H., Chimonyo, V. G. P., Sibanda, M., Kunz, R., Masemola, C. R., Modi, A. T., Mabhaudhi, T. (2021). Evaluation of Land Suitability Methods with Reference to Neglected and Underutilised Crop Species : A Scoping Review. *Land, 10(2)*, 21–24.

Mukaka, M. M. (2012). Statistics Corner : A guide to appropriate use of Correlation coefficient in medical research. *Malawi Med Journal, 24(*09), 69–71.

Munsell Colour, C. (1975). *Munsell Soil Colour Charts*.Munsell Colour Company, Balitmore.

Narkhede, S. (2018). *Understanding AUC - ROC Curve*. Retrieved from: https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

Nations, U. (1994). Convention on Desertification. *Information Programme on Sustainable Development.*

Newby, T., Modiselle, S., Chaminuka, P., Paterson, G., Verschoor, A. (2018). *Framework for Preservation and Development of Agricultural Land in South Africa: A Discussion Paper*. *ARC SCW Internal Report: GW/A/2018/11*. Pretoria, South Africa.

New Zealand Ministry of Works (1969). *Land Use Capability Survey Handbook*.

Nicholson, C. F., Stephens, E. C., Kopainsky, B., Thornton, P. K., Jones, A. D., Parsons, D., Garrett, J. (2021). Food security outcomes in agricultural systems models : Case examples and priority information needs. *Agricultural Systems*, *188*, 103030.

Nkonya, E. M., Karsenty, A., Msangi, S., Souza Jr, C., Shah, M., Von Braun, J., Galford, G., Park, S. (2012). Sustainable land use for the 21 century. United Nations Report. In *Sustainable land use for the 21st century*.

Nkurunziza, L., Watson, C. A., Öborn, I., Smith, H. G., Bergkvist, G., Bengtsson, J. (2020). Socio-ecological factors determine crop performance*. Scientific Reports, 4232.*

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, 24(12), 1565–1568.

O'Geen, A. T., & Southard, S. B. (2005). A revised Storie Index modeled in NASIS. *Soil Survey Horizons*, *46*(3), 98–109.

O'Geen, A. T., Southard, S. B., Southard, R. J. (2008). *Revised Storie Index for Use with Digital Soils Information*. Univeristy of California, Publication 8335.

Omer, M., Idowu, O.J., Ulery, A.L., Van Leeuwen, D., Guldan, S.J. (2018) Seasonal Changes of Soil Quality Indicators in Selected Arid Cropping Systems. Agriculture, 8 (124).

Padarian, J., Minasny, B., Mcbratney, A. B. (2019). Machine learning and soil sciences : A review aided by machine learning tools. *Soil, 6,* 35–52.

Pakawanich, P., Udomsakdigool, A., Khompatraporn, C. (2020). Robust production allocation model for an agricultural cooperative with yield uncertainty and similar revenue constraints. *Computers and Electronics in Agriculture*, *168*(01),105090.

Pease, J.R., Coughlin, R. E. (1996). *Land Evaluation and Site Assessment: A Guidebook for Rating Agricultural Lands, second ed*. Soil and Water Conservation Society, Ankeny, IA.

Pegalajar, M. C., Ruiz, L. G. B., Sánchez-marañón, M., Mansilla, L. (2020). A Munsell colour-based approach for soil classification using Fuzzy Logic and Artificial Neural Networks. *Fuzzy Sets and Systems*, *401*, 38–54.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal*, *316* (1236).

Phillips, D. L., Marks, D. G. (1996). *Spatial uncertainty analysis : propagation of interpolation errors in spatially distributed models*. *91*, 213–229.

Qian, F., Lal, R., Wang, Q. (2021). Land evaluation and site assessment for the basic farmland protection in Lingyuan County, Northeast China. *Journal of Cleaner Production*, *314*(06), 128097.

Rabia, A. H., Neupane, J., Lin, Z., & Lewis, K. (2021). Principles and applications of topography in precision agriculture. In *Advances in Agronomy* (1st ed.). Elsevier Inc.

Ramothar, A; Marais, J; Coetzee, J. (2021). *The repeal of the Subdivision of Agricultural Land Act – so close, yet so far*. Retrieved from https://www.fasken.com/en/knowledge/2021/07/26-the-repeal-of-the-subdivision-of-agricultural-land-act.

Ransom, J. and Mattern, R. (2011). Flooding/waterlogging Will Affect Crop Development. North Dakota State University: NDSU Agriculture Communication.

R Core Team. (2013). *A language and environment for statistical computing.* R Foundation for Statistical Computing. http://www.r-project.org/

Republic of South Africa. (1970). *Sub-division of Agricultural Land Act, 70 of 1970*. Pretoria: Government Printer.

Republic of South Africa. (1983). *Act No. 43 of 1983. Conservation of Agricultural Resources Act, 1983*. *April 1983*, 24. Pretoria: Government Printer.

Republic of South Africa. (1997). *Subdivision of Agricultural Land Act Repeal Bill, 1997.* Pretoria: Government Printer.

Republic of South Africa. (1998). *National Environmental Management Act, 107 of 1998.* Pretoria: Government Printer.

Republic of South Africa. (2001). *Conservation of Agricultural Resources Act, 43 of 1983. Amended Regulations 15 and 16. (Government Gazette No 22166, Vol. 429).* Pretoria: Government Printer.

Republic of South Africa. (2008). *Act 62 of 2008. Environmental Management Amendment Act , 2008. 523*(31789), 20. Pretoria: Government Printer.

Republic of South Africa. (2013). *Spatial Planning and Land Use Management Act 16 of 2013*. Pretoria: Government Printer.

Reynolds, T. W., Anderson, C. L., Slakie, E.,Gugerty, M. K. (2015). *How Common Crop Yield Measures Misrepresent Productivity among Smallholder Farmers.* Proceddings: International Conference of Agricultural Economists 2015.

Riley, S. J., DeGloria, S. D., Elliot, R. (1999). A terrain ruggedness that quantifies topographic heterogeneity. *Intermountain Journal of Science*, *5*(1-4), 23–27.

