Using remote sensing and geographical information systems to classify local landforms using a pattern recognition approach for improved soil mapping

by

Jonathan Tom Atkinson

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

at

Stellenbosch University
Department of Soil Science, Faculty of AgriSciences

Supervisor: Dr Willem Petrus de Clercq
Co-Supervisor: Dr Andrei Borisivich Rozanov

April 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.

This dissertation includes three (3) original papers published in peer-reviewed journals or books and one (1) paper under review, and (1) paper accepted and in print. The development and writing of the papers (published and unpublished) were my principal responsibility.

A full citation for the resulting publications is provided at the beginning of each chapter and in the Appendices.
April 2022

Summary

Presently, a major focus of digital soil mapping (DSM) in South Africa is unlocking the soil-landscape relationships of legacy soil data by disaggregating the only source of contiguous soil information for South Africa, the National Land Type Survey (LTS) (ARC, 2003). Each land type is best defined as a homogenous mapping unit with a unique combination of terrain type, soil pattern and macroclimate properties (Paterson *et al.*, 2015). One of the prevailing reasons for the LTS longevity and continual temporal-interoperability is that terrain description is expressly related to a suite of catenary soil property descriptions (Milne, 1936). These terrain types are further divided into terrain morphological units (TMUs) representing a sequence of patterns based on a 5-unit landscape model of 1-crest, 2-scarp, 3-midslope, 4-footslope and 5-valley bottom. Importantly, dominant soil distribution patterns are defined by terrain units relying on an elementary terrain topo-sequence pattern approach, with much of the work done on modelling soil variation related to variation in terrain (van Zijl, 2019). Whilst the LTS remains a source of national interest, there is immense opportunity to build on the existing soil inventory data rather than only focus on "breaking it down" (disaggregation). However, what is needed is a standard operating procedure that not only leverages the ability of digital elevation models (DEM) to explicate soil-landscape associations beyond the limited 5-unit landscape model but allows better refinement of soil descriptions with landscape features. Only once the nuances of optimal DEM parametrisation under controlled conditions are fully understood can the complete scope of DSM and digital geomorphological mapping (DGM) applications be explored.

This dissertation attempts to synthesise knowledge on theory, methods, and applications of using remote sensing (RS) and geographical information systems (GIS) to classify local landforms using a pattern recognition approach for improved soil mapping in the context of multiscale problems of digital terrain analysis in KwaZulu-Natal. The dissertation is divided into three parts. Part one (Chapter 2) represents the DEM pre-processing and generalisation method and establishes the protocols for soil-landscape covariate application derived from various sensor platforms and spatial scales. Part two (Chapter 3) introduces the concept of improved terrain unit mapping through the geomorphon approach and describes DEM optimisation for standardised geomorphon representation for uniformly describing soil-landscape properties for inputs to DSM applications. Finally, part three (Chapters 4 & 5) looks at applications of DEM sources and geomorphons first from a holistic landscape context by linking digital terrain and soil-landscape analysis to geodiversity. Finally, the benefit of improved RS and GIS combined with quantitative modelling approaches on improving natural resource predictions are explored by modelling soil-ecotope and soil type mapping units and proposing improvements to an existing DSS designed for KwaZulu-Natal Natal. Specifically, this research is organised into four (4) research chapters with an overview of each chapter's contribution outlined hereafter.

Chapter 2 accounts for the recognition and requirements of DEM generalisation from high to medium resolution RS platforms and the influence these pre-processing approaches have on the extraction of a wide range of terrain attributes. Digital elevation data are elemental in deriving primary topographic attributes that are input variables to various regional soil-landscape models. DEMs' utility to extract different topographic indices as primary inputs to DSM allows the generalised soil-formative relationship between topography and soil characteristics to be measured quantitatively. Traditional landscape-scale approaches to extracting and analysing soils remain subjective and an expensive last resort for large-scale regional soil distribution and variability prediction. Selecting the right DEMs is a critical step in the development of any soil-landscape model. Therefore, the ability to represent soil-landscape relationships rapidly and objectively between soil properties and landscape position using emerging technologies and elevation data in a digital environment and

at varying scales is fundamental for using soil-landscape mapping as a regional planning tool. There is, however, still varied consensus on the effect of DEM source and resolution on the application of these topographic attributes to landscape and geomorphic characterisation within South Africa. However, Atkinson et al. (2017) have shown that topographic variable extraction is highly dependent on the DEM source and generalisation approach. However, while higher resolution DEMs may represent the "true" landscape surface more accurately, they do not necessarily offer the best results for all extracted terrain variables for modelling soil-landscape outputs. Given the convenience of a wide range of open-source elevation data for South Africa, there is a need to quantify the impact that DEM generalisation approaches have on simplifying detailed DEMs and compare the accuracy and reliability of results between high resolution and coarse resolution data on the extraction of localised topographic variables as a primer for soil-landscape or digital soil models.

Chapter 3 explores the harmonisation of geomorphons derived from various RS platforms to define the landscape character in central KwaZulu-Natal. Robust DGM approaches that can simplify and translate the inclusion of "human knowledge" to automatic terrain classification across a broader spectrum of terrain morphological units and a range of DEM spatial scales offer great potential for improved topographic and landscape analysis and must have their utility investigated. Continual advances in quantitative modelling of surface processes, combined with new spatio-temporal and geo-computational algorithms, have revolutionised the auto-classification and mapping of landform components through the automated analysis of high-quality DEMs. Therefore, a thorough assessment of the effects that different pixel resolution (grain size) and DEM sources have on replicating observed geomorphic spatial patterns and representing selected terrain parameters using advanced automated geomorphometric mapping approaches is necessary. Specifically, it would be valuable to interrogate the self-adapting ability of these automated mapping approaches under regional conditions to quantitatively analyse how the choice of terrain model and scale influences the extraction, generalisation, and representation of digitally derived terrain attributes such as slope gradient, elevation and terrain unit feature extent. Equally important is understanding how the variation in resulting terrain unit representation is limited by spatial resolution discontinuities that ultimately influence the extraction and representation of elementary soil properties.

Chapter 4 is a shift from the technical aspects of digital terrain preprocessing and modelling and instead attempts to explore the contribution of gridded soil-landscape products to the abiotic landscape development agenda. It would be worthwhile to contextualise and decode these technical aspects of terrain and soil analyses to a holistic landscape development agenda. It is argued that current global environmental problems and questions demand exploration into new scientific perspectives and improved related paradigms and methodologies. Geodiversity (abiotic complexity) has not received the same level of attention as biodiversity (biotic complexity) despite its intrinsic and indivisible linkages to ecosystem and landscape richness characterisation. The ability to better describe the substrate in which biological and human activities occur is of top standing and must have its potential explored. To date, only one landmark study has successfully investigated the influence of environmental factors on geodiversity mapping in South Africa (Kori *et al.*, 2019). Using an array of multimodal environmental covariates, including hydrographic, lithostratigraphic, pedological, climatic, topographic, solar morphometric and geomorphic variables, I aim to provide further confirmation to regional and international geodiversity research agendas.

Chapter 5 culminates in applying quantitative DSM methods, with improved terrain representation, to classify productive soil units (ecotopes) as a proposed methodology to improve the current Bioresource Report Writer (BRW) soil-landscape recommendations. In KwaZulu-Natal, it has been accepted that detailed natural resource information based on scientifically accurate and relevant criteria is required to develop spatial layers that planners, developers, local government, and other stakeholders can use to guide future development. At present, the KwaZulu-Natal Department of Agriculture and Rural Development (KZNDARD) can provide high-level crop production approximations for various crops based on BioResource Units (BRU). However,

the BRW has not seen a significant revision for over two decades. Still, the natural resource information it contains provides land managers, policymakers and farmers with invaluable access to regional and farm level qualitative estimations of agricultural productivity. There is a need to preserve this information while simultaneously providing modern measures of land management recommendation at multiple scales to the end-user. Against this backdrop, access to readily interpretable soil and crop information is increasingly being prioritised by provincial planning commissions as critical inputs to DSS for sustainable land management within KwaZulu-Natal.

Opsomming

Tans ontsluit 'n groot fokus van digitale grond kartering (DSM) in Suid-Afrika die grond landskap verhoudings van nalatenskap grond data deur die enigste bron van aaneenlopende grond inligting vir Suid-Afrika, die Nasionale Grondtipe-opname (ARC, 2003) te distreun. Elke land tipe word die beste gedefinieer as 'n homogene karterings eenheid met 'n unieke kombinasie van terrein tipe, grondpatroon en makro klimaat eienskappe (Paterson *et al.* , 2015) . Een van die heersende redes vir die LTS-langlewendheid en voortdurende temporale interoperabiliteit is dat terrein beskrywing uitdruklik verband hou met 'n reeks katalise grondeiendom beskrywings (Milne, 1936). Hierdie terrein tipes word verder verdeel in terrein morfologiese eenhede (TMUs) wat 'n reeks patrone verteenwoordig wat gebaseer is op 'n 5-eenheid landskap model van 1-kuif, 2-serp, 3-midslope, 4-voet en 5-vallei bodem. Belangrik, dominante grond verspreidings patrone word gedefinieer deur terrein eenhede wat staatmaak op 'n elementêre terrein topo-volgorde patroon benadering, met baie van die werk gedoen op modellering grond variasie wat verband hou met variasie in terrein (van Zijl, 2019). Terwyl die LTS bly 'n bron van nasionale belang; daar is enorme geleentheid om voort te bou op die bestaande grond voorraad data eerder as om net te fokus op "afbreek" (disaggregasie). Wat egter nodig is, is 'n standaard bedryfsprosedure wat nie net die vermoë van digitale hoogte modelle(DEM) gebruik om grond landskap verenigings buite die beperkte 5-eenheid landskap model te vererger nie, maar beter verfyning van grond beskrywings met landskap kenmerke moontlik te maak. Slegs sodra die nuanses van optimale DEM parametrisasie onder beheerde toestande ten volle verstaan word, kan die volledige omvang van DSM- en digitale geomorfologiese kartering (DGM) aansoeke ondersoek word.

Hierdie verhandeling poog om-kennis oor teorie, metodes en toepassings van ute sintetiseer om afstand waarneming (RS) en geografiese inligtingstelsels (GIS) tesing om plaaslike land vorms te klassifiseer deur 'n patroonherkenning benadering vir verbeterde grond kartering in die konteks van multiskaal probleme van digitale terrein analise te klassifiseer. In KwaZulu-Natal. Die verhandeling word in drie dele verdeel. Deel een (Hoofstuk 2) verteenwoordig die DEM-voor verwerker- en veralgemenings metode en vestig die protokolle vir grondlandskap-kovariaat toediening afgelei van verskeie sensor platforms en ruimtelike skale. Deel twee (Hoofstuk 3) stel die konsep van verbeterde terrein eenheid kartering deur die geomorfon benadering bekend en beskryf DEM-optimalisering vir gestandaardiseerde geomorfon verteenwoordiging om grond landskap eienskappe eenvormig te beskryf vir insette tot DSM-toepassings. Ten slotte, deel drie (Hoofstukke 4 & 5) kyk na toepassings van DEM bronne en geomorfon eerste vanuit 'n holistiese landskap konteks deur die koppeling van digitale terrein en grond landskap analise aan geodiversiteit. Ten slotte word die voordeel van verbeterde RS en GIS gekombineer met kwantitatiewe modellerings benaderings op die verbetering van natuurlike hulpbron voorspellings ondersoek deur grond-ekopeïen- en grondtipe karterings eenhede te modelleer en verbeterings voor te stel aan 'n bestaande DSS wat vir KwaZulu-Natal ontwerp is. Spesifiek, tsy navorsing is organiseer in vier (4) navorsing hoofstukke met 'n oorsig van elke hoofstuk se bydrae wat hierna uiteengesit word.

Hoofstuk 2 is verantwoordelik vir die erkenning en vereistes van DEM veralgemening van hoë tot medium resolusie RS platforms en die invloed wat hierdie preprocessing benaderings het op die onttrekking van 'n wye verskeidenheid van terrein eienskappe. Digitale hoogte data is elementêr in die afleiding van primêre topografiese eienskappe wat inset veranderlikes aan verskeie plaaslike grond landskap modelle is. DEMs se nut om verskillende topografiese indekse as primêre insette tot DSM te onttrek, laat die algemene grond vormende verhouding tussen topografie en grondeienskappe kwantitatief gemeet word. Tradisionele landskap skaal benaderings tot die onttrekking en ontleding van grond bly subjektief en 'n duur laaste uitweg vir grootskaalse streeks grond verspreiding en veranderlikheid voorspelling. Die keuse van die regte DEMs is 'n

kritieke stap in die ontwikkeling van enige grond landskap model. Daarom is die vermoë om grond landskap verhoudings vinnig en objektief tussen grondeienskappe en landskap posisie te verteenwoordig deur opkomende tegnologieë en hoogte data in 'n digitale omgewing te gebruik en op verskillende skale fundamenteel vir die gebruik van grond landskap kartering as 'n streeksbeplanning instrument. Daar is egter steeds uiteenlopende konsensus oor die uitwerking van DEM-bron en resolusie oor die toepassing van hierdie topografiese eienskappe aan landskap- en geomorfiese karakterisering binne Suid-Afrika. Atkinson *et al*. (2017) het egter getoon dat topografiese veranderlike onttrekking baie afhanklik is van die DEM-bron en veralgemenings benadering. Alhoewel hoër resolusie-DEMs die "ware" landskap oppervlak meer akkuraat kan verteenwoordig, bied hulle nie noodwendig die beste resultate vir alle onttrokke terrein veranderlikes vir die modellering van grond landskap-uitsette nie. Gegewe die gerief van 'n wye verskeidenheid oopbron-hoogte data vir Suid-Afrika, is dit 'n behoefte om die impak wat DEM-veralgemenings benaderings het op die vereenvoudiging van gedetailleerde DEMs te kwantifiseer en die akkuraatheid en betroubaarheid van resultate tussen hoë resolusie en growwe resolusie data te vergelyk oor die onttrekking van gelokaliseerde topografiese veranderlikes as 'n primer vir grond landskap of digitale grond modelle.