Riquier, J., Bramao, D. L., & Cornet, J. P. (1970). *A new system of soil appraisal in terms of actual and potential productivity*. Soil Resources, Development and Conservation Service, Land and Water Development Division

Robertson, G. P. (2008). *GS +: Geostatistics for the Environmental Sciences.* Gamma Design Software.

Rodrigues, R., Souza, D., Toebe, M., Sergio, V., Cargnelutti, A., Dal, A., Lúcio, C., Chuquel, A., Lima, F. De Liberalesso, G. (2022). Soybean yield variability per plant in subtropical climate : sample size definition and prediction models for precision statistics . *European Journal of Agronomy*, *136*(11), 126489.

Rodriguez-galiano, V. F., & Chica-rivas, M. (2014). Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. *International Journal of Digital Earth, 7(6),* 492-509.

Rossiter, D. G. (1988). *The Automated Land Evaluation System: a microcomputer program to assist in land evaluation*. Cornell University.

Rossiter, D. G. (1990). ALES: A framework for land evaluation using a microcomputer. *Soil Use & Management*, *6*(1), 7–20.

Rossiter, D. G. (1994). Lecture Notes : “Land Evaluation". College of Agriculture & Life Sciences. Department of Soil, Crop, & Atmospheric Sciences. Cornell University.

Rossiter, D. G. (1996). A theoretical framework for land evaluation. *Geoderma*, *72*, 165–190.

Rossiter, D. G. (2003). Biophysical Models in Land Evaluation. *Encyclopedia of Life Support Systems (EOLSS)*, *January 2003*, 1–16.

Rossiter, D. G. (2009). Book reviews. *Geoderma*, *148*(3–4), 428–429.

Rossiter, D. G. (2018). Past, present & future of information technology in pedometrics. *Geoderma*, *324*(March), 131–137.

Rossiter, D. G., & Wambeke, A. R. Van. (1997). ALES Version 4 . 65 User ' s Manual February 1997 Printing. *System*, *1*(02).

Sadler, E. J., Busscher, W. J., Bauer, P.J., Karlen, D.L. (1998). Spatial scale requirements for precision farming: A case study in the southern USA. *Agronomy Journal*, *90*, 191–197.

SAGL. (2016a). *South African Maize Crop: Quality Report 2015/2016 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2016b). *South African Soybean Crop: Quality Report 2015/2016 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2017a). *South African Maize Crop: Quality Report 2016/2017 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2017b). *South African Soybean Crop: Quality Report 2016/2017 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2018a). *South African Maize Crop: Quality Report 2017/2018 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2018b). *South African Soybean Crop: Quality Report 2017/2018 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2019a). *South African Maize Crop: Quality Report 2018/2019 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2019b). *South African Soybean Crop: Quality Report 2018/2019 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2020a). *South African Maize Crop: Quality Report 2019/2020 Season*. South African Grain Laboratory. Pretoria, South Africa.

SAGL. (2020b). *South African Soybean Crop: Quality Report 2019/2020 Season*. South African Grain Laboratory. Pretoria, South Africa.

Sadras, V. O., Calvino, P. A. (2001). Quantification of Grain Yield Response to Soil Depth in Soybean, Maize, Sunflower, and Wheat. *Agronomy Journal*, *93*, 577–583.

Sanchez, P. A., Couto, W., & Buol, S. W. (1982). The fertility capability soil classification system: Interpretation, applicability and modification. *Geoderma*, *27*(4), 283–309.

Sanchez, P. A., Palm, C. A., & Buol, S. W. (2003). Fertility capability soil classification: a tool to help assess soil quality in the tropics. *Geoderma*, *114*, 157–185.

Schiefer, J., Lair, G. J., Blum, W. E. H. (2016). Potential and limits of land and soil for sustainable intensification of European agriculture. *Agriculture, Ecosystems and Environment*, *230*, 283–293.

Schoeman, J. L., & Macvicar, C. N. (1978). A method for evaluating and presenting the agricultural potential of land at regional scales. *Agrochemophysica*, *10*(3), 25–37.

Schoeman, J. L., & Scotney, D. M. (1987). Agricultural potential as determined by soil, terrain and climate (South Africa). *South African Journal of Science*, *83*(5), 260–268.

Schoeman, J. L., Van der Walt, M., Monnik, K. A., Thackrah, A., Malherbe, J., Le Roux, R. E. (2002). *Development and application of a land capability classification system for South Africa*. *04*, 1–61.

Schulze, R. E. (1995). *Hydrology and Agrohydrology: A Text to Accompany the ACRU 3.00 Agrohydrological Modelling System.*

Schulze, R. E. (1997). *South African Atlas of Agrohydrology and -Climatology* (p. 279). Water Research Commission.

Scotney, D. M. (1971). Klassifikasie van grondvermoe - ’n basis vir boerdery bew aringsbeplanning. *Proceedings of the Annual Congresses of the Grassland Society of Southern Africa*, *6*(1), 101–107.

Scotney, D. M., Ellis, F., Nott, R. W., Taylor, K. P., Van Niekerk, B. J., Verster, E., & Wood, P. C. (1987). *A System of Soil and Land Capability for Agriculture in South Africa.* (p. 62). Report compiled by a Task Team appointed by the Multilateral Technical Committee for Agriculture and Environment Affairs (MTC/AGEN).

Scotney, D. M., Ellis, F., Nott, R. W., Taylor, K. P., Van Niekerk, B. J., Verster, E., & Wood, P. C. (1991). *A System of Soil and Land Capability for Agriculture in South Africa*. Report compiled by a Task Team appointed by the Multilateral Technical Committee for Agriculture and Environment Affairs (MTC/AGEN).

SCWG. (1991). *Soil Classification: A Taxonomic System for South Africa. Department*. Department of Agricultural Development.

SCWG. (2018). *Soil Classification: A Natural and Anthropogenic System for South Africa.* ARC-Institute for Soil, Climate and Water.

Shepherd, G. T., Stagnari, F., Pisante, M., Benites, J. (2008). *Visual Soil Assessment Field Guide: Annual Crops*. Food and Agriculture Organization of the United Nations. Rome, Italy.

Shepherd, G. T. (2010). *Visual Soil Assessment Field Guide: Maize, Part 1*. Food and Agriculture Organization of the United Nations. Rome, Italy.