Hoofstuk 3 ondersoek die harmonisering van geomorfon wat van verskeie RS-platforms afkomstig is om die landskap karakter in Sentraal-KwaZulu-Natal te definieer. Robuuste DGM benaderings wat die insluiting van "menslike kennis" kan vereenvoudig en vertaal na outomatiese terrein klassifikasie oor 'n breër spektrum van terrein morfologiese eenhede en 'n verskeidenheid DEM ruimtelike skale bied groot potensiaal vir verbeterde topografiese en landskap analise en moet hul nut ondersoek. Voortdurende vooruitgang in kwantitatiewe modellering van oppervlak prosesse, gekombineer met nuwe spatio-temporale en geo-berekenings algoritmes, het die ou toklassifikasie en kartering van land vorm komponente omwentel deur die outomatiese analise van hoë gehalte DEMs. Daarom is 'n deeglike assessering van die effekte wat verskillende pixel resolusie (graan grootte) en DEM-bronne het op die replisering van waargenome geomorfiese ruimtelike patrone en verteenwoordig geselekteerde terrein parameters met behulp van gevorderde outomatiese geomorfon metriese karterings benaderings nodig. Spesifiek, dit sal waardevol wees om die self-aanpassing vermoë van hierdie outomatiese kartering benaderings onder streeks toestande te ondervra om kwantitatief te analiseer hoe die keuse van terrein model en skaal die onttrekking, veralgemening en voorstelling van digitaal afgeleide terrein kenmerke soos hellings gradiënt, hoogte- en terrein eenheid-funksie omvang beïnvloed. Ewe belangrik is om te verstaan hoe die variasie in gevolglike terrein eenheid verteenwoordiging beperk word deur ruimtelike resolusie-stakings wat uiteindelik die onttrekking en voorstelling van elementêre grondeienskappe beïnvloed

Hoofstuk 4 is 'n verskuiwing van die tegniese aspekte van digitale terrein voor verwerking en modellering en poog eerder om die bydrae van geroosterde grond landskap produkte na die abiotiese landskap ontwikkelings agenda te verken. Ek sou die moeite werd wees om hierdie tegniese aspekte van terrein- en grond ontledings na 'n holistiese landskap ontwikkelings agenda te kontekstualiseer en te dekodeer. Daar word aangevoer dat huidige globale omgewingsprobleme en vrae eksplorasie in nuwe wetenskaplike perspektiewe en verbeterde verwante paradigmas en metodologieë vereis. Geodiversiteit (abiotiese kompleksiteit) het nie dieselfde vlak van aandag as biodiversiteit (biotiese kompleksiteit) ontvang nie, ten spyte van sy intrinsieke en ondeelbare verbande met ekosisteem- en landskap ryke karakterisering. Die vermoë om die substraat waarin biologiese en menslike aktiwiteite voorkom, beter te beskryf, is van bostaande en moet sy potensiaal ondersoek. Tot op hede het slegs een ander landmerk studie die invloed van omgewingsfaktore op geodiversiteits kartering in Suid-Afrika (Kori *et al.* , 2019). Met behulp van 'n verskeidenheid multimodale omgewings kovariaat, insluitend hidrografiese, lithostratigraphic, pedologiese, klimaat-, topografiese, son morfometriese en geomorfiese veranderlikes, beoog ek om verdere bevestiging te gee aan streeks- en internasionale geodiversiteits navorsing agendas.

Hoofstuk 5 kulmineer in die toepassing van kwantitatiewe DSM-metodes, met verbeterde terrein verteenwoordiging, om produktiewe grondeenhede (ekotipes) te klassifiseer as 'n voorgestelde metodologie om die huidige BRW-grondlandskap aanbevelings te verbeter. In KwaZulu-Natal is aanvaar dat gedetailleerde

natuurlike hulpbron inligting gebaseer op wetenskaplik akkurate en relevante kriteria nodig is om ruimtelike lae te ontwikkel wat beplanners, ontwikkelaars, plaaslike regering en ander belanghebbendes kan gebruik om toekomstige ontwikkeling te lei. Tans kan die KwaZulu-Natal Departement van Landbou en Landelike Ontwikkeling (KZNDARD) hoëvlak-gewasproduksie-benaderings vir verskeie gewasse op grond van BRUs verskaf. Die BRW het egter vir meer as twee dekades nie 'n beduidende hersiening gesien nie. Tog bied die natuurlike hulpbron inligting wat dit bevat, grond bestuurders, beleidmakers en boere van onskatbare waarde toegang tot streeks- en plaasvlak kwalitatiewe beramings van landbou produktiwiteit. Daar is 'n behoefte om hierdie inligting te bewaar, terwyl dit terselfdertyd moderne maatreëls van grondbestuur aanbeveling op verskeie skale aan die eindgebruiker verskaf. Teen hierdie agtergrond word toegang tot geredelik interpreteerbare grond- en gewas inligting toenemend deur provinsiale beplanningskommissie geprioritiseer as kritiese insette tot DSS vir volhoubare grondbestuur binne KwaZulu-Natal.

A generalist knows less and less about more and more until eventually he or she knows nothing about everything. A specialist knows more and more about less and less until eventually he or she knows everything about nothing.

Nicholas Butler

I dedicate this thesis to my wife, Kylie, and my daughters Lilly and Layah. My reasons for being.

To my Parents, May and Pieter. Thank you for making me the man I am today.

To JP and Inga. You never gave up on me. I love you both.

Biographical Sketch

I, Jonathan Atkinson, graduated with a BSc degree in Applied Environmental Science from the University of KwaZulu-Natal in December 2006, with meritorious distinction in Advanced Spatial Modelling. In December 2007, I completed my BSc (Hons) in Applied Environmental Science under the tutelage of Professor Trevor Hill, Professor (Biogeography), Professor Mike Savage (Agrometeorology), Professor Emeritus Roland Schulze (Hydrology) and Professor Onnisimo Mutanga (GIS and Remote Sensing). In September 2012, after a self-imposed five-year academic hiatus, I completed my MSc (Cum Laude) in Environmental Issues and Society with the University of Pretoria's Department of Geography, Geoinformatics and Meteorology.

I currently serve as the Sustainable Land Management Project Coordinator with the International Soil Reference Information Centre (ISRIC-World Soil Information). My primary focus is on facilitating the management, coordination and implementation of the Land, soil and crop Information Services for East Africa under the European Union's Development Smart Innovations through Research in Agriculture Partnership. I also represent ISRIC on the Executive Team for the World Overview of Conservation Approaches and Technologies (WOCAT), focusing on land degradation neutrality (LDN) and sustainable land management practices.

Before joining ISRIC, I was employed as the Director of Operations for the National Land Restitution Evaluation Study with the University of Cape Town's Southern Africa Labour and Development Research Unit (SALDRU). The study, commissioned by the Department of Rural Development and Land Reform and the International Institute for Impact Evaluation, aims to evaluate the impact of compensation (cash & agricultural land) on Land Restitution Beneficiaries as part of the Land Reform Program and Agrarian Transformation Policies of South Africa.

Before joining SALDRU, I occupied the position of Professional Scientist with the KwaZulu-Natal Department of Agriculture and Rural Development's (KZNDARD) Natural Resources Sub-Directorate. My primary role was to characterise and evaluate the province's natural resources and provide strategic specialist support to Ministerial Echelons regarding the feasibility and sustainability of National Development Projects related to Agri-parks, Operation Phakisa Agri-Villages and River Catalytic Projects for rural development.

I am a registered Geospatial Professional with the South African Council of Natural and Professional Scientists (SACNASP) and the South African Geomatics Council (PLATO). I also hold a position on the National Soil Classification Working Group tasked with curating all matters relating to the National Classification of soils in RSA. Since starting my PhD, my research has focused on developing rapid assessment approaches for evaluating the agronomic potential of rural areas using a soil-landscape method for sustainable livelihood purposes. In the broadest sense, my research aims to test novel forms of combining field-based survey approaches with GIS and Remote Sensing technologies to improve the characterisation and evaluation of natural resources and quantify the risk and vulnerability associated with shifting climatic patterns.

Acknowledgements

First and foremost, I would like to give thanks to my Lord and Saviour, Jesus Christ. Your grace is enough for me, for your power is made perfect in my weakness. I thank you for blessing me with a faithful, supportive, and loving soulmate, Kylie, whom you had allowed to be my rock when my tank was empty. You have taken this journey with me every step of the way, and for that, I share this achievement. To my two little girls, Lilly and Layah - my proudest achievement, your love and affection for Daddy are beyond comparison. I love you both so very much, and I thank you for having the understanding and patience to allow Daddy to complete this PhD. Your hugs and cuddles are what got Daddy over the finish line!

Completing this research would not have been possible without my supervisor's support, guidance, and intellectual investment, Professor Willem de Clercq. Who would have thought our paths would be so intertwined the first time I met in 2014 at STIAS? Your wise counsel, both spiritually and academically, and friendship has prepared me to be a better scientist and a better husband and father. Our many "think-tank" sessions and discussions pushing the boundary on new methods will forever be remembered. To my co-supervisor, Dr Andrei Rozanov, I recall the first time I heard you deliver a talk at STIAS in 2014. I could not understand a word you were saying. That is because ***I*** was not listening! Your wealth of soil, geospatial science and geomorphology knowledge and strategic insights to novel research have allowed me to become a better" listener". I consider myself extremely fortunate to have shared this journey with you. Your humour and willingness to provide the necessary support and guidance over a cup of coffee have been a highlight of our relationship. I am reminded of our field trip to the Drakensberg and your determination to return to Cape Town with our 20 kg fossilised mementoes – a highlight of my PhD.

- I wish to further express my sincere gratitude and appreciation (in no particular order) to the following persons and institutions:
- To Sappi Southern Africa, particularly Dr Jacob Crous, for your support, expert input and data contribution regarding access to the Forestry Soil Database repository.
- Dr Riyad Ismail of Sappi Southern Africa supervised my MSc, exposed me to the practical utility of spatial and statistical data analysis, and taught me the benefits of being a thorough researcher. *Allah yirħama.*
- To Kurt Barichievy for introducing me to coffee, the middle-mouse button and for knowing when to sneak me away for a spot of fishing. Your effortless support, wisdom and encouragements during my PhD have been a central component to my completion. I have come far together, and I value your friendship above all.
- Professor Adriaan van Niekerk, for your guidance, support and constructive criticism when I needed it.
- To Theo Pauw and the Team at the Centre for Geographical Analysis (CGA), Stellenbosch University, for providing me with high-quality Stellenbosch University DEM's.
- To Professor Martin Kidd, from the Centre for Statistical Consultation, your guidance and statistical contributions were much appreciated.
- To Professor Jeff Hughes for inspiring my passion for soil science.
- To Professor Trevor Hill, you unequivocally are the reason why I pursued my PhD. It has taken me a while, but I hope that I have made you proud!
- To Dr Gary Paterson of the ARC-ISCW and his support staff for their continued support and willingness to engage on soil-related matters and providing access to the wealth of Land Type Survey data.

- The National Research Foundation (NRF). The well-timed financial support provided towards this research is hereby acknowledged. I must note that the opinions expressed, and conclusions arrived at are those of the author and are not necessarily attributed to the NRF.
- Lastly, I would like to thank the KwaZulu-Natal Department of Agriculture and Rural Development for the partial bursary funding I received from them. Thank you for providing the necessary financial support to pursue the PhD, travel when required and purchase the necessary research equipment.

Table of Contents

List of Figures

List of Tables

Abbreviations and acronyms

AGL	Above ground level
AEZ	Agro ecological zones
ANOVA	Analysis of Variance
ARC	Agricultural Research Council
ASA	Average size area
ASTER	Advanced spaceborne thermal emission and reflection radiometer
BPS	Biome Protection Status
BRG	Bioresource Group
BRU	Bioresource Units
BRW	Bioresource Report Writer Version 9
CGA	Centre for Geographical Analysis
CGIAR	Consultative Group on International Agricultural Research Consortium
DEM	Digital elevation model
DGM	Digital geomorphic mapping
DSA	Digital soil assessment
DTM	Digital terrain model
DSM:	Digital soil mapping
DsM	Digital surface model
EarthENV	EarthEnv-DEM90
ERD	Effective rooting depth
ESRI	Environmental Systems Research Institute
FAO	Food and Agricultural Organisation of the United Nations
FSD	Forest Industry Soils Database
GDEM	Global DEM
GDIx	Geodiversity Index
GDSS	Geomorphic decision support systems
GIS	Geographical information systems
GIT	Geographic information technology
GLSDEM	Global Land Survey Digital Elevation Model
GLCF	Global Land Cover Facility
GNSS	Global network of satellite systems
GSHC	Gentle slope, high convexity
GSLC	Gentle slope, low convexity
GPS	Global Position systems

HCD	Hydrological corrected topo-to- raster
HD	Horizontal displacement
HSD	Honestly significant difference
ICC	Intra-correlation coefficient
IFSAR	Interferometric Synthetic Aperture Radar
IDW	Inverse distance weighted
KZNDARD	KwaZulu-Natal Department of Agriculture and Rural Development
LD	LiDAR-derived
LiDAR	Light detection and ranging
LOS	Line-of-sight
LS	LS factor
LTS	Land Type survey
MA	Mean cell aggregation
MAD	Mean Absolute Deviation
MAUP	Modifiable area unit problem
MCC	Matthews Correlation Coefficient
METI	Ministry of Economy, Trade, and Industry of Japan
MLA	Maximum location accuracy
NASA	National Aeronautics and Space Administration
NCS	National Biodiversity Conservation Status
NED	National Elevation Dataset
NEMA	National Environmental Management Act of 2008
NHRA	National Heritage Resource Act of 1999
NLCap	National Land Capability
NLCov	National Land Cover
NLD	Non-LiDAR-derived
NOAA	National Oceanic and Atmospheric Administration
NNR	Nearest neighbour resampling
NSEC	Nash-Sutcliffe efficiency coefficient
NTB	National Terrestrial Biomes
OOB	Out of bag error rate
PDALB	Preservation and Development of Agricultural Land Bill of 2020
PlanCurv	Planform curvature
ProCurv	Profile curvature
PSM	Predictive soil mapping
PTF	Pedotransfer functions
QGIS	Quantum GIS
RF	Random Forest classification

RMSE	Root mean square error
RS	Remote sensing
SABC	South African Bionomical Classification System
SAHRA	South African Heritage Resources Agency
SAGA	System for geoscientific analysis
SD	Standard deviation
SEM	Standard Error of Mean
SlopAsp	Slope aspect
SlopGrad	Slope gradient
SOTER	SOil and TERrain digital database
SRTM	Shuttle radar topography mission
SSHC	Steep slope, high convexity
SSLC	Steep slope, low convexity
SUDEM	Stellenbosch university digital elevation model
SPD	Soil profile database
SPLUMA	Spatial Planning and Land Use Management Act of 2013
TBS	Tugela Basin Soils
TMU	Terrain morphological unit
TSC	Terrain surface classification
TST	Terrain surface texture
TRI	Terrain ruggedness index
TPI	Topographic position index
TWI	Topographic wetness index
USDA	United States Department of Agriculture
USGS	United States Geological Survey
UTDM	uThukela District Municipality
WGS 84	World Geodetic system 1984
WRB	World Reference Base

CHAPTER 1

1.1 INTRODUCTION

There is a renewed focus, both locally and internationally, to further understand the environmental and geomorphic drivers of soil-landscape variability and complexity (McKenzie *et al.*, 2000). The need to better appreciate and manage soil resources is the driving force for developing and enhancing geospatial technologies that can rapidly describe and segment the landscape into refined homogenous soil-landscape units for digital soil mapping (Mulder *et al.*, 2011). Given the now global effort towards sustainable development and resource management, it is unsurprising that the need for information on precise quantitative relationships between soils and critical environmental factors is essential to understanding our soil systems' resilience to future threats and how we can better manage the Earth's limited resources sustainably. Frankly, the demand for soil data in response to cross-cutting local, national, and global policy issues of accelerated soil loss, land use degradation, biodiversity and climate change is profoundly recognised as the "rule" rather than the "exception." To maintain overall biodiversity, including ecosystem protection, Grunwald *et al.* (2015) stress that much of the focus on global challenges of this generation, such as food and soil security, will not be met without up-to-date, high-quality, high resolution, spatio-temporal, and continuous soil and environmental data. This is needed to characterise the physio-chemical, biological, and hydrologic conditions of ecosystems at varying spatial decision-support scales. Notably, managing these resources requires knowledge and information about them (de Carvalho Junior *et al.*, 2014). Such opportunities are welcomed, particularly in South Africa, where there is often a lack of quality open-source terrain data and novel software development for performing detailed landscape analyses, and much of the digital geomorphometric desiderata remains elusive to the local soil science fraternity.