Simbahan, G. C., Dobermann, A. (2004). Screening Yield Monitor Data Improves Grain Yield Maps. *Agronomy Journal*, *96*(4), 1091–1102.

Simpson, G. B., Badenhorst, J., Jewitt, G. P. W., Berchner, M., Davies, E. (2019). Competition for land: The water-energy-food nexus and coal mining in Mpumalanga province, South Africa. *Frontiers in Environmental Science*, *7*(6), 1–12.

Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing Journal*, *97*, 105524.

Singh, D., & Singh, B. (2022). Feature wise normalization : An effective way of normalizing data. *Pattern Recognition*, *122*, 108307.

Singh, S. K., Hoyos-villegas, V., Iii, J. H. H., & Fritschi, F. B. (2012). Influence of artificially restricted rooting depth on soybean yield and seed quality. *Agricultural Water Management*, *105*, 38–47.

Smith, J. M. B. (1997). *Crop, Pasture and Timber Yields: Estimates for KwaZulu-Natal*. Department of Agriculture, Cedara, Natal.

Smith, J. M. B. (2006). *The Farming Handbook*. University of KwaZulu-Natal Press.

Sonneveld, M. P. W., Hack-ten Broeke, M. J. D., van Diepen, C. A., Boogaard, H. L. (2010). Thirty years of systematic land evaluation in the Netherlands. *Geoderma*, *156(3–4)*, 84–92.

Southeastern Wisconsin Regional Planning Commission. (2007). *Land Evaluation and Site Assessment (LESA) Analysis of Farmlands in Ozaukee County. 170*.

Statistics South Africa. (2020). Census of Commercial Agriculture CoCA 2017. *Stats SA*, *April*. Retrieved from: http://www.statssa.gov.za/publications/Report-11-02-01/CoCA 2017 Fact Sheets.pdf

Steiner, F., Dunford, R., Dosdall, N. (1987). The use of the agricultural land evaluation and site assessment system in the United States. *Landscape and Urban Planning*, (*14)*, 183–199.

Storie, R. E. (1978). Storie index soil rating (revised). Special Publication 3203. *Division of Agricultural Science ,University of California, Berkeley, CA*.

Sun, W., Whelan, B., McBratney, A. B., Minasny, B. (2013). An integrated framework for software to provide yield data cleaning and estimation of an opportunity index for site-specific crop management. *Precision Agriculture*, *14*(4), 376–391.

Sys, C. (1985). *Land Evaluation: Parts I, II and III*.University of Ghent.

Takoutsing, B., Weber, J., Aynekulu, E., Rodríguez Martín, J. A., Shepherd, K., Sila, A., Tchoundjeu, Z., & Diby, L. (2016). Assessment of soil health indicators for sustainable production of maize in smallholder farming systems in the highlands of Cameroon. Geoderma, 276, 64–73.

Tesfa, T. K., Tarboton, D. G., Chandler, D. G., Mcnamara, J. P. (2009). *Modeling Soil Depth from Topographic and Land Cover Attributes Modeling soil depth from topographic and land cover attributes*. *Water Resources Research, 45 (10).*

Tesfagiorgis, G. B. (2004). *An investigation into land capability classification in Eritrea : the case study of Asmara city environs. MSc Thesis.* University of KwaZulu-Natal, Pietermaritzburg.

Taluja, C.,Thakur, R. (2018). Recent Trends Of Machine Learning In Soil Classification : A Review. *International Journal of Computational Engineering Research, 8(9).*

TIBCO Software Incorperated. (2018). *Statistica (data analysis software system)*. http://tibco.com.

Thacker, B. H., Doebling, S. W., Hemez, F. M., Anderson, M. C., Pepin, J. E., & Rodriguez, E. A. (2004). Concepts of Model Verification and Validation. Los Alamos Laboratory. Retrived from: https://www.doe.gov/bridge.

United Nations. (2019). World Population Prospects 2019: Data Booklet. Retrieved from: https://population.un.org/wpp/Publications/Files/WPP2019_DataBooklet.pdf.

USDA: Natural Resources Conservation Service. (2019). *National Soil Survey Handbook: Title 430*. Retrieved from: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054242

USDA: Soil Survey Staff. (1975). *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*. United States Department of Agriculture.

USDA. (2011). National Agricultural Land Evaluation and Site Assessment (Lesa) Handbook. *United States Department of Agriculture Natural Resources Conservation Service*, *1*, 91.

Ustaoglu, E., Sisman, S., Aydınoglu, A. C. (2021). Determining agricultural suitable land in peri-urban geography using GIS and Multi Criteria Decision Analysis (MCDA) techniques. *Ecological Modelling*, *455*(June), 109610.

van Diepen, C. A., Van Keulen, H., Wolf, J., & Berkhout, J. A. A. (1991). Land evaluation: from intuition to quantification. In B. A. Stewart (Ed.), *Advances In Soil Science* (pp. 139–204). Springer, New York.

van Niekerk, A. (2010). A comparison of land unit delineation techniques for land evaluation in the Western Cape, South Africa. *Land Use Policy*, *27*(3), 937–945.

van Niekerk, B. J. (1983). *Developments in land capability mapping with particular reference to southern Africa*. Proc. 10th Nat. SSSSA Congress.

van Zijl, G. M., & Botha, J. O. (2016). In pursuit of a South African national soil database: potential and pitfalls of combining different soil data sets. *South African Journal of Plant and Soil*, *33*(4), 257–264.

van Zijl, G., Turner, D., Paterson, G., Koch, J., van Tol, J., Barichievy, K., Clarke, C., du Plessis, M., van Deventer, P. (2020) The new soil classification system in South Africa, its history, important changes made and implications for users. *South African Journal of Plant and Soil*, 37:5, 331-342,

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer, New York.

Verheye, W. H. (2009). Land Evaluation. In W. H. Verheye (Ed.), *Land Use, Land Cover and Soil Sciences* (p. Volume II, 1-28). Eolss Publishers.

Viana, C. M., Santos, M., Dulce Freire, P. A., & Rocha, J. (2021). Evaluation of the factors explaining the use of agricultural land : A machine learning and model-agnostic approach. *Ecological Indicators*, *131*(108200).

Wang, C., Guo, L., Li, Y. and Wang, Z. (2012). Systematic Comparison of C3 and C4 Plants Based on Metabolic Network Analysis. BMC Systems Biology 6 (Supplement 2), S9.

Wang, C., Linderholm, H. W., Song, Y., Wang, F., Liu, Y., Tian, J., Xu, J., Song, Y., & Ren, G. (2020). Impacts of drought on maize and soybean production in northeast China during the past five decades. *International Journal of Environmental Research and Public Health*, *17*(7).

Wells, M. R., King, P. D. (1989). *Land capability assessment methodology for rural-residential development and associated agricultural land uses*. Department of Primary Industries and Regional Development, Western Australia, Perth. Report 1.

Wilhelm, R. C., Es, H. M. Van, Buckley, D. H. (2022). Predicting measures of soil health using the microbiome and supervised machine learning. *Soil Biology and Biochemistry*, *164(10),* 108472.

Wineman, A., Anderson, C. L., Reynolds, T. W., Biscaye, P. (2019). Methods of crop yield measurement on multi-cropped plots: Examples from Tanzania. *Food Security*, *11*, 1257–1273. 5

Xiaohu, Z., Li, J., Xiaolei, Q., Jianxiu, Q., Juan, W., Yan, Z. (2016). An improved method of delineating rectangular management zones using a semivariogram-based technique. *Computers and Electronics in Agriculture*, *121*, 74–83.

Yesilnacar, E., Topal, T. (2005). *Landslide susceptibility mapping : A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey)*. Engineering Geology, *79*, 251–266.

Zhang, L., Yang, L., Ma, T., Shen, F., Cai, Y., Zhou, C. (2021). Geoderma A self-training semi-supervised machine learning method for predictive mapping of soil classes with limited sample data. *Geoderma*, *384*(11), 114809.

APPENDIX A:

YIELD BUFFER RESULTS USING AVERAGE MEDIAN YIELD AND ASSOCIATED STANDARD DEVIATIONS AT VARYING DISTANCE INTERVALS

Maize															
	2016			2017			2018			2019			2020		
Distance (m)	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation
5	6	7.68	0.64	8	10.72	1.09	7	11.5	0.6	5	6.77	0.45	6	10.55	0.68
6	9	7.93	0.83	12	10.84	1.16	9	11.4	0.8	8	7.11	0.52	9	10.45	0.82
7	12	8.12	1.16	15	10.88	1.24	13	11.4	0.8	10	7.26	0.61	13	10.45	0.89
8	17	8.10	1.23	20	10.91	1.35	17	11.4	0.9	13	6.92	0.71	16	10.44	1.02
9	21	8.07	1.35	26	10.88	1.41	22	11.5	0.9	17	6.54	0.77	19	10.72	1.07
10	25	8.06	1.37	32	10.87	1.43	26	11.4	0.9	22	6.62	0.79	23	10.68	1.06
11	31	7.97	1.35	38	10.86	1.45	32	11.4	1.0	26	6.65	0.88	29	10.67	1.10
12	37	7.90	1.37	46	10.94	1.46	37	11.4	1.0	31	6.62	0.94	34	10.70	1.18
13	43	7.90	1.48	53	10.91	1.51	44	11.4	1.0	35	6.63	1.00	41	10.75	1.20
14	49	7.87	1.45	61	10.94	1.55	50	11.4	1.0	40	6.62	0.97	49	10.79	1.20
15	58	7.90	1.49	70	10.87	1.58	58	11.4	1.0	46	6.50	0.95	55	10.83	1.26
20	104	7.73	1.65	121	10.94	1.64	102	11.3	1.1	82	6.33	1.02	99	10.94	1.52
25	163	7.65	1.60	189	10.93	1.65	158	11.3	1.3	127	6.24	1.03	150	10.98	1.79
30	231	7.63	1.62	271	10.82	1.77	225	11.3	1.5	179	6.18	1.13	217	11.00	1.80

Soybean															
	2016			2017			2018			2019			2020		
Distance (m)	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation	n	Avg. Median Yield	Std. Deviation
5	5	2.25	0.24	8	3.34	0.32	6	3.26	0.21	5	2.53	0.20	11	3.24	0.31
6	7	2.28	0.27	11	3.36	0.43	10	3.34	0.24	7	2.47	0.23	15	3.22	0.31
7	10	2.34	0.30	16	3.30	0.44	14	3.36	0.27	9	2.50	0.26	20	3.21	0.35
8	13	2.43	0.34	20	3.23	0.44	18	3.42	0.27	12	2.50	0.28	27	3.21	0.41
9	18	2.32	0.35	25	3.20	0.44	21	3.42	0.25	16	2.50	0.30	34	3.22	0.41
10	21	2.34	0.35	30	3.18	0.47	24	3.44	0.25	21	2.49	0.30	43	3.20	0.41
11	26	2.30	0.38	36	3.25	0.47	31	3.44	0.27	26	2.50	0.31	53	3.17	0.40
12	32	2.26	0.42	43	3.28	0.47	36	3.45	0.26	29	2.47	0.32	62	3.20	0.44
13	36	2.24	0.43	51	3.26	0.52	43	3.45	0.26	34	2.48	0.32	74	3.18	0.47
14	42	2.21	0.44	60	3.31	0.53	51	3.48	0.28	40	2.49	0.34	86	3.15	0.50
15	46	2.22	0.44	68	3.29	0.53	60	3.43	0.31	47	2.54	0.36	101	3.21	0.53
20	84	2.21	0.47	118	3.18	0.60	104	3.47	0.39	81	2.56	0.41	175	3.20	0.66
25	126	2.24	0.50	180	3.14	0.64	164	3.52	0.54	123	2.59	0.47	275	3.19	0.70
30	179	2.27	0.51	257	3.17	0.65	232	3.48	0.54	172	2.55	0.52	377	3.19	0.71

APPENDIX B:

CROP ECOTOPE RESULTS FOR ALL BUFFERED YIELD POINTS (FCL FARMING)

Average median maize yield and average SNV per Crop Ecotope Class over five growing Seasons (2016-2020).