In South Africa, existing regional and national agency-operated soil databases are neither precise nor exhaustive to support the requisite soil information necessary for digital soil mapping (DSM) spatial data infrastructure. Traditional soil surveying in South Africa was primarily focused on developing resource inventories, usually presented as paper-based choropleth soil class maps. These map resources were instrumental for characterising land for agronomic development needs, driven by the need to prioritise the country's more productive zones, specifically the soil resource suitability for cultivation, at nominal scales between 1:50 000 to 1:250 000 (ARC, 2003). Much of the legacy soil data available is now obsolete since current soil data needs are widely more diverse, with the temporal influence of land use and anthropogenic changes in soil attributes through biophysical process likewise rendering many legacy data sites inappropriate for DSM (Kidd *et al.*, 2020). Moreover, the current "stock" of digitised legacy soil data explicitly incorporates tacit pedological knowledge (Grunwald, 2010) with purposive sampling without sufficient spatial location that reproducibility is nearly impossible. This is further compounded by the current lack of standardisation of descriptive and analytical soil property data collection between private and state-owned entities. However, what is needed is a standard operating procedure that not only leverages the ability of DEMs to explicate soil-landscape associations beyond the limited 5-unit landscape model but allows better refinement of soil descriptions with landscape features. In truth, the substantive contribution of geographical information systems (GIS) and spatial statistics has provided a platform for users to access spatial soil information readily, examine the spatial context of soils, and apply process models to spatially distributed data that offer a range of possibilities for landscape-scale pedology (Pennock & Veldkamp, 2006).

Many regional soil prediction analyses, ecological modelling and management applications still rely on producing soil maps by combining conventional field investigation with the manually delineated Land Type Survey (LTS) (ARC, 2003) to describe the distribution of soils across the landscape. As the need for improved land use planning and decision-making increases in South Africa, driven mainly by legislative and socio-economic developments, the importance of soil information is being strategically prioritised to effectively manage high-value agricultural land, albeit from a top-down National approach. In these instances, it will be necessary to develop, test and refine soil-landscape approaches that can readily relate soil properties with accessible remotely sensed digital elevation model (DEM) data to reach a broader spatial planning directive. The LTS is only applicable for use on a countrywide basis and thus is not designed to provide the detailed (high attribute and spatial resolution) soil information required for regional or landscape specific models. Furthermore, According to Van den Bergh *et al.* (2009), the LTS contains unacceptably large areas with either over or underestimated terrain unit proportions for a particular land unit inconsistent with surrounding land types. Naturally, this presents a suite of challenges to the National Development Agenda, of which natural resource management and land use planning feature heavily (KZNPPC, 2013). Information about the natural resource distribution of a country is crucial for various purposes, from local and regional planning to economic forecasting to food security and environmental preservation (Paterson *et al.*, 2015). Terrain, climate, and soil, three of the most critical resources for agriculture, directly impact production and the prospect of sustainable livelihoods. Several studies have shown that recognition of geomorphic landscape features and derivation of associated terrain and soil covariates is strongly dependent on the spatial resolution of the input terrain data and the adequate interpretation of topography and classification of terrain into landform elements (Grunwald, 2010; McBratney *et al.*, 2011; Wilson, 2012; Florinsky, 2016).

DEMs provide a common approach to digital representations of topographic variables helpful in describing the bare Earth surface (DTM) or a digital surface model that includes vegetation and built structures (DsM). With DEMs, users can analyse and model the land surface and relationships between the landscape's topography and geological, hydrological, biological and anthropogenic components (Florinsky *et al.*, 2002). Typically, DSM techniques leverage topography's formative influence on soil formation and distribution to spatially analyse the precise quantitative relationships between soil attributes and landscape covariates. Therefore, the success of environmental correlation depends on the strength of relationships between soil and environmental variables. Accordingly, terrain analysis will be most useful in environments where the topographic shape is strongly related to soil formation processes, and the environmental variables are readily and remotely measurable (McKenzie *et al.*, 2000). While considered separate disciplines, from an applied perspective, pedology and geomorphology are inexplicably linked (geopedology) to form a single indivisible system – the soil-landscape system (Grunwald, 2016).

Geopedology aims to develop an explicit analogue between geomorphology and pedology to analyse soil-landscape relationships and map soils as they occur on the landscape (Zinck, 2013). From this point of view, landscape models have shown that the landscape elements are predictable, and the geomorphic component primarily controls a large part of the non-random spatial variability of the soil cover. According to Zinck (2013), the geopedologic paradigm considers geomorphology as the “container” for terrain covariate map units, while pedology provides their taxonomic components, i.e. the “content”. Incidentally, geopedologic map units are considered more comprehensive than conventional soil maps since they contain information about the geomorphic context in which soils are found and developed. Thus, soil-landform classifications based on remotely acquired geomorphometric parameters (digital terrain data) derived from DEMs are important inputs for the digital exploration of variation in landscapes (Mohamed, 2020). Specifically, parallel advances in digital geomorphological mapping (DGM) have modulated further contemporaneous exploration in geomorphological and pedological research.

Vogel *et al.* (2013) further acknowledge that improved insight into topographic structure formation processes would enable better model parameterisation from structural features that can be independently measured or inferred from typical structural attributes of functional soil types.

In the last several decades, advances in remote sensing (RS) and geographic information technology (GIT) have catalysed the evolution of digital mapping systems such as digital soil mapping (DSM), predictive soil mapping (PSM) and digital soil assessment (DSA). These formalised approaches are widely considered the benchmark for generating numerical or statistical digital soil resource data by predicting soils' spatial distribution based on theoretical soil-landscape models that characterise soils as a function of ancillary environmental variables, including climate, organisms, topography, parent material, time and spatial location, recognised as the primary factors governing soil formation (Jenny, 1941; Grunwald, 2010; McBratney *et al.*, 2011). Recently, Arrouays *et al.* (2020) have reviewed the significant advantages of DSM methods over conventional soil mapping approaches. These benefits, collectively with the development of improved geospatial computing power, new predictive techniques, including machine learning and greater availability of covariates and environmental data, support the realisation of DSM as a pragmatic, cost-efficient operational approach resulting from the "perfect storm" of ideas and technologies over time (Kidd *et al.*, 2020). The DSM approach avoids many limitations of conventional soil mapping methods. However, these methods are not without their criticisms, limitations and concerns, which are constructively detailed by Minasny & McBratney (2016), Kidd *et al.* (2020) and Arrouays *et al.* (2020). The most notable of these concerns in the Southern Africa context is the nexus of sparse, incomplete, and inconsistent contiguous soil data available to the end-user for DSM applications at multiple scales. Deployable DSM requires a repository of unfragmented spatially adequate and calibrated soil data. Unfortunately, to fill this data vacuum, collecting "raw" qualitative and analytical field data for training, prediction and validation in optimal locations and density requires considerable resources.

Developments in RS have led to the availability of DEMs of high resolution (1 m cell^{-1}) for an increasing number of local areas and medium resolution (10-1000 m cell^{-1}) for most of the terrestrial landmass (Jasiewicz & Stepinski, 2013). Over this period, land-surface analysis and classification have seen rapid improvements in the rate and quality of geomorphometric computational approaches, a key factor for the utility of digital geomorphic mapping. The advent of digital soil and geomorphological mapping and a transition from qualitative to more quantitative soil mapping methods have increased the utility of computationally efficient techniques for auto-classification and mapping of landform components from a DEM. Specifically, DGM focuses on the extraction of measures (land-surface parameters) and spatial features (land-surface objects) from digital topography (Wilson, 2012). With modern DGM approaches, users can now quantify better and represent landscape morphology, evaluate surface-biophysical associations, and implicitly explore spatial landscape complexity through a broad suite of diversity and contagion indices of landscape diversity. In this regard, DGM presents an *avant-gard*e solution for modern geomorphological analysis using contemporary geospatial and geo-computational techniques. Here, DGM is understood to include three-dimensionally distributed geo-referenced databases, the capabilities of dynamic visualisation and virtual reality platforms, geomorphometry and digital terrain modelling, landscape evolution models, information-extraction technologies, and an assortment of modern parallel subfields (James *et al.*, 2012). Specifically, advances in RS of land surface terrain features and geospatial modelling systems that support numerical modelling of surface biophysical processes have reformed traditional geomorphic analyses and mapping. The focus now is on integrating quantitative topographic information, on a cell-by-cell basis, with some form of geospatial platform with an ever-increasing reliance on geostatistical approaches and artificial intelligence to describe the dynamic landscape and topographic change, patterns, and complexity. These digital systems, data and tools, often referred to as geomorphic decision support systems (GDSS), are increasingly favoured for their contributions to diagnostic assessments and modelling of the terrain to achieve enhanced interpretations of scale, patterns, processes and diversity of landscape features and systems.

In South Africa, where traditional soil surveys' costs and effort remain high, the importance of RS's role in DGM and soil-landscape analysis cannot be overstated. As expected, the reliance on high-resolution (temporal and geometric) imagery has proliferated. It now dominates much of the automated landform classification research and operational domain from national to local scales. Indeed, RS technology provides new and systematic capabilities for users to analyse continuous land-surface morphometric conditions and efficiently map large areas within a wide range of spatial scales (i.e., Spatio-temporal scales at which surface processes operate) with credible results that are both objective and repeatable to soil-landscape applications. RS further provides a segway to endless opportunities to merge knowledge from different domains (e.g., software, mathematics, engineering, geomorphology) to study landscapes in a more interdisciplinary manner (Napieralski *et al.*, 2013). Notably, in the present study, DEM classifications based on landscape position include landforms on slopes, ridges, spurs or valleys. In contrast, classifications based on the complete environment include landforms in fluvial, coastal, aeolian, glacial, and even volcanic landscapes explicitly excluded from this research agenda (Napieralski *et al.*, 2013).

In a digital era where the rate of new products and techniques far exceeds the commensurate rate of knowledge assimilation and upskilling, users are called upon to have a wide range of formalised understanding of the practical and empirical underpinnings of such technologies and multi-system approaches. However, a problem arises in that as we consider soil-landscape relationships, we immediately are faced with the scales at which we observe soils and the landscapes in which they occur. Importantly, when modelling soil-landscape assemblages, it is essential to acknowledge the gradient between phenomena scales and select the most appropriate predictor covariates for the scale of interest since phenomena governing soil formation and distribution function at different scales (Schoorl *et al.*, 2000). Further, landscape patterns observed at one analysis scale are often not readily observed at other analysis scales. So predictive relationships developed at one scale using a specific algorithm might not be useful in predicting different scales. These phenomena are typically known as the scale effect of the modifiable area unit problem (MAUP) (Amrhein, 1995; Jelinski & Wu, 1996). Despite any resistance to the contrary, DEM algorithms must be applied over various resolutions to determine the scale on which the desired landscape elements can be retrieved (Mulder *et al.*, 2011). Therefore, higher levels of terrain generalisation can, in some cases, provide more explanation of a spatial variable than “bare earth” DEMs that are often considered *ipso facto* optimal for many soil-landscape applications.

Currently, the number of freely accessible DEMs with extensive global coverage is increasing, and the need for standardised DGM desiderata on data structures, tools, analytical protocols, visualisation symbology, and reporting errors is snowballing as the free exchange of data and analytical systems proliferate (James *et al.*, 2012). This is particularly evident in a South African context where approaches such as DGM and DSM to soil-landscape analysis are still in their infancy, with significant research potential to investigate and further calibrate model predictions to local geographical settings.

1.2 BACKGROUND

1.2.1 DEM quality and understanding DEM error

Global elevation datasets are inevitably subjected to errors, mainly due to the methodology followed to extract elevation information and the various processing steps the models have undergone. It is widely accepted that the primary focus of DEM pre-processing of DEMs is to achieve three main objectives: to remove gross errors and artefacts, to make a better approximation of the land surface, and to make a better approximation of the

hydrological/ecological processes (e.g. flow, radiation, deposition) (Reuter *et al.*, 2009). The most common of these processing steps include measurement, digitisation, typing, interpretation, classification, generalisation and interpolation errors (Nikolakopoulos *et al.*, 2006; Heuvelink, 2014). Furthermore, these errors' nature and extent are often unknown and not readily available to spatial data users. So, the classic problem of DEM data quality, integrity and uncertainty remain. Creating DEMs from new and traditional sources is a complex process. It is important to recognise the empirical nature of many spatial analysis and modelling forms, affecting various assumptions and approaches (Goodchild, 2011; Bishop *et al.*, 2012).

Our lack of knowledge about these errors constitutes model uncertainty. Importantly, understanding the limitations of these errors on model outputs will determine if a given map fits the intended use. By the same token, any DEM uncertainty and data error will circulate into its derived topographic parameters, which will lead to varied predictions since model coefficients, model variables, and or model structure may change (Wechsler & Sciences, 2007). So it is equally necessary to quantify the errors and uncertainties associated with a map and communicate these to map users (Heuvelink, 2014). This is an essential yet overlooked research topic in spatial data quality assessment, especially for DGM and DSM applications' integrity to better transfer soil mapping approaches to other landscapes. Dobos (2006) emphasise that one of the most limiting factors of using DEM data is the inconsistency in data collection using different standards and methods.

Many critical considerations persist but are often overlooked as nuanced semantics before a map can even be derived. There are still many questions, and scientists and practitioners must know about the advantages and disadvantages of various representations, metrics, indices, and spatial modelling approaches and their applicability for scientific investigations. Wilson (2018) outlines a series of mainstream questions that at the highest or most general level include the following:

1. What is the proper way of representing the land surface?
2. Which scale is preferred and why?
3. Which elevation sources would be most beneficial for the opportunity or problem at hand?
4. How should a DEM be pre-processed to produce a useful product?
5. How will the DEM error propagate, and what should be done with this uncertainty throughout the subsequent analysis?
6. How should specific land surface parameters be calculated?
7. How should specific land surface objects be delineated?
8. Is it necessary to develop new land surface parameters and objects to address particular problems?
9. Which metrics or indices are best suited to a particular mapping application? Do they even exist?
10. Are there adequate models, or do we have to create them or modify them?

1.2.2 Digital geomorphological modelling: the geomorphon approach to terrain unit mapping

Despite significant progress in specific and general DGM applications globally, Minar & Evans (2008) point out that very few investigations have attempted to formalise landform taxonomy and ontology. In South Africa, no *consensus omnium* exists to promote a systematic digital approach to accurately mapping landforms and geomorphic systems with a different climate, geological and topographic setting. Central to this outcome is the continual refinement of inordinate applications from state-of-the-art active laser and radar RS platforms that provide data of immense interest to DSM, and some not. Mass-produced globally synoptic Interferometric Synthetic Aperture Radar (IFSAR) can be flown at high altitudes and fast speeds, covering large areas with few weather problems and offering a range of open-source DEM resolution products to the end-user (Evans,

2012). In contrast, access to airborne light detection and ranging (LiDAR) technology has increased so rapidly in the last decade that coverage of laser-based DEMs at 1 m resolution is the operational "gold standard" for many sector-specific applications. This is attributed mainly to LiDAR's capability of providing sub-meter DEMs with absolute vertical accuracy better than 0.3 m and, depending on the laser point cloud density. It is now possible to represent terrain gridding at 1 m generalisations (Evans, 2012). Increased accessibility to both LiDAR and conventional technologies provides rapidly increasing capabilities to acquire very exact, high spatial resolution, DEM data quickly and cost-effectively (MacMillan *et al.*, 2003). However, the use of high-resolution terrain models derived from LiDAR technology related to soil-landscape morphometrics and land type modelling has not been fully explored under local conditions due to the data often being prohibitively expensive and limited to small areas (Mashimbye, 2013).

Furthermore, many countries have developed their own national DEM datasets that provide high-quality regional digital surface products, as is the case in South Africa. The Stellenbosch University DEM (SUDEM)(Van Niekerk, 2016), as well as the 2 m digital elevation model of South Africa (DEMSA), is gaining favour from a wide range of users for its accurate representation of the South African landscape since it incorporates photogrammetric contours (ranging from 5 m to 20 m), 1:50 000 spot height data, 1-arc second SRTM DEM, and photogrammetric-derived DSMs and DTMs. With the range in sensor platforms and various accessibility options, in addition to the highly prioritised selection of DEM spatial resolution, users are increasingly challenged to align their usage expectations within DEM radiometric and geometric limitations. This presents both challenges and opportunities for the application of DGM within soil-landscape applications. Digital geomorphology is a rapidly evolving and complex field with a growing number of parameters and algorithms that can be used to track the features that describe the terrain (objects) and geographic boundaries (parameters)(Wilson, 2018). There is thus a need to investigate the potential of developing region-specific computationally efficient techniques of auto-classification and mapping of landform elements derived at various spatial resolutions.

With the over-reliance on "black-box" modelling, front-end applications, and packaged mosaiced image datasets, the diverse audience and wide variety of users are often unaware of the necessary pre-processing required to "prime" DEMs for quality model outputs – which can have severe consequences for decision making (Heuvelink, 2014). Since DEMs only provide a model of reality that inexplicably contains nonconformities of the true surface or errors, the recognition of geomorphic elements depends on the quality of input data used, which is influenced by DEM spatial resolution and accuracy (Mulder *et al.*, 2011). DEM variability and application differ in spatial coverage, data resolution, quality, and noise treatment. Therefore, it is essential to examine their qualities and understand possible errors and other characteristics before extracting different information from these DEM products (Li *et al.*, 2017). Furthermore, DEM data preparation for many DSM and DGM analyses can be complex since the elevation is typically not the dominant attribute of interest. The factual geomorphological accuracy of DEMs can often only be truly assessed by measuring derived surface parameters and objects as drainage lines, landforms or viewsheds in the field and then comparing their shapes, distributions, and location with the values obtained by geomorphometric analysis (Wilson, 2012). Therefore, the quality of the DEM surface inextricably determines the quality of the geomorphic analysis.

From a parallel perspective, the tensions between soil taxonomic units and landscape-based soil mapping units with a perceivable global shift from soil-pedon to more soil-use research have inevitably seen a gradual decline in taxonomically focused pedological research (Grunwald, 2016). However, there is little doubt that landscape processes should not be studied in isolation from their spatial and temporal context (Pennock & Veldkamp, 2006). A significant driver for this paradigm shift is the realisation that field results from classical, small plots of homogenous soils often provide a narrow perspective when applied to non-level, heterogeneous landscapes. In fact, as Gruber *et al.* (2017) highlight, many soil field description guidelines for soil classification schemes

require the characterisation of landform and topography of soil profile sites. For instance, in Food and Agriculture Organization (FAO) guidelines, soil descriptions include details regarding major landforms, the site's relative position within the landscape, slope form, and slope angle (FAO, 2006). Likewise, the United States Department of Agriculture (USDA) *Fieldbook for Describing and Sampling Soils* (Schoeneberger *et al.*, 2012) outlines the procedures for collecting a deluge of geomorphic information, including physiographic and surface morphometry descriptions, including drainage patterns.

Similarly, in South Africa, there are no less than several acceptable guidelines (excluding sector-specific protocols, i.e., mining, forestry, wetland conservation) for soil and land evaluation for agronomic purposes (ARC, 1991). Indeed, the diversity of soils and landforms has marked quantitative and qualitative effects on the landscape. This is because the development and variation of soil properties are strongly influenced by how water and soil materials interact with the land surface and often co-evolve with the local topography (Jenny, 1941; McBratney *et al.*, 2003). Given the overarching influence of relief as a pedogenetic state factor on soil formation, this principle alone is sufficient to explore soil-landscape descriptions further. This is usually done by extracting a suite of common topographic attributes (i.e., slope, aspect, drainage area and network, curvature, topographic wetness index) generated from various DEM sources, using advanced GIS and RS applications for terrain parametrisation. In fact, the global soil and terrain database, initiated by the International Society of Soil Science, commonly known as the SOil and TERrain digital database (SOTER) (Oldeman & Van Engelen, 1993), is explicitly focused on mapping the global landscape using polygons to presenting homogeneous physiographic and lithologic units with soil information assigned to the polygons as soil associations developed at a scale of 1:1 million (Van Engelen & Dijkshoorn, 2013). A potential solution for improving soil-landscape inferential modelling in a South African context may require a multi-modal approach: first include a deep-dive into DEM pre-processing applications that aim to prepare DEM surfaces for DGM and DSM implementation and second, appraise the utility of a dynamic semi-automated framework, using remote sensing and GIS technologies, which identify entire topographic patterns that correspond to specific landforms. In this context, morphometric and geospatial analysis using remote sensing techniques offer a possible solution for further investigating the correlation between soil characteristics and landforms. More precisely, the utility of characterising the landscape into *Local Ternary Patterns* to identify homogenous local landform elements known as geomorphons (Julesz, 1981; Jasiewicz & Stepinski, 2013) must be investigated.

Geomorphons can improve the representation of terrain units and their properties related to the landscape pattern-process by exploring the relationship that geomorphons exhibit with selected soil-landscape indices of land structure (pattern) and landscape function (process). Landscape structure includes size, shape, composition, number and position of terrain-unit patches within a particular landscape (Zaragozí *et al.*, 2012). I believe the geomorphons approach offers great potential to improve terrain classification, providing an operational alternative to the current LTS landform classification at a local to semi-regional scale in South Africa. Furthermore, the application of geomorphons to DSM, specifically soil-landscape modelling, is yet to be fully explored under Southern Hemisphere conditions where unique soil-landscape patterns inevitably influence a diverse range of ecological processes. While many studies have already shown the approach's usefulness under international conditions, limited local studies have sought to capitalise on the approach's computing and classification efficiency. Data interoperability and integration methods must be evaluated to gain further insight into the utility of the geomorphon approach to landscape analysis in South Africa.

Furthermore, the contribution of geomorphons in the assessment of somewhat renewed concepts such as geodiversity in the broader ambit of sustainable development has yet to be quantified appropriately and developed in methodological terms. The point of departure for including geodiversity in this research is based on the understanding that geodiversity is increasingly recognised to underpin and deliver essential ecosystem services (landscape functioning) linking people, landscapes and their heritage to the broader benefit of the

natural and built environment as outlined in the Millennium Ecosystem Assessment (MEA, 2015). Comer *et al.* (2015) showcased that land units with high geodiversity are relatively resistant to degrading living conditions and well equipped to cope with disturbances such as climatic changes than low-geodiversity land units. Therefore, most important to decision-makers is prioritising the endogenous linkages and mapping the spatial variability between geodiversity and geological, hydrogeological, topographic, climatic, geomorphological and pedological covariates. This will facilitate a more holistic approach to mitigate cross-cutting sustainable land management issues of current concern, including habit loss and food insecurity in the face of increased natural and anthropogenic pressures linked to economic development. Thus, the present work aims to provide a first approximation methodology for geodiversity assessment adapted to the regional context with possible scaling-up potential to a national level. Here, a proper application of the concept by producing a geodiversity index (*GDIx*) map highlighting the richness of selective (limited) abiotic elements at a landscape scale is presented. The *GDIx* digital dataset would then be a flexible decision-support tool, even for the most agnostic users, allowing straightforward interpretation regardless of specialist background. To my knowledge, there has been no analysis of diverse areas in KwaZulu-Natal that has quantified the linkages between geodiversity and land-use systems spatially using a framework for quantifying geodiversity.

Perhaps the most critical and most challenging aspect of this study is an attempt to identify a defined set of soil-landscape relationships that can derive a rule-based protocol for producer and user-driven soil prediction using terrain features. Only once the nuances of parameterisation under these controlled conditions are fully understood can the full scope of interoperable applications be explored in broader, ungauged regions. However, this study will undoubtedly lay a foundation for future research in various landscape and ecological applications in South Africa simply because understanding the spatial variations of terrain properties and their relationship with landform elements is key to defining appropriate land management and assessment. Achieving this task could expose the broader academic, professional and citizen science community to DGMs benefits in particular regions like KwaZulu-Natal. Furthermore, while it is not disputed that geomorphons may indeed have a strong relationship with selected terrain features, it is postulated that these relationships will vary with varying pixel resolution of the terrain variables, as studies by Atkinson *et al.* (2017) have shown. However, it cannot be denied that establishing criteria with fixed parameters for specific landforms combined with high-resolution DEMs and geospatial platforms present a rejuvenated frontier for morphometric analysis in South Africa.

1.2.3 The importance of soil-landscapes and Decision Support Systems to the sustainable development agenda

Soil and its associated condition (health), as well as physical, chemical and biological properties, are widely accepted to be inextricably complementary to the Sustainable Development Goals (SDG) (UN, 2020) of the Agenda 2030 of the United Nations (Lal *et al.*, 2021). Of the 17 SDGs, which include a further 169 targets, soils feature extensively given their linkages between the atmosphere, the geosphere, the hydrosphere and the biosphere (Blum, 2016). These linkages directly influence the roles played by soils in addressing the resource, environmental, health and social problems that humanity is currently facing. While soils are not explicitly prioritised in the SDGs, their connection to sustainable management and use of natural resources features prominently intending to realise many of the SDGs. Most notably, SDG 1 or "No Poverty", SDG 2 or "Zero Hunger", SDG 3 or "Good Health and Wellbeing", SDG 11 "Sustainable Cities and Communities, SDG 12 or "Responsible Consumption and Production", SDG 13 or "Climate Action", and SDG 15 or "Life on Land".

Of particular relevance to the present body of work presented herein is SDG 2 since it targets food security-related initiatives to end hunger, achieve food security and improve nutrition. In fact, the nexus between soils

and food security (zero hunger) is relatively straightforward, with a clear association to food production and sustainable agricultural management (Barrett, 2010). Of course, this depends on the availability of natural resources, such as productive land surfaces, topography, soil distribution, soil quality, availability of water, biota, and favourable climatic conditions (Mueller *et al.*, 2010). These globally defined conditions to sustainable food security must be prioritised within the physio-geographic conditions at a local level (Mueller *et al.*, 2018).

Bouma (2020) points out that traditionally, soil surveys and land evaluations are based on expert opinion evaluating the relative suitability or limitations of soils for a range of land uses without providing specific advice about how to resolve limitations. However, the impact of expert knowledge has been extended by recent interventions such as Decision Support Systems (DSS) that systematically collect resource management information to address modern land use questions. Notably, the true benefit of DSS is the ability to adapt and integrate new quantitative approaches into resource management. Therefore, existing databases designed from significant intellectual investments of expert knowledge yet provide a flexible digital platform for the inclusion of quantitative methods and apply to a wide range of users must not only be preserved but prioritised to address contemporary and competing land use challenges on agricultural production. An excellent example of such a DSS is the Bioresource Writer Program (BRW)(Camp, 1995), developed exclusively for the Province of KwaZulu-Natal. The diversity of the natural resources of KwaZulu-Natal is extensive. This significant variation in the natural resources, in turn, leads to variations in the forms of land use and results in a leading challenge in providing for sound ecologically-based development planning in both the conservation, agricultural and all forms of planning that involves natural resources (Smith, 2006). The program's primary aim is to achieve a sound matching of all forms of land use with the sustainable use of natural resources in the province. Operationally, it does this by mapping the natural resources, i.e., soil, vegetation, terrain, climate and water, and land potential (for various crops). It also provides a regional land use appraisal system to the end-user (Camp, 1999). KwaZulu-Natal has been mapped and described at three levels under the Bioresource Programme based on the spatial scale described from the largest to the smallest land management unit: Bioresource Group (BRG), Bioresource Unit (BRU) and ecotopes, (crop and veld) all contained within GIS DSS. In the context of agricultural productivity, soil ecotopes are the most relevant because they represent a class of land defined by soil form, texture, and depth, with a narrow range of possible farming enterprises, potential yield and the methods of production for each (Camp, 1999).