KZN Ecotope*	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)
B11	23	11.24	56.61
B12	6	10.98	38.10
B13	3	10.38	46.35
B21	107	10.33	13.43
B22	18	9.51	-38.24
B23	9	8.97	-28.20
D11	30	9.99	14.94
D12	4	10.74	44.80
D13	5	9.83	11.90
D21	79	10.07	10.39
D22	47	9.36	-19.82
D23	3	10.81	70.94
E11	3	10.90	68.04
E12	4	9.01	-48.12
E13	8	10.26	45.00
E22	10	8.40	-55.44
E23	19	9.57	-10.73
E24	4	9.97	-19.35
H13	3	7.55	-125.29
H22	2	7.55	-93.79
I13	1	7.03	-194.35
I24	1	7.44	-172.62
J13	4	8.36	-87.55
J22	3	8.30	-16.44
J23	14	8.79	-61.84
J24	8	8.52	-36.76

*Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class

Average median soybean yield and average SNV per Ecotope Class over five growing Seasons (2016-2020).

KZN Ecotope*	Buffered Point n	Buffered Point Avg. Median Yield (t.ha^{-1})	Buffered Point Avg. SNV (%)
B11	11	2.98	24.90
B12	4	3.26	102.43
B13	2	2.58	-7.91
B21	49	2.84	27.51
B22	17	2.52	-51.14
B23	9	2.98	-10.87
D11	14	2.82	9.34
D12	1	1.9	-90.53
D13	4	2.65	-60.81
D21	27	2.79	7.65
D22	34	2.62	-12.38
D23	3	3.22	112.72
E11	2	2.53	-30.63
E12	3	3.16	25.55
E13	8	3.33	23.12
E22	8	2.97	-26.83
E23	10	3.27	14.96
E24	6	2.48	-65.67
H13	2	2.07	-87.05
H22	2	3.56	42.12
I13	1	2.43	27.10
J13	1	1.86	-112.50
J22	5	2.68	-22.14
J23	13	2.60	-30.15
J24	4	3.11	-8.14

*Ecotope Code consists of Soil Functional Group. Topsoil Clay Content Class. Effective Depth Class

APPENDIX C:

CORRELATIONS BETWEEN CROP YIELD (SNV) AND TERRAIN ATTRIBUTES

Maize Correlations: Terrain Attributes

		Maize Yield (SNV)	Elevation	Aspect (Rads)	Slope Length Factor	Aspect (Degrees)	Flow Accumulation	Convergence Index	Planform Curve	Profile Curve	Relative Slope Position	Slope (Degrees)	Slope (Percent)	Terrain Roughness Index	Terrain Position Index	Terrain Wetness Index
Maize Yield (SNV)	Pearson Correlation	1	0.055	-0.022	-0.054	-0.002	-0.067	0.022	-0.055	0.052	.130**	-0.050	-0.059	-0.074	-.141**	-0.007
	Sig. (2-tailed)		0.229	0.629	0.238	0.966	0.145	0.635	0.229	0.259	0.005	0.278	0.202	0.106	0.002	0.878
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Elevation	Pearson Correlation	0.055	1	.150**	-0.055	.120**	-.142**	.148**	0.078	.147**	.467**	0.005	-0.087	-.102*	0.056	-.127**
	Sig. (2-tailed)	0.229		0.001	0.234	0.009	0.002	0.001	0.089	0.001	0.000	0.913	0.059	0.027	0.227	0.006
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Aspect (Rads)	Pearson Correlation	-0.022	.150**	1	-0.043	.621**	-0.067	-0.063	0.023	0.061	0.027	-0.009	-0.009	-0.017	-0.057	0.038
	Sig. (2-tailed)	0.629	0.001		0.347	0.000	0.143	0.170	0.616	0.183	0.555	0.847	0.847	0.718	0.214	0.415
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Slope Length Factor	Pearson Correlation	-0.054	-0.055	-0.043	1	.092*	-0.031	-.216**	-.122**	-.163**	-.122**	.750**	.726**	.680**	-.180**	-0.061
	Sig. (2-tailed)	0.238	0.234	0.347		0.046	0.497	0.000	0.008	0.000	0.008	0.000	0.000	0.000	0.000	0.182
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Aspect (Degrees)	Pearson Correlation	-0.002	.120**	.621**	.092*	1	-0.084	-0.027	0.085	0.051	0.013	.181**	.172**	.231**	0.008	-.170**
	Sig. (2-tailed)	0.966	0.009	0.000	0.046		0.068	0.555	0.065	0.272	0.773	0.000	0.000	0.000	0.856	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Flow Accumulation	Pearson Correlation	-0.067	-.142**	-0.067	-0.031	-0.084	1	-.117*	-.115*	-0.056	-.094*	-0.021	-0.053	-0.065	-0.063	0.005
	Sig. (2-tailed)	0.145	0.002	0.143	0.497	0.068		0.011	0.013	0.222	0.041	0.654	0.254	0.160	0.170	0.914
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Convergence Index	Pearson Correlation	0.022	.148**	-0.063	-.216**	-0.027	-.117*	1	.723**	.319**	.133**	-.134**	-.117*	-.122**	.326**	-.092*
	Sig. (2-tailed)	0.635	0.001	0.170	0.000	0.555	0.011		0.000	0.000	0.004	0.004	0.011	0.008	0.000	0.045
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Planform Curve	Pearson Correlation	-0.055	0.078	0.023	-.122**	0.085	-.115*	.723**	1	.206**	0.047	0.073	0.074	0.079	.286**	-.225**
	Sig. (2-tailed)	0.229	0.089	0.616	0.008	0.065	0.013	0.000		0.000	0.312	0.113	0.110	0.087	0.000	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Profile Curve	Pearson Correlation	0.052	.147**	0.061	-.163**	0.051	-0.056	.319**	.206**	1	.330**	-0.078	-.169**	-.155**	.246**	-.108*
	Sig. (2-tailed)	0.259	0.001	0.183	0.000	0.272	0.222	0.000	0.000		0.000	0.090	0.000	0.001	0.000	0.019
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Relative Slope Position	Pearson Correlation	.130**	.467**	0.027	-.122**	0.013	-.094*	.133**	0.047	.330**	1	0.029	-0.034	-0.054	0.050	-.209**
	Sig. (2-tailed)	0.005	0.000	0.555	0.008	0.773	0.041	0.004	0.312	0.000		0.527	0.454	0.237	0.278	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Slope (Degrees)	Pearson Correlation	-0.050	0.005	-0.009	.750**	.181**	-0.021	-.134**	0.073	-0.078	0.029	1	.833**	.779**	.126**	-.579**
	Sig. (2-tailed)	0.278	0.913	0.847	0.000	0.000	0.654	0.004	0.113	0.090	0.527		0.000	0.000	0.006	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Slope (Percent)	Pearson Correlation	-0.059	-0.087	-0.009	.726**	.172**	-0.053	-.117*	0.074	-.169**	-0.034	.833**	1	.944**	0.027	-.345**
	Sig. (2-tailed)	0.202	0.059	0.847	0.000	0.000	0.254	0.011	0.110	0.000	0.454	0.000		0.000	0.553	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Terrain Roughness Index	Pearson Correlation	-0.074	-.102*	-0.017	.680**	.231**	-0.065	-.122**	0.079	-.155**	-0.054	.779**	.944**	1	0.037	-.338**
	Sig. (2-tailed)	0.106	0.027	0.718	0.000	0.000	0.160	0.008	0.087	0.001	0.237	0.000	0.000		0.427	0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Terrain Position Index	Pearson Correlation	-.141**	0.056	-0.057	-.180**	0.008	-0.063	.326**	.286**	.246**	0.050	.126**	0.027	0.037	1	-.422**
	Sig. (2-tailed)	0.002	0.227	0.214	0.000	0.856	0.170	0.000	0.000	0.000	0.278	0.006	0.553	0.427		0.000
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473
Terrain Wetness Index	Pearson Correlation	-0.007	-.127**	0.038	-0.061	-.170**	0.005	-.092*	-.225**	-.108*	-.209**	-.579**	-.345**	-.338**	-.422**	1
	Sig. (2-tailed)	0.878	0.006	0.415	0.182	0.000	0.914	0.045	0.000	0.019	0.000	0.000	0.000	0.000	0.000	
	N	473	473	473	473	473	473	473	473	473	473	473	473	473	473	473