Therefore, the lucid benefits of the BRW to land and agricultural practitioners in meeting the objectives of sustainable land and agricultural management should be self-evident. Even so, the BRW is plagued by several material limitations. First, the computer support software of the BRW is obsolete and so has not seen a significant update in the last decade. This severy limits the interoperability with a range of efficient and robust GIS platforms. Second, soil ecotopes are not spatially defined and can only be mapped once they have been identified in a field survey. This limits the application of DSM approaches and quantitative accuracy assessment techniques in evaluating soil resource maps. Finally, the spatial extent of a BRU can vary from 10 000 ha to more than 100 000 ha, with resource information considered to be uniformly represented from wall to wall within the BRU. This approach to representing the landscape for inventorying land resources to prioritise agricultural practices was a revolutionary developmental agenda for the 20th century. However, the increased demand for a range of equally critical socio-economic priorities requires the assessment of land with better-refined detail and precision, enabling users to make the best-informed land use decision and maximising their benefits from their landholdings. Surprisingly, not much effort has been invested in reviewing or overhauling the BRW program, which indeed would be a massive undertaking. Still, the platform does allow for the refinement of geospatial approaches. It would readily benefit from the introduction of improved DEM for terrain unit representation and quantitative approaches to predictive soil and ecotope mapping.

These updates would provide a much richer perspective on land management priorities in KwaZulu-Natal from regional through to local level.

1.3 RESEARCH AIMS AND OBJECTIVES

This study aims to develop a sustainable and practical approach for integrating landform information into digital terrain analysis for soil assessment and mapping using advanced geomorphometric and remote sensing approaches. By emphasising the importance of land-type orientation coupled with soil-surveying techniques, I hope to broaden and strengthen the soil science community's understanding of how to optimise better and incorporate proven geospatial services and tools that have become increasingly mainstream in terrain mapping under regional circumstances. The overview of this research is summarised in the research design in **Figure 1.1.** The main objectives of the research were:

Objective1 (Chapter 2)

- Explore how spatial resolutions and search window generalisation influence the characterisation of topographic and terrain data generalisation and ultimately geopedological inputs to DSM.
- Compare the variation in topographic indices derived from the most common and widely accepted DEM against the base DEM data.
- Offer a simple yet effective solution for screening the selection of optimal DEM before terrain analyses and topographic feature characterisation.

Objective 2 (Chapter 3)

- Explore how spatial resolutions and geomorphon parameters influence the characterisation of geomorphon features and topographic covariate extraction.
- Compare the geomorphon surfaces and extracted soil-scape features from high spatial resolution soil data and the most common and widely accepted DEM surfaces.
- Highlight the operational pitfalls and practical considerations of utilising DEM surface products not fit-for-purpose to extract selected topographic properties to DSM & DGM.
- Quantitatively show that DEM resolution and geomorphon model parameterisation significantly influence geomorphon characterisation and generalised soil-landscape representation in the Drakensberg, KZN.

Objective 3 (Chapter 4)

- Develop a geospatial grid-based approach to quantify landscape geodiversity by introducing the idea of specific regional partial indices that contribute to landscape geodiversity.
- Highlight the overarching influence of pedological and topographic influence on geodiversity quantification.
- Evaluate the relationship between landscape geodiversity and mainstream land use planning spatial coverages currently in use for National and Provincial strategic land management.

Objective 4 (Chapter 5)

- Spatially define soil and crop ecotope regions for KwaZulu-Natal within the BRU Program using readily accessible GIS protocols

- Qualitatively assess the relationships between geomorphon units, derived from 3-arc second Shuttle Radar Topography Mission DEM, to better describe the spatial relationship between ecotope units and TMUs within the selected BRU currently limited to five landform descriptions.

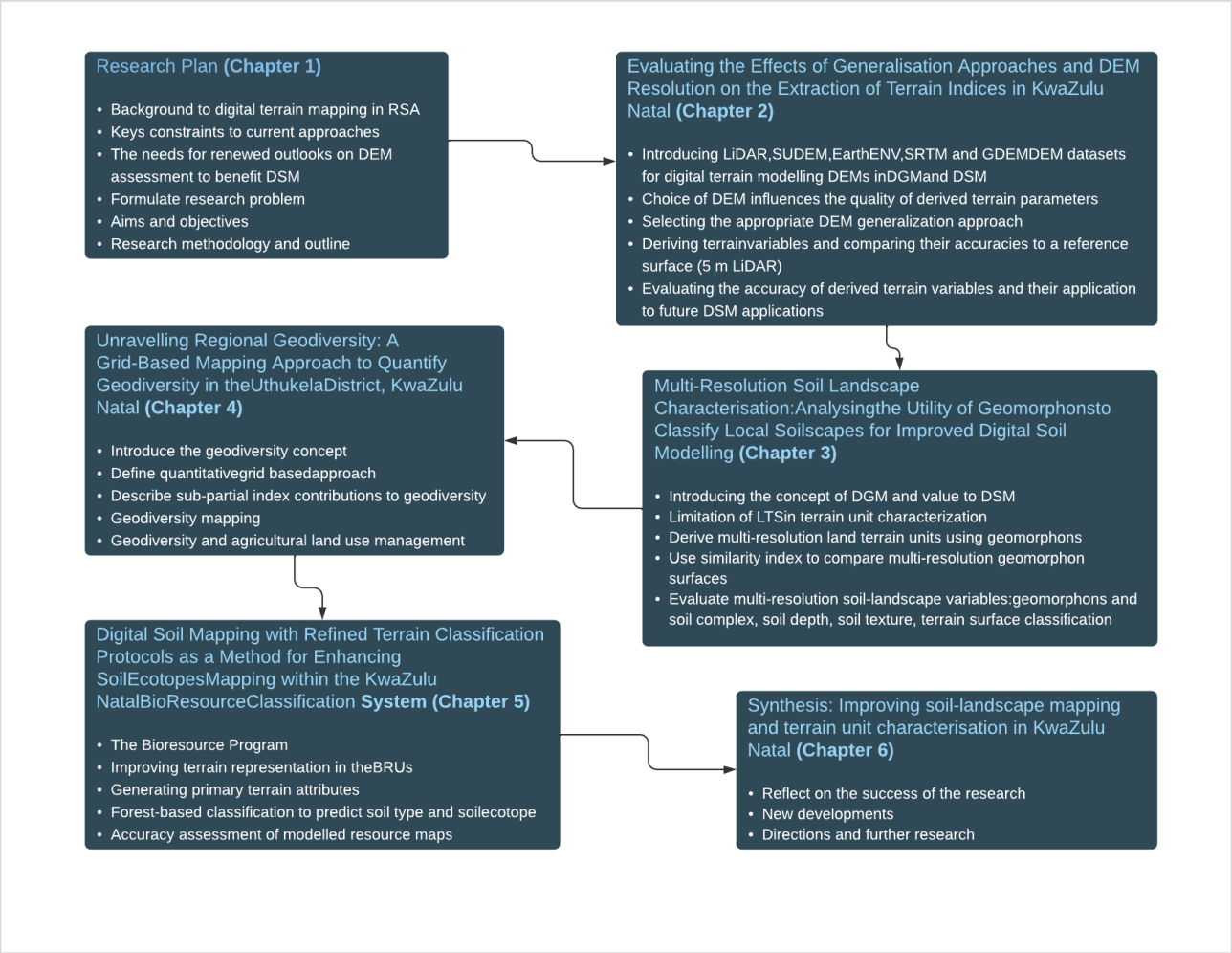


Figure 1.1: Research design workflow

CHAPTER 2

2 EVALUATING THE EFFECTS OF GENERALISATION APPROACHES AND DEM RESOLUTION ON THE EXTRACTION OF TERRAIN INDICES IN KWAZULU-NATAL[2]

"It is intuitively appealing to think that DEMs with high spatial resolution have low uncertainty. After all, they are better able to describe small scale landscape variation than DEMs with lower resolution. However, it is too simplistic to equate high resolution to high accuracy and coarse resolution to poor accuracy[1]

[1] Temme, A., Heuvelink, G., Schoorl, J., Claessens, L. (2009). Geostatistical simulation and error propagation in geomorphometry. *Developments in soil science, 33*, 121-140.

[2]This work was published in the South African Journal of Geomatics: Atkinson, J.T., Rozanov, A.B., De Clercq, W.P., 2017. Evaluating the effects of generalization approaches and DEM resolution on the extraction of terrain indices in KwaZulu-Natal, South Africa. South African Journal of Geomatics 6(2), 245-261, and as a book chapter in International Research in Environment, Geography and Earth Science Vol. 6: Atkinson, J.T., De Clercq, W.P., 2020. Assessment and Evaluating the Effects of Generalisation Approaches and DEM Resolution on the Extraction of Terrain Indices in Kwazulu Natal, South Africa. In: K.F.A. Lo (Ed.), International Research in Environment, Geography and Earth. Book Publisher International, London, UK.

The work was also presented at the annual Combined Congress of the Soil Science Society of South Africa, the South African Weed Science Society, the South African Society of Crop Production and the Southern African Society for Horticultural Sciences (SASHS), 23-26 January 2017, Bela Bela, Limpopo, South Africa as well as Geomatics Indaba, 21-23August 2017, Durban ICC, South Africa.

2.1 INTRODUCTION

The application of predictive digital soil mapping (DSM), defined here as the acquisition and population of spatial soil information by field and laboratory methods combined with spatial and non-spatial soil inference systems (McBratney *et al.*, 2003), has grown in popularity and has shown to be a viable alternative to traditional soil surveying and mapping approaches (Van Zijl *et al.*, 2014; Paterson *et al.*, 2015; Van Zijl & Botha, 2016; Wiese *et al.*, 2016). However, computer platforms, digital spatial covariates, and numerical models are not novel to soil mapping and have been used since the 1990s (Skidmore *et al.*, 1991; Bell *et al.*, 1992; McKenzie & Austin, 1993). However, in recent years, there has been significant interest in improving the scale and accuracy at which soil-landscape relationships (Hudson, 1992) are modelled and represented, from a local through to a national scale (Maynard & Johnson, 2014). The origins of the soil-landscape concept are based on C.F. Marbuts (Marbut, 1937) idea that soils are landscapes and profiles. Typically, surveyors rely on their expert knowledge, combined with ancillary site descriptions, to develop conceptual and preliminary soil-landscape models that are later refined. This traditional approach is entrenched in the concept of soil-landscape units, similar to landforms but more precisely defined by Hudson (1990). Concomitantly, high on the research agenda is the need for developing novel ways of exploring the complexity associated with deterministic characterisation and modelling approaches of topographical landforms (Bishop *et al.*, 2012). An added benefit would be to relate various soil properties to readily available spatial data such as digital elevation data (Mashimbye *et al.*, 2014; Atkinson *et al.*, 2017).

Digital Elevation Models (DEM) provide a convenient and representative interpretation of the Earth's surface. The DEM is a 3D mathematical (or digital) model of the terrain surface (Li *et al.*, 2004) provided in a raster (grid) data model that presents the end-user with a 'field view' representation of the landscape, defined as a regular rectangular array of cells with some aggregate value of the terrain recorded for each cell. The resolution of the data stored in this format is, in turn, a function of the grid cell size, or DEM pixel resolution (De Smith *et al.*, 2007). DEMs are generally considered the *de-facto* dataset(s) for various terrain and spatial analyses. Therefore, they are integral for constructing a principle digital topo-sequence (catena) for DSM (Malone *et al.*, 2017). The topographic studies of endogenic and exogenic pedo-geomorphological processes using DEMs are well established and are central to the catenary concept for soil formation (Hook & Burke, 2000). All widely-recognised soil mapping paradigms for rationalising the spatial distribution and dynamics of soil properties within a region acknowledge topography as the dominant pedogenetic state forming factor – the most temporally stable and influential component of geosystems. These include the conceptual soil-forming factor theory proposed initially by Dokuchaev (1892) and later refined by Jenny (1941), or the more well-known *SCORPAN* function model (McBratney *et al.*, 2003) or even the enhanced quantitative STEP-AWBH model proposed by Grunwald (2010). DEMs are particularly effective at representing land components as landform elements with constant elevation values or having constant values of several readily interpretable morphometric variables with borders separated by lines of feature discontinuities (Minar & Evans, 2008). Studies have reported using different sets of morphometric variables as predictors ranging from only elevation (Leenaers *et al.*, 1990; Baxter & Oliver, 2005) to as many as 69 variable morphometric indices (Behrens *et al.*, 2005). Still, the most popular indices are elevation, slope gradient and topographic index (Florinsky, 2016).

Land component boundaries are highly correlated with both geodiversity (abiotic) and biodiversity (biotic) environmental properties (MacMillan *et al.*, 2004; Van Niekerk, 2010). Therefore, using scale-specific DEMs to derive key topographic indices specifically related to soil formation is essential for spatially representing the statistical associations with quantitative soil physical, chemical and biological properties and soil descriptive taxonomic units across varying topographic phenomena (Maynard & Johnson, 2014). DEMs are particularly crucial for describing the processes of soil genesis, or pedogenesis, in the spatial domain. However, as Debella-Gilo & Etzelmüller (2009) points out, none of the current DSM methods or sets of secondary data

used for DSM is necessarily simple to apply, free of defects and generically applicable. Despite much success since its inception, DSM remains a paradigm in progress. Continual application of DSM to accurate, translatable and pragmatic outcomes requires a solid understanding of the proper spatial scale and grid resolution application for feature (model covariate) selection and information generalisation. Indeed, there remains an opportunity to explore the limits of uncertainties and complexities using different methods and data sources under varying spatial scenarios

Since the inception of DSM, a constant limiting factor has been selecting the appropriate spatial scale of analysis and representation. According to Addiscott (1998) and Atkinson & Tate (2000), this remains partially unresolved due to diverging meanings of multi-level feature representation complexity. Cavazzi *et al.* (2013) draw on an extensive range of sources to highlight that due to the complexity of the scale-resolution dichotomy, determining this optimal combination of scale-relationships is often dependent on the context of the study and remains in large an unresolved issue with limited practical guidelines available. The main challenges with DEM pixel resolution are that first, at finer resolutions, the terrain variables contain too much unnecessary detail or "noise" that may lead to lower modelled accuracy. Secondly, at coarse resolutions, terrain variables may show too much generalisation or "smoothing" and not adequately represent terrain attributes or the land surface, ultimately compromising soil-landscape models' predictive capacity (Cavazzi *et al.*, 2013).