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Soybean Correlations: Terrain Attributes

		Soybean Yield (SNV)	Elevation	Aspect (Rads)	Slope Length Factor	Aspect (Degrees)	Flow Accumulation	Convergence Index	Planform Curve	Profile Curve	Relative Slope Position	Slope (Degrees)	Slope (Percent)	Terrain Roughness Index	Terrain Position Index	Terrain Wetness Index
Soybean Yield (SNV)	Pearson Correlation	1	-0.091	-.177**	-0.009	0.002	-.171**	0.046	0.087	-0.080	-.177**	-0.016	0.024	0.028	0.038	-0.074
	Sig. (2-tailed)		0.128	0.003	0.881	0.973	0.004	0.447	0.148	0.182	0.003	0.793	0.692	0.637	0.531	0.219
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Elevation	Pearson Correlation	-0.091	1	.256**	0.051	.281**	-.186**	0.032	0.051	0.096	.541**	0.106	0.044	0.047	0.094	-.199**
	Sig. (2-tailed)	0.128		0.000	0.390	0.000	0.002	0.597	0.398	0.108	0.000	0.077	0.458	0.429	0.116	0.001
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Aspect (Rads)	Pearson Correlation	-.177**	.256**	1	-0.075	.612**	-0.020	-0.010	-0.021	.119*	0.077	0.005	-0.017	-0.029	-0.007	-0.029
	Sig. (2-tailed)	0.003	0.000		0.209	0.000	0.739	0.869	0.732	0.046	0.195	0.939	0.773	0.630	0.905	0.626
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Slope Length Factor	Pearson Correlation	-0.009	0.051	-0.075	1	0.076	-.129*	-.185**	0.003	-.127*	-0.095	.772**	.726**	.659**	-.124*	-.168**
	Sig. (2-tailed)	0.881	0.390	0.209		0.203	0.030	0.002	0.961	0.033	0.111	0.000	0.000	0.000	0.038	0.005
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Aspect (Degrees)	Pearson Correlation	0.002	.281**	.612**	0.076	1	-.188**	-0.046	0.060	0.064	.136*	.178**	.188**	.229**	0.026	-.220**
	Sig. (2-tailed)	0.973	0.000	0.000	0.203		0.002	0.440	0.315	0.286	0.023	0.003	0.002	0.000	0.664	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Flow Accumulation	Pearson Correlation	-.171**	-.186**	-0.020	-.129*	-.188**	1	.149*	-0.025	0.030	-.139*	-.168**	-.212**	-.249**	-.126*	.317**
	Sig. (2-tailed)	0.004	0.002	0.739	0.030	0.002		0.012	0.677	0.618	0.020	0.005	0.000	0.000	0.034	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Convergence Index	Pearson Correlation	0.046	0.032	-0.010	-.185**	-0.046	.149*	1	.698**	.325**	0.043	-.147*	-.130*	-0.077	.329**	0.017
	Sig. (2-tailed)	0.447	0.597	0.869	0.002	0.440	0.012		0.000	0.000	0.473	0.014	0.030	0.199	0.000	0.776
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Planform Curve	Pearson Correlation	0.087	0.051	-0.021	0.003	0.060	-0.025	.698**	1	.190**	-0.072	.131*	.123*	.156**	.354**	-.191**
	Sig. (2-tailed)	0.148	0.398	0.732	0.961	0.315	0.677	0.000		0.001	0.227	0.028	0.039	0.009	0.000	0.001
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Profile Curve	Pearson Correlation	-0.080	0.096	.119*	-.127*	0.064	0.030	.325**	.190**	1	.158**	-0.071	-.169**	-.148*	.245**	0.009
	Sig. (2-tailed)	0.182	0.108	0.046	0.033	0.286	0.618	0.000	0.001		0.008	0.233	0.004	0.013	0.000	0.881
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Relative Slope Position	Pearson Correlation	-.177**	.541**	0.077	-0.095	.136*	-.139*	0.043	-0.072	.158**	1	0.031	-0.028	-0.023	.129*	-.210**
	Sig. (2-tailed)	0.003	0.000	0.195	0.111	0.023	0.020	0.473	0.227	0.008		0.602	0.639	0.702	0.031	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Slope (Degrees)	Pearson Correlation	-0.016	0.106	0.005	.772**	.178**	-.168**	-.147*	.131*	-0.071	0.031	1	.823**	.766**	.184**	-.658**
	Sig. (2-tailed)	0.793	0.077	0.939	0.000	0.003	0.005	0.014	0.028	0.233	0.602		0.000	0.000	0.002	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Slope (Percent)	Pearson Correlation	0.024	0.044	-0.017	.726**	.188**	-.212**	-.130*	.123*	-.169**	-0.028	.823**	1	.933**	0.109	-.435**
	Sig. (2-tailed)	0.692	0.458	0.773	0.000	0.002	0.000	0.030	0.039	0.004	0.639	0.000		0.000	0.067	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Terrain Roughness Index	Pearson Correlation	0.028	0.047	-0.029	.659**	.229**	-.249**	-0.077	.156**	-.148*	-0.023	.766**	.933**	1	.130*	-.450**
	Sig. (2-tailed)	0.637	0.429	0.630	0.000	0.000	0.000	0.199	0.009	0.013	0.702	0.000	0.000		0.030	0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Terrain Position Index	Pearson Correlation	0.038	0.094	-0.007	-.124*	0.026	-.126*	.329**	.354**	.245**	.129*	.184**	0.109	.130*	1	-.416**
	Sig. (2-tailed)	0.531	0.116	0.905	0.038	0.664	0.034	0.000	0.000	0.000	0.031	0.002	0.067	0.030		0.000
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Terrain Wetness Index	Pearson Correlation	-0.074	-.199**	-0.029	-.168**	-.220**	.317**	0.017	-.191**	0.009	-.210**	-.658**	-.435**	-.450**	-.416**	1
	Sig. (2-tailed)	0.219	0.001	0.626	0.005	0.000	0.000	0.776	0.001	0.881	0.000	0.000	0.000	0.000	0.000	
	N	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

APPENDIX D:

PROPOSED FUNCTIONAL GROUPS (FG) FOR THE 2018 SOIL CLASSIFICATION SYSTEM FOR SOUTH AFRICA

FUNCTIONAL GROUP CODE	
Deep humic soils	A
Well and Moderately drained Soils	B
Unconsolidated sediments	C
Mottled and moderately drained soils	D
Mottled and poorly drained soils	E
Black (Margalitic) soils	F
Black (Margalitic) poorly drained soils	G
Young soils	H
Gleyed soils	I
Duplex soils	J
Soft and/or hard carbonates	K
Dorbank	L
Man-made soils	M
Organic soils	O
Podzols	Z

PEAT TOPSOIL					
Topsoil Horizon	**Subsoil Horizon**	**Subsoil Horizon**	**Form**	**Code**	**FG**
Peat	Gley	-	**Mfabeni**	Mf	O
	Albic	-	**Nhlangu**	Nh	O
	Soft Carbonate	-	**Muzi**	Mz	O
	Hard Rock	-	**Kromme**	Ko	O

ORGANIC TOPSOIL					
Topsoil Horizon	**Subsoil Horizon**	**Subsoil Horizon**	**Form**	**Code**	**FG**
Organic	Gley	-	**Champagne**	Ch	O
	Albic	Gley	**Manguzi**	Mg	O
	Soft Carbonate	-	**Makhasana**	Mh	O
	Hard Rock	-	**Didema**	Dd	O

VERTIC TOPSOIL					
Topsoil horizon	**Subsoil Horizon**	**Subsoil Horizon**	**Form**	**Code**	**FG**
Vertic	Gley	-	**Rensburg**	Rg	G
	Pedocutanic (thick)		**Glen**	Gl	F
	Soft Carbonate	Gley	**Zondereinde**	Zo	G
	Soft Carbonate	Hard Carbonate	**Nietverdiend**	Nv	F
	Soft Carbonate	Lithic	**Bakwena**	Bk	F
	Hard Carbonate	-	**Waterval**	Wv	F
	Alluvium (thick)		**Mkuze**	Mk	F
	Lithic	-	**Arcadia**	Ar	F
	Hard Rock	-	**Rustenburg**	Rs	F

MELANIC TOPSOIL					
Topsoil horizon	Subsoil Horizon	Subsoil Horizon	Form	Code	FG
Melanic	Gley	-	**Willowbrook**	Wo	G
	Red Structured	Lithic	**Stanger**	St	F
	Pedocutanic	Gleyic	**Lauriston**	Lr	G
	Pedocutanic	Alluvium	**Potsdam**	Pd	F
	Pedocutanic	Lithic	**Darnall**	Da	F
	Pedocutanic (thick)		**Bonheim**	Bo	F
	Neocutanic (thick)		**Marolong**	Ml	F
	Soft Carbonate	-	**Steendal**	Sn	F
	Hard Carbonate	-	**Immerpan**	Im	F
	Alluvium (thick)		**Inhoek**	Ik	F
	Lithic	-	**Mayo**	My	F
	Hard Rock	-	**Milkwood**	Mw	F

HUMIC TOPSOIL					
Topsoil Horizon	Subsoil Horizon	Subsoil Horizon	Form	Code	FG
Humic	Yellow-Brown Apedal	Gleyic	**Dartmoor**	Dm	A
	Yellow-Brown Apedal	Red Apedal	**Kranskop**	Kp	A
	Yellow-Brown Apedal	Soft Plinthic	**Eland**	El	A
	Yellow-Brown Apedal	Lithic	**Longtom**	Lg	A
	Yellow-Brown Apedal (thick)		**Magwa**	Ma	A
	Red Apedal	Gleyic	**Highmoor**	Hm	A
	Red Apedal	Soft Plinthic	**Netherley**	Ne	A
	Red Apedal	Lithic	**Gangala**	Ga	A
	Red Apedal (thick)		**Inanda**	Ia	A
	Neocutanic	Soft Plinthic	**Umvoti**	Um	A
	Neocutanic	Lithic	**Henley**	He	A
	Neocutanic (thick)		**Sweetwater**	Sr	A
	Lithic	-	**Nomanci**	No	A
	Hard Rock	-	**Graskop**	Gp	A