Furthermore, soil-landscape pattern-process functions vary between spatial scales and may even be independent of each other (Phillips, 1988). This should concern unenlightened practitioners relying on terrain-based modelling for DSM since terrain attributes may only be valid within a limited range of spatial scales. Constraints include either the observed scale-dependent soil-landscape pattern-process or because the models used to predict and represent the environmental phenomena themselves are scale-dependent (Zhang & Montgomery, 1994; Florinsky & Kuryakova, 2000). Consequently, a persistent research agenda item for DSM is defining the range of data resolution for which DEM models are applicable. DEMs vary in resolution and accuracy by production method used. Thus, the DEM characteristics directly impact the spatial traits of the resulting soil maps. Choosing the most representative DEM is critical to any soil-landscape and DSM analysis (Zhang *et al.*, 2008). As Wechsler & Kroll (2006) discuss, many DEM users accept that DEM uncertainty affects their applications' outcomes. However, DEM error effects on elevation and derived parameters or even landscape analysis, including DSM, are often not quantified or integrated by DEM users since methods to address DEM error have not been systematically combined with GIS software packages.

Moreover, many freely available DEM products represent elevation surfaces that may provide visually acceptable interpretations of the land surface built into them. Still, they may not be scale-appropriate or provide sufficient data accuracy for spatially realistic models to adequately represent the soil-landscape continuum (Vaze *et al.*, 2010). As Boixadera *et al.* (2016) point out, the quality of the available soil-landscape model is of supreme importance to the quality of the mapping results, whether using conventional soil mapping techniques (Lagacherie *et al.*, 1995; Legros, 1996; De Bruin *et al.*, 1999), DSM approaches (McBratney *et al.*, 2003; Lagacherie, 2008) or even of landscape evolution models (Minasny *et al.*, 2015). With DSM, there is a general acknowledgement that users are increasingly challenged to integrate available technologies to provide more objective decision-support systems in the management and monitoring of soil resources. However, this must recognise the fundamental issues of scale effects in data and data processing - the essential aspects of any research (Quattrochi & Goodchild, 1997; Baumgardner, 1999).

DEM data may not always readily meet users' demands for terrain analysis at diverse levels and various applications. So, they require complementary pre-processing approaches to generate new data with different resolutions. This process is usually achieved through map generalisation (Ai & Li, 2010). Map generalisation is the simplified representation of digital elevation details appropriate to the scale or the map's purpose. The

new model remains a proper substitution of reality or a copy of the physical environment to a reasonable interpretation. A central consideration for the DEM application of generalisation approaches is to derive a less detailed terrain surface from a more detailed one. This may be necessary to reduce data volume, mitigate data redundancy issues, or when soil-landscape studies require coarser analytical scales of assessment (Zhou & Chen, 2011). Therefore, the generalisation process should be seen as an information abstraction rather than a data compression intervention (Ai, 2007). The integrity of the metric quality of terrain features is still retained.

Specific terrain attributes, such as curvature, topographic wetness index (TWI) and surface roughness, demand a more detailed DEM, which can be generated by integrating additional higher quality data (Podobnikar, 2005). In contrast, terrain attributes such as slope gradient, slope aspect and terrain texture are optimised by DEM with fewer details produced by the original DEM's generalisation. Both Zakšek & Podobnikar (2005) and Ai & Li (2010) provide comprehensive reviews of the most popularised DEM generalisation approaches. The most common is nearest neighbour resampling of a high-resolution, square-grid DEM for discrete data or bilinear and cubic interpolation for continuous data. These studies further highlight two critical aspects of DEM generalisation. Firstly, the influence that various generalisation approaches have on final predicted DEM accuracy and, secondly, how DEM origin (sensor platform) influences terrain representation accuracy by generalisation. McMaster & Monmonier (1989) further extend the description of raster generalisation tasks into four primary categories: a) structural generalisation, b) numerical generalisation, c) numerical categorisation, and d) categorical generalisation. Their research provides an outline of fundamental GIS-based operators and evaluates various techniques used for raster-mode generalisation. Therefore, DSM practitioners must understand that generalisation approaches used for altering the scale and representation of DEMs contain inherent interpolative errors, and these errors can propagate through to the model predictions.

Gallant & Hutchinson (1997) highlight the practical limitations of DEM generalisation by reporting an overall decrease in terrain attributes' accuracy due to the discretisation effects of altering DEM grid sizes through DEM coarsening (smoothing). Additionally, a study conducted by Zhang *et al.* (2008) concluded that DEMs with different resolutions and accuracies generated various topographic and hydrological surface attribute values, affecting overall soil erosion model performance. Most conventional studies employ a three by three moving window neighbourhood approach to calculate DEM terrain attributes (Anderson *et al.*, 2006; Smith *et al.*, 2006). However, as Thompson *et al.* (2001), Maynard & Johnson (2014), and Atkinson *et al.* (2020) point out, as grid-size increases, so too do the neighbourhood extents, thereby making it difficult to evaluate the effects of changing grid-size and neighbourhood extent on DEM accuracy. Therefore, the study conducted by Cavazzi et al. (2013) provides necessary insight into DEMs' behaviour to various scale and neighbourhood algorithms. Their research highlights that coarser-resolution DEMs (140 m) performed better in low-detail flat homogenous areas, regardless of window search sizes. Whereas in morphologically more varied regions with abrupt changes in topography, finer resolution DEMs (30 m) with smaller windows outperformed coarser DEMs and larger window search radii.

Presently, the primary inventories of curated open-source synoptic DEM data available to the public for use in service-oriented soil and landscape mapping are provided by either the National Aeronautics and Space Administration NASA/ US Geological Survey (USGS) (https://lpdaac.usgs.gov/), the Consortium for Spatial Information (CSI-CGIAR) (http://srtm.csi.cgiar.org/), the National Oceanic and Atmospheric Administration (NOAA)(https://www.ngdc.noaa.gov/mgg/global/) and the Ministry of Economy, Trade, and Industry (METI) of Japan (https://asterweb.jpl.nasa.gov/gdem.asp). These platforms remain popular for ease of access, negligible or *gratis* cost, and complete global coverage at medium to high resolutions. However, a significant caveat to their utility is the inherent error in these DEMs for specific regions, resulting from the data source (satellite) and the creation process (Shi *et al.*, 2012). An alternative to the freely available DEM products is the DEM created from high-resolution Light Detection and Ranging (LiDAR) data. LiDAR, or interferometry, data represents the point-cloud results of fast pulses of a focused infrared laser from a remote object, either a

plane or UAV flying over a specified flight path. The distance between the ground and the plane is determined by measuring the first and second-return times between the transmitted and received signal. In turn, the actual (true) surface elevation is derived (French, 2003).

Consequently, the data collected by LiDAR have an excellent resolution, both horizontally and vertically, with a high level of terrain accuracy. Because of their fine resolution and precision, LiDAR-based DEMs are naturally considered the gold standard for many DSM applications. The ability of sensors such as LiDAR to readily provide very high-resolution bare surface models has allowed researchers to investigate whether terrain attributes derived at finer resolutions are better correlated to pedo-geomorphic properties at high-resolutions or generalised, lower resolutions (Barber & Shortridge, 2005; Chow & Hodgson, 2009). Notwithstanding, Thompson *et al.* (2001) contend that higher resolution DEMs might not be necessary for producing functional soil-landscape models and often provide high levels of signal-to-noise ratio artefacts due to the high volume point-cloud data. Although a far superior product in terms of representing terrain-based phenomena, the volume of data acquired and the often repetitive and insignificant contribution of terrain discontinuities to topographic attributes combined with the pre-processing requirements creates a problem for inexperienced and practitioners. Both Anderson *et al.* (2006) and Liu (2008) further point out that the added technical advantages of LiDAR come with their own inherent operational and financial limitations that ensure that open-source platforms remain routinely attractive alternatives where applicable. These DEM datasets include the readily available high-resolution near-global DEMs such as the shuttle radar topography mission (SRTM), the global EarthEnv-DEM90 and the more region-specific Stellenbosch university digital elevation model (SUDEM). Predictably, current research hypotheses on the contribution of LiDAR to DSM seems varied. In particular, Hodgson *et al.* (2005) reported that specific terrain matrices derived from LiDAR displayed higher accuracy than SRTM DEM data derived from synthetic aperture radar (SAR). These DEMs, therefore, need further assessment to compare the quality of specific terrain products derived under varying generalisation methods to resulting spatial resolutions. This should allow users to quantitatively evaluate terrain feature error before DEM selection for inclusive digital geomorphometric, soil-landscape and DSM applications in South Africa. The lack of direct local evidence that they can significantly contribute to quantitative DEM surface error analysis prohibits an easy decision of prioritising a type of DEM for service-oriented soil mapping and soil-landscape applications. At the same time, comparisons between LiDAR-based DEM and satellite-based DEM are well documented in the literature, e.g., Hodgson *et al.* (2005), (Hengl & Reuter, 2011) and (Brubaker *et al.*, 2013), very few studies offer a comprehensive and intuitive assessment of DEM source with terrain generalisation approaches within a southern Africa context. Thus, their results are hardly applicable to regionalised soil-landscape and DSM applications. There is a definite need to quantify the impact that generalisation approaches have on simplifying detailed DEMs and to compare the accuracy and reliability of results between high resolution and coarse resolution data on the extraction of localised topographic variables as a primer for future digital landscape assessment. Therefore, a key research question is how the topographic parameters computed from DEMs are affected by the DEM resolution, DEM source and DEM generalisation and how can they be compared.

In this regional study, the relationship between DEM sensor efficiency on the extraction of selected local terrain characteristics such as terrain altitude (*Elev*), slope aspect (*SlopAsp*), slope gradient (*SlopGrad*), profile curvature (*ProCurv*), planform curvature (*PlanCurv*), topographic wetness index (*TWI*), terrain surface texture (*TST*), LS factor (*LS*), and lastly terrain ruggedness index (*TRI*) is investigated. This study builds on the clear research incentive to investigate how DEM source, resolution, and terrain generalisation affect the accuracy of the derived terrain-based attributes. Specifically, I benchmark the performance of the SRTM 1-and 3-arc second, the SUDEM and the EarthEnv DEMs, at different spatial resolutions, to the results of an accurate surface model using three generalisation algorithms, namely: mean cell aggregation (*MA*), nearest neighbour resampling (*NNR*) and hydrological corrected topo-to- raster (*HCD*). The most suitable resolution of DEM to apply in soil-landscape predictions depends on the scale of the processes controlling pedogenesis and landform

relationships and is ultimately landscape-dependent (Cavazzi *et al.*, 2013). Consequently, there is a real need for more in-depth assessments of the quality of landscape-level derived digital terrain covariate data. The purpose of this research is to demonstrate and evaluate how suitable sensor choice and spatial scale influence DEM surfaces that are "fit-for-purpose" for describing the scale-dependent pedo-geomorphic inputs relevant to soil-landscape modelling. More specifically, the study identified two primary objectives. First, compare the variation in selected topographic indices derived from high a resolution base DTM (1 m LiDAR data) using three generalisation approaches across different resampled surface resolutions. Secondly, I compare these LiDAR-derived products with those derived from the most common and widely accepted regional and near-global DEM surfaces to examine the variation in topographic representation within a sub-tropical eastern coastal region of the Republic of South Africa.

2.2 MATERIALS AND METHODS

2.2.1 Study site

The study area is the commercial forestry plantation known as Braemar (*30º 23' 18.53''E, 30º 22' 35.06" S*) situated in the vicinity between Kenterton and Braemar, south of the town of Durban in the Ugu District Municipality within the Kwazulu-Natal Province of South Africa (**Figure 2.1**). The mono-cultured Eucalyptus dominated plantation is owned and managed by Sappi Southern Africa's Forestry Division and covers a region of approximately 5 200 ha. The study site was selected due to it being one of the few inland/coastal forestry plantations with LiDAR data that could be combined with the gridded soil data from many of Sappi's wood plantations, including Braemar. The region has a temperate sub-tropical climate characterised by warm, humid temperatures and summer rainfall, with most rainfall experienced during the spring and summer months (October to March). The highest rainfall is typically experienced over December and January of any year. Mean monthly precipitation for the summer rainfall period ranges from 113-120 mm, whilst the mean annual precipitation (MAP) is just under 1 000 mm with an average annual humidity of 73% (Camp, 1995). Climatic conditions vary significantly between the coast and inland, with temperatures associated with the coastal areas being moderated by the warm Indian Ocean's effects. Average summer temperatures range from 21-23 ºC, with winter temperatures reaching approximately 16 ºC.

The terrain morphology within the study area is both diverse and varied, typical for KwaZulu-Natal. The Drakensberg mountain range inland and to the west has many rivers that drain from its watershed into river catchments that follow an easterly course dissecting the corridor before draining into the Indian Ocean on the KwaZulu-Natal South Coast. The river catchments dominate the landscape, creating a closed hills and mountains terrain, with a moderate to high relief, deep valleys and hilltop plains characterised by extensive river gorges and undulating hills. The region generally has a rolling or undulating topography dominated by moderate to steep slopes (> 12%) (Macfarlane *et al.*, 2014). The altitude range for the region varies from sea level to 950 m, while in the study site, the altitude range is from 200 m to 820 m.a.s.l. The geology is equally diverse and directly influences the topography of the region. The region is underlain by a distinctive combination of geological forms ranging from Granite and Gneiss of the Natal Metamorphic Group, which occur along the coastline, to fluvial Sandstone of the Cambrian to Ordovician Natal Group that characterise the many table-top cliffs in the region. Glacial deposits comprised of thick units of Tillite of the Dwyka Group overlay the Natal Group sediments and dominate the Mkomazi and Mzimkulu rivers' southern areas. Smaller pockets of Ecca Shale outcrops can be found spread across the coastline (Henning, 2018).

Pedologically, shallow, weakly weathered soils and yellow-brown block clay subsoils and yellow mottled waterlogged subsoil dominate the topographic and geological catenary sequence within the study site (SASRI, 1999). Major soil groups include, without restriction, combinations of Haplic Leptosols, Cutanic Lixisols and Leptic Cambisols that dominate the crests and midslope convex creep slopes with *Rhodic Ferralsols*, *Cambisols* and *Arenosols* as well as *Xanthic Ferralsols*, *Arenosols* and *Cambisols* dominating the well-drained flat regions and the lower altitude regions with lower-lying regions displaying continuous bands of *Xanthic* and *Rhodic* Hydromorphic pedogenesis. The study area straddles quaternary catchments U80G and U80H, which form part of the South and Central Coast River hydrologic reference catchment of the Mvoti to Umzimkulu Water Management Area. Two main perennial rivers, the Fafa and Hlangankulu (DWS, 2013), drain into the catchment with several other seasonal streams completing the catchment's hydrographic character.

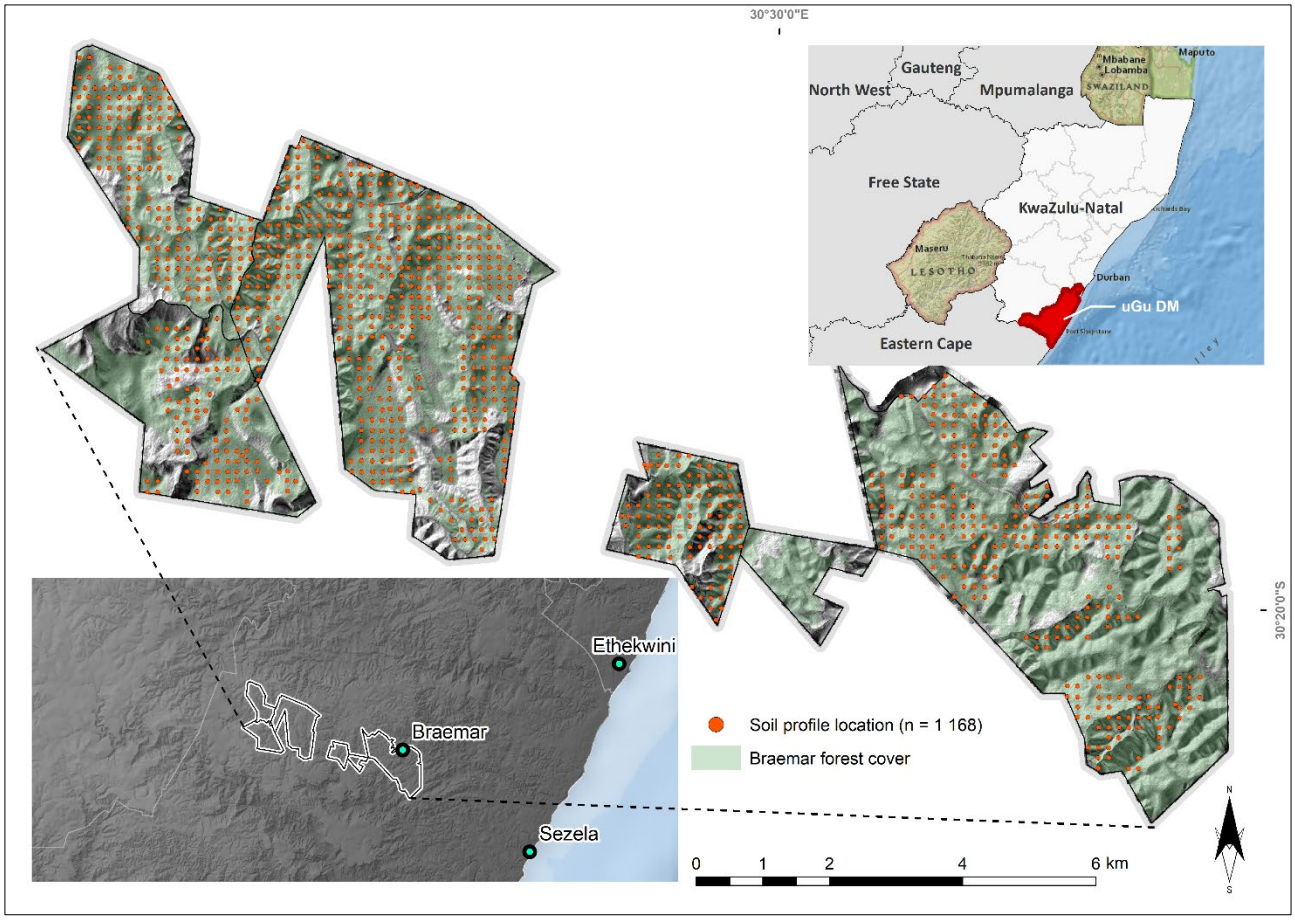


Figure 2.1: Location of study site near Braemar (KZN) showing the locations of the soil profile sites of the Forestry Soil Database. The red region in the upper-right inset map represents the Ugu District Municipality.

2.2.2 DEM datasets

2.2.2.1 Airborne LiDAR

Airborne-based LiDAR for the study site was acquired using the Leica ALS50-2 multi-phase laser scanner with a CUS6-"uIRS" inertial measurement unit and XP embedded data logger on the 24 and 25 of January 2015 by Land Resources International Pty (Ltd). The LiDAR imagery was collected at an average flight height of 800 m above ground level (AGL) with a swath width of 324 m and a minimum flight line overlap of 50% and 25° field of view. Point cloud densities ranged from 4.3 pulse m^{-2} to 13 pulses m^{-2} with an overall average density of 7.4 pulses m^{-2}. The vendor did all post-processing of the acquired data using Terrapos® (TERRATEC, 2016) to process the raw LiDAR data and final trajectories to projected format and correct positional accuracy. The TerraScan® suite of tools (TERRASOLID, 2016) was then employed to process, classify and clean the point cloud data to yield the final ground and non-ground LiDAR products. Extensive QA/QC was conducted on the LiDAR data to ensure that the correct geoidal model (SAGeoid2010) was applied as a height correction from ellipsoidal height mean sea level.

Furthermore, the survey was conducted using a real-time fix to the TrigNET and VRS broadcast system with horizontal calibration using the SCOT (Scottburgh) TrigNET station and the Hartebeesthoek 94 Datum. The final 1 m resolution LiDAR DTM product was exceptionally good at representing the terrain surface with an overall maximum height error limited to 0.17 m with an average magnitude of 0.06 m and a root mean square error (RMSE) of 0.08 m across all elevation data collected.

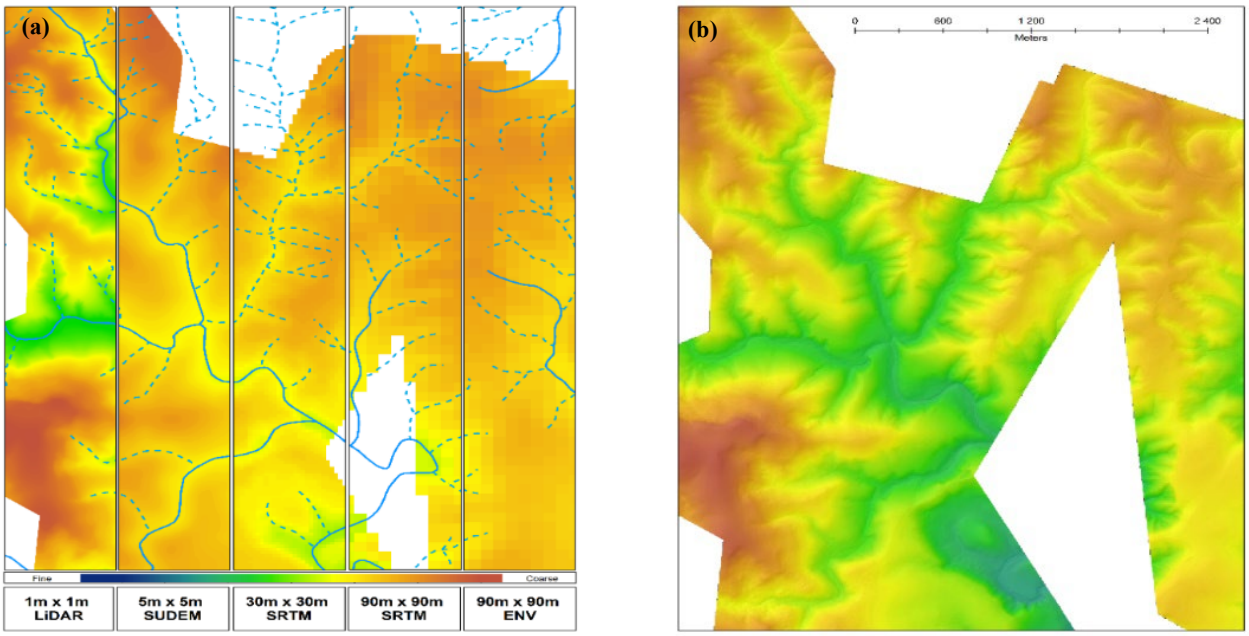


Figure 2.2: (a) Image-slice comparison of LiDAR with non-LiDAR DEM surfaces for Braemar. Note the absence of drainage lines and reduced detailed as pixel resolution decreases compared to the original LiDAR DTM (b).

The original LiDAR DEM, including the resulting LiDAR-derived (LD) products, were compared to the same extracted topographic attributes from other freely available broad-scale satellite-derived DEMs. These medium-resolution elevation products, hereafter referred to as non-LiDAR-derived (NLD) products, included: both the global 1 arc-sec (30 m) and 3 arc-sec (90 m) void-filled Shuttle Radar Topographic Mission (SRTM V4.1) (Jarvis *et al.*, 2008) surface models, the 5 m Stellenbosch DEM (SUDEM) (Van Niekerk, 2015) and a relatively new global DEM known as the EarthEnv-DEM90 (Robinson *et al.*, 2014) derived from fusing three freely available elevation datasets namely the 1 arc-sec Advanced Spaceborne Thermal Emission and Reflection Radiometer Global DEM2 (ASTER GDEM2) (NASA, 2011), 3 arc-sec CGIR CSI SRTM v4.1 and 3 arc-sec Global Land Survey Digital Elevation Model (GLSDEM)(USGS, 2008) into a quality enhanced grid of elevation estimates covering approximately 91% of the globe. **Figure 2.2** highlights the difference in terrain elevation detail. The LiDAR surface at 1 m displays the most elevation detail with smoother surfaces visible as DEM resolution decreases, providing a glimpse of the variation expected in terrain parameter extraction due to differences in pixel resolution. Note that as the pixel resolution decreases, the detail of landscape features such as drainage lines, valleys and even ridges and hilltops also decreases, to the point where certain features may not even be represented. Nelson *et al.* (2009) have compared DEM data across various data sources, and their findings provide a valuable comparison of the major specification for the DEM models used in this study. **Table 2.1** has been adapted to highlight only the critical components of the DEM data sets used in this study. For detailed descriptions and further applications of the sensors used in this study, readers are referred to studies conducted by (Farr & Kobrick, 2000; Rodriguez *et al.*, 2005; Farr *et al.*, 2007; Prasannakumar *et al.*, 2011; Gesch *et al.*, 2012; Van Niekerk, 2012).

Table 2.1: Key properties of datasets used in the study (Modified from (Nelson *et al.*, 2009) in (Wilson, 2012).

Source	Resolution (m)	Accuracy	Footprint (Km^2)	Post-Processing required	Elevation/Surface
LiDAR	1 to 3	0.15–1 m vertical, 1 m horizontal	User-defined	High	Surface
SUDEM	2 to 5	2.1-4.6 m vertical	Provincial/ User defined	Low	Surface/ Elevation
SRTM, Band C	90, 30	16 m vertical, 20 m horizontal	Near global, 60° N to 58° S	Potentially high	Surface/ Elevation
SRTM, Band X	30	16 m vertical, 6m horizontal	Similar to Band C, but only every second path is available	Potentially high	Surface/ Elevation
EarthEnv90-DEM	90	7–50 m vertical, 6–50 m horizontal	Near global	Medium	Surface/Elevation

2.2.2.2 EarthEnv-DEM90

The EarthEnv-DEM90 (hereafter EarthEnv) is the result of the union of the following DEM datasets: the 1 arc-sec ASTER GDEM2, 3 arc-sec CGIAR CSI SRTM and 3 arc-sec GLSDEM. EarthEnv aims to provide a product that combines the advantages of the source DEMs (high quality of the SRTM data and global coverage of the ASTER GDEM2) whilst reducing the inherent shortfalls of their respective pervasive quality issues. This chapter will not address the methodological detail of how the three datasets were smoothed, blended and filtered other than mentioning that an adaptive multi-scale Gaussian smoothing algorithm was applied using a weighted average approach to fuse the input DEM datasets. Readers are referred to Robinson *et al.* (2014) for a detailed description of the EarthEnv genesis. The final 90 m DEM compilation surface consists of different zones relating to the input datasets' respective mosaiced coverage.

As a result of rigorous quality control techniques in both the ASTER and SRTM zones, the vertical and horizontal accuracy is similar to that of the original GDEM2 and SRTM v1.4 datasets with RMSE accuracies of 0.99 and 0.99 for the fused ASTER GDEM2 and SRTM v1.4 products, respectively. The final EarthEnv DEM is, however, not without limitations. Contemporaneous issues that plague ASTER GDEM2 and SRTM v4.1 products, particularly erroneous spikes and wells, are carried through to the respective “zones” in the DEM surface. The basis for this limitation is that the adaptive multi-scale smoothing algorithm applied to EarthEnv is geared toward noise-reduction rather than the correction of erratic and bizarre elevation values. To my knowledge, no regionalised study has investigated the utility of using a fused global DEM such as the EarthEnv90 product for local terrain variable extraction. Although the final EarthEnv DEM is well correlated with its standalone input datasets, it is still necessary to evaluate whether the fusing of the individual input datasets has impacted the variance in topographic attribute outputs within a regionalised coastal context such as Braemar. To better understand the primary DEM specifications of the EarthEnv, an overview of its respective constituents is necessary, and a brief description of the specification for each dataset are listed hereafter.

2.2.2.3 SRTM

The SRTM v1.4 (hereafter SRTM) project, managed by the CGIAR, and freely distributed by the USGS, provides global topographic data covering approximately 80% of the Earths terrain at 1- and 3 arc-second resolution (30 m and 90 m along the equator, respectively) with a reported vertical RMSE of 16 m and horizontal circular error of 20 m at a 90% confidence interval (Farr & Kobrick, 2000; Rabus *et al.*, 2003; Rodriguez *et al.*, 2005; Farr *et al.*, 2007). Both SRTM DEMs were acquired using an interferometry technique with C- and X-band synthetic aperture radar (SAR) to derive a digital surface model where elevation heights represent the top of reflective features such as vegetation, anthropogenic features or bare soil, and are defined above the WGS 84 geoidal datum (Jarvis *et al.*, 2008; Prasannakumar *et al.*, 2011). This top-of-canopy limitation can be problematic depending on the end-user requirements as landcover features may influence elevation data. Earlier versions of the SRTM DEM data were known to have large data voids where water or heavy shadow prevented the quantification of elevation (Sharma *et al.*, 2010). These errors were usually less than 5 data pixels in size (Hall *et al.*, 2005) and were associated either with flat water bodies or steep slopes in mountainous terrain and incredibly limiting in fields of hydrological modelling.

However, due to their widely used applications, there has been continual data improvement efforts and rigorous quality control of the SRTM products (Jarvis *et al.*, 2008). This has led to a highly trusted and widely used data product with the current version(s) of the SRTM now offered as void filled products, and these include both the 30 m and 90 m elevation datasets used in this study.

2.2.2.4 ASTER DEM

The ASTER GDEM2 (hereafter GDEM2) is a product of a joint venture between NASA and METI. The DEM was acquired aboard NASA's Terra Satellite in 1999 using downward and reward facing single-pass stereoscopic correlation techniques and cameras capable of capturing overlapping near-infrared images at virtually any global extent. By augmenting an additional 260 000 image scenes to the already compiled 1,2 million scene-based DEMS, the GDEM2 can improve on its predecessor, the ASTER GDEM1, through the use of smaller correlation kernels to reduce artefact incidence and improve spatial resolution and the accuracy of water masking (Tachikawa *et al.*, 2011). The GDEM2 is a complete mapping product of the earth surface, covering almost 99% of the planet with 14 different electromagnetic spectrum bands, ranging from visible to thermal infrared light. The absolute vertical accuracy was calculated by comparing the GDEM2 data with 18 000 geodetic ground control points from the US National Geodetic Survey as well as pixel-to-pixel based differencing with other 1 arc-sec DEMs for conterminous United States (CONUS), the US National Elevation Dataset (Gesch & Maune, 2007) and SRTM data (Farr *et al.*, 2007) with supplementary calibrations from global space-borne laser altimeter data. To fully characterise the vertical accuracy of the GDEM2 surface, GPS benchmarking with surface control points was included for the full extent of CONUS, similar to the validation approach conducted for the original GDEM Version 1 in 2009 (Gesch *et al.*, 2012). The absolute vertical accuracy reported for the GDEM2 is 7 m at a 95% confidence interval with RMSE improvements in horizontal resolution between 71 and 82 m, making the GDEM2 comparable to the SRTM 30 m DEM on a pixel-by-pixel basis. Although void and artefact removal is significantly improved in GDEM2 compared to GDEM1, the data is still characterised by a variety of anomalies and artefacts that require processing before use, particularly at higher latitudes (> 60N) and lower latitudes (< 60S) (Robinson *et al.*, 2014). It is also clear that GDEM2 still includes non-ground level elevations for areas that have above-ground features (tree canopies and built structures). Additionally, the increase in vertical and horizontal accuracies come at the cost of increasing noise

in the datasets as well as sensitivity to land cover and reflective surfaces, with accuracy measurement being lower for forested areas, buildings etc. compared to bare areas (Tachikawa *et al.*, 2011; Gesch *et al.*, 2012).

2.2.2.5 GLSDEM

The GLSDEM is a compilation of several public and commercial DEM products of varying resolution and quality produced by NASA and the USGS to support the *Global Land Cover Facility*. All contributing datasets are resampled using a cubic convolution algorithm to 3 arc-sec pixels and collaged together to form a global DEM product. Strict data merging and validation techniques were not applied to create the GLSDEM since the product's primary focus is to provide ancillary surface data to help fill voids in high latitude GDEM2 (Robinson *et al.*, 2014). The final GLSDEM product is supplied in a UTM projection, WGS 84 horizontal datum referencing the EGM96 geoid. The GLSDEM data used in the EarthEnv DEM fusion is the WRS-2 with 185 km x 185 km tiles span and has a grid spacing of 90 m. Since the GLSDEM is only used to supplement the GDEM2 by in-filling voids in high latitude areas, a detailed description of the accuracy and quality of GLSDEM is not necessary for this study.

2.2.2.6 The SUDEM

The SUDEM 2015 (hereafter SUDEM) is offered as a series of products developed by the Centre for Geographic Analysis (CGA) at Stellenbosch University, South Africa. By processing large scale contours and spot heights with local interpolation approaches, high-resolution DEM products are derived and available as four different products: levels 1, 2, 3 and 4 (Van Niekerk, 2012). The level-1 product(s) rely solely on using local vertical contour intervals and spot heights as input variables to derive the final DEM resulting in a 5 m resolution bare surface model or DTM. The level-2 product(s) augment the contours and spot heights with the 1 arc-sec SRTM DEM (research-grade version) in lower relief areas, with fewer contour and spot height densities, resulting in a 5 m DTM that is sporadically influenced by reflective surfaces. Elevation data for the level-3 product(s) are extracted from 0.5 m resolution stereo aerial images, only where extraction failures occurred due to shadow issues, resulting in a 1-2 m resolution digital surface model (DsM). Finally, the level four product(s) are presented as a 1-2 m DsM (virtually the Level 3 product) that has been converted to a DTM by removing all objects with heights of more than 0.5 m (Van Niekerk, 2015). The SUDEM product suite provides perhaps the most accurate and readily available localised high-resolution DEM for South Africa. Although there are spurious issues relating to contour and spot height values, many of these errors (gaps) do not feature in the final interpolated product. Most of the more significant errors were corrected in preparation for the SUDEM 2015 product. Issues such as banding and wave effects may persist in Level-1 products; however, these may be rectified by employing generalisation or filtering approaches to reduce these artefacts. The expected SRTM 30 m DEM errors addressed earlier in this chapter (see section *2.2.2.3*.), such as voids and spikes, are corrected in the SUDEM 2015; however, issues relating to banding and striping have not been comprehensively addressed. Finally, geospatial processing limitations to the pre-processing software at the time of publication resulted in the SUDEM 2015 not being hydrologically corrected to remove sinks and peaks from the original DEM spatial constituents. Overall SUDEM data accuracy is assessed by systematically comparing the DEMs with available reference elevation data such as surveyed points obtained by a global network of satellite systems (GNSS) for the City of Cape Town. Vertical accuracies for the four products range from 0.35 to 1.77 m. Simultaneously, the SRTM data used in this study is reported to have a vertical accuracy of 3.22 m confirming Van Niekerk (2012) suggestion that improvements in relative and vertical accuracy can indeed be achieved when combining SRTM DEM data with contours and elevation points in flat areas.

2.2.3 Data preparation

This study's first objective is to evaluate the effect of DEM resolution and spatial extents on the extraction of selected soil-landscape relevant terrain attributes. This is done using a similar approach proposed by Gillin *et al.* (2015) and Maynard & Johnson (2014) by comparing the values of terrain attributes derived from a reference surface, at 1 m resolution using the airborne LiDAR DEM surface, to the same terrain-attribute values of six alternative resolution LiDAR-derived DEMS, at 5, 10, 20, 30, 60 and 90 m grid resolution intervals, hereafter referred to as LiDAR-derived (LD) products. These lower resolution LiDAR DEMs are further committed to three specific interpolation methods; namely, *mean aggregation* (MA), *nearest neighbour resampling* (NNR) and a *hydrological corrected topo-to-raster* approach (HCD) to determine which generalisation method and pixel resolution combination best correlate with the 1 m LiDAR reference terrain parameters (true ground measurement). The result of this analysis is 22 unique resolution-generalisation combinations in total, with 252 topographic variables extracted. Cavazzi *et al.* (2013) point out that the most suitable resolution of DEMs applied to soil-terrain predictions depends on the scale (resolution) of the processes driving pedogenesis in the landscape since distinct pedogenic factors and landscape processes are applicable at different spatial scales (Florinsky & Kuryakova, 2000). It is worth outlining then, *ex-ante,* that particular resolutions of terrain attributes will be better suited than others for soil-terrain predictions, predominantly at a landscape or regional level. This study's extent-neighbourhood DEM combination results will allow further insight into grid resolution effects using conventional interpolation approaches where both the grid resolution and neighbourhood extent influence terrain data extraction. This study has expressly adopted a similar approach to the work by Wise (2000) and Wise (2007), which is to conduct our spatial analyses using GUI-based methods that form the fabric of many generalisations and interpolation approaches that are readily available in many commercial geospatial software platforms (e.g. ArcMap, SAGA, GRASS, Quantum GIS). These methods may not be the most superior in terms of model robustness or their ability for further refinement or customisation in widely-used data analysis platforms such as R (R Core Team, 2013) (Mitášová & Hofierka, 1993; Heuvelink, 1999). However, they represent the *de facto* methods used to produce many DEM surfaces from a wide variety of data sources. Due to their ease of use and out-the-box application, they will most likely appeal to non-specialists for DEM creation for their specific applications.

2.2.3.1 Mean cell aggregation

The first interpolative approach is accessed from the Earth Science Research Institutes ArcGIS® (ESRI, 2021)(hereafter ArcMap) *Spatial Analyst's* generalisation toolbox. Mean cell aggregation (MA) techniques fall within the statistical category of generalisation approaches to determine each output cell's value. The tool works by aggregating the individual values of a group of cells, determined by a specified neighbourhood, and producing a single, coarser-resolution cell of that value (Gillin *et al.*, 2015). Output aggregated elevation values are then represented as summary values either as a minimum, maximum, mean, median, mode, or weighted average, calculated from a window of full-resolution elevation cells (Gesch *et al.*, 2012). Incidentally, the present study presents the generalised DEM surface based on the mean cell aggregated DEM values.

Importantly, there is no concept of a specified neighbourhood in the aggregate function. The neighbourhood and the output data frame are the same, are always rectangular, and encompass the same cell locations. The size of the block in the aggregate function is determined by the aggregation of cells necessary to reach the desired resolution defined by the end-user (ESRI, 2011). For instance, generalising the 1 m LiDAR DEM using the MA approach to create a 3 m DEM would require a neighbourhood size of 9 cells (centre cell and eight surrounding cells), and 25 neighbourhood cells would be aggregated to create a 5 m DEM (**Figure 2.3**).

Using the MA tool in ArcMap, I systematically interpolated the original 1 m LiDAR DEM dataset to lower resolutions with cell factors of 5, 10, 20, 30, 60 and 90 m, respectively (Gillin *et al.*, 2015). These resolutions were selected to thoroughly assess terrain feature representations across various solutions and evaluate how generalisation varies between LiDAR DEM attributes (5 to 20 m) and between satellite sensor platforms (30 to 90 m). The arithmetic mean of all pixels within each kernel is used as the new image pixel's value. By applying the MA approach, all pixels are equally weighted. Like all metrics using *mean* values, the output pixel values will be strongly influenced by outlier or extreme values belonging to the primary kernels (Przydatek *et al.*, 2003; Wagner, 2004; Li *et al.*, 2005).

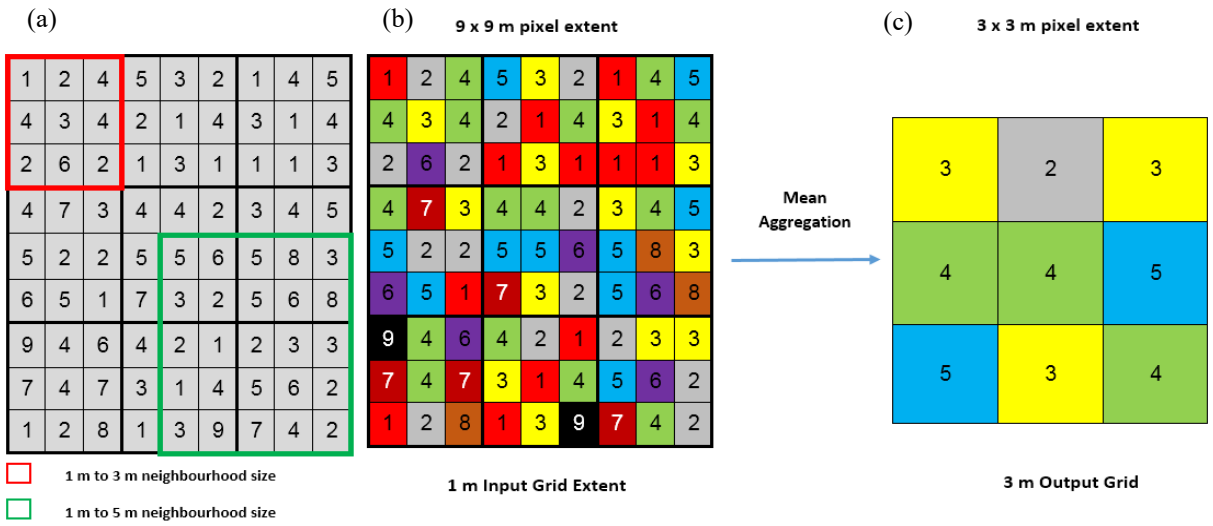


Figure 2.3: Example of applying a 3 x 3 mean aggregation approach to a 9x9 data frame (a). Note the aggregations of 9-pixel values, (b) to a single mean output pixel value (c). Modified from (ESRI, 2011).

2.2.3.2 Nearest neighbour resampling

The second generalisation approach used on the LiDAR DEM is the nearest neighbour (NNR) resampling. NNR resampling is perhaps the most commonly applied raster generalisation approach, and it functions by matching a pixel from the original image to its corresponding position in the resized image (Studley & Weber, 2011). If no corresponding pixel is available, the pixel nearest is used instead (**Figure 2.4**). NNR is particularly effective at representing discrete or categorical data as it does not alter the value of the output cells, i.e., the values that go into the grid stay precisely the same. A value of 2 comes out as a 2, and 11 will remain as 11 in the final output (ESRI, 2011). The input DEM is resampled using an output cell size (x-width, y-height), and for each new output cell, a value needs to be derived from the original input raster.

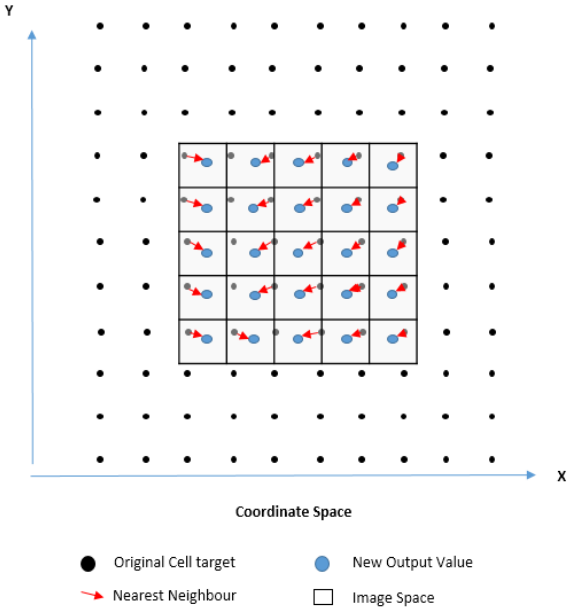


Figure 2.4: Example of applying the nearest neighbour approach to a 5x5 cell datasets. Note how the cell data frame remains