ORTHIC TOPSOIL					
Topsoil horizon	Subsoil Horizon	Subsoil Horizon	Form	Code	FG
Orthic	Gley	-	**Katspruit**	Ka	I
	Albic	Gley	**Kroonstad**	Kd	I
	Albic	Yellow-brown Apedal	**Constantia**	Ct	B
	Albic	Red Apedal	**Shepstone**	Sp	B
	Albic	Neocutanic	**Vilafontes**	Vf	B
	Albic	Soft Plinthic	**Longlands**	Lo	E
	Albic	Hard Plinthic	**Wasbank**	Wa	E
	Albic	Podzol/Unconsolidated material with wetness	**Lamotte**	Lt	P
	Albic	Podzol/ Lithic	**Houwhoek**	Hh	P
	Albic	Podzol	**Concordia**	Cc	P
	Albic	Prismacutanic	**Estcourt**	Es	J
	Albic	Pedocutanic	**Klapmuts**	Km	J
	Albic	Neocarbonate	**Kinkelbos**	Kk	K
	Albic	Lithic	**Cartref**	Cf	H
	Albic	Hard Rock	**Iswepe**	Is	H
	Albic (thick)		**Fernwood**	Fw	C
	Yellow-Brown Apedal	Gleyic	**Pinedene**	Pn	D
	Yellow-Brown Apedal	Red Apedal	**Griffin**	Gf	B
	Yellow-Brown Apedal	Soft Plinthic	**Avalon**	Av	D
	Yellow-Brown Apedal	Hard Plinthic	**Glencoe**	Gc	D

	Yellow-Brown Apedal	Soft Carbonate	**Molopo**	Mp	K
	Yellow-Brown Apedal	Hard Carbonate	**Askham**	Ak	K
	Yellow-Brown Apedal	Lithic	**Clovelly**	Cv	K
	Yellow-Brown Apedal	Hard Rock	**Carolina**	Ca	B
	Yellow-Brown Apedal (thick)		**Ermelo**	Er	B
	Red Apedal	Gleycutanic	**Bloemdal**	Bd	D
	Red Apedal	Soft Plinthic	**Bainsvlei**	Bv	D
	Red Apedal	Hard Plinthic	**Lichtenburg**	Lc	D
	Red Apedal	Soft Carbonate	**Kimberley**	Ky	K
	Red Apedal	Hard Carbonate	**Plooysburg**	Py	K
	Red Apedal	Dorbank	**Garies**	Gr	L
	Red Apedal	Lithic	**Nkonkoni**	Nk	B
	Red Apedal	Hard Rock	**Vaalbos**	Vb	B
	Red Apedal (thick)		**Hutton**	Hu	B
	Red Structured	Lithic	**Magudu**	Md	B
	Red Structured	Hard Rock	**Nshawu**	Ns	B
	Red Structured (thick)		**Shortlands**	Sd	B
	Soft Plinthite	Gleyic	**Westleigh**	We	E
	Hard Plinthite	-	**Dresden**	Dr	E
	Podzol	Unconsolidated material with wetness	**Witfontein**	Wf	P
	Podzol	Lithic	**Groenkop**	Gk	P
	Podzol (thick)		**Pinegrove**	Pg	P
	Prismacutanic	Gleyic	**Idutywa**	Id	J
	Prismacutanic	Pedocutanic	**Heilbron**	Hb	J
	Prismacutanic	Alluvium	**Utrecht**	Ut	J
	Prismacutanic	Lithic	**Sandile**	Sa	J
	Prismacutanic	Hard Rock	**Cookhouse**	Ck	J
	Prismacutanic (thick)		**Sterkspruit**	Ss	J
	Pedocutanic	Gleyic	**Sepane**	Se	J
	Pedocutanic	Alluvium	**Queenstown**	Qt	J
	Pedocutanic	Lithic	**Swartland**	Sw	J
	Pedocutanic	Hard Rock	**Spioenberg**	Sb	J
	Pedocutanic (thick)		**Valsrivier**	Va	J
	Neocutanic	Gleyic	**Tukulu**	Tu	D
	Neocutanic	Neocarbonate	**Makgoba**	Mb	K
	Neocutanic	Soft Carbonate	**Etosha**	Et	K
	Neocutanic	Hard Carbonate	**Gamoep**	Gm	K
	Neocutanic	Gypsic	**Soutvloer**	Sv	K
	Neocutanic	Dorbank	**Oudtshoorn**	Ou	L
	Neocutanic	Alluvium	**Quaggafontein**	Qf	B
	Neocutanic	Unconsolidated material with wetness	**Tshiombo**	Ts	D
	Neocutanic	Lithic	**Tubatse**	Tb	B
	Neocutanic	Hard Rock	**Bethesda**	Be	B
	Neocutanic (thick)		**Oakleaf**	Oa	B
	Neocarbonate	Soft Carbonate	**Addo**	Ad	K
	Neocarbonate	Hard Carbonate	**Prieska**	Pr	K
	Neocarbonate	Gypsic	**Sendelingsdrif**	Sf	K
	Neocarbonate	Dorbank	**Trawal**	Tr	K
	Neocarbonate	Alluvium	**Motsane**	Mt	K
	Neocarbonate	Unconsolidated material with wetness	**Montagu**	Mu	K
	Neocarbonate	Lithic	**Burgersfort**	Bg	K
	Neocarbonate	Hard Rock	**Hofmeyr**	Hf	K
	Neocarbonate (thick)		**Augrabies**	Ag	K
	Soft Carbonate	Unconsolidated material with wetness	**Kolke**	Ko	K
	Soft Carbonate	Hard Carbonate	**Olienhout**	Oh	K
	Soft Carbonate	Gypsic	**Koiingnaas**	Ks	K
	Soft Carbonate	-	**Brandvlei**	Br	K
	Hard Carbonate	-	**Coega**	Cg	K

	Gypsic	-	**Rooiberg**	Ro	K
	Dorbank	-	**Knersvlakte**	Kn	L
	Alluvium (thick)		**Dundee**	Du	C
	Regic Sand (thick)		**Namib**	Nb	C
	Lithic	-	**Glenrosa**	Gs	H
	Hard Rock	-	**Mispah**	Ms	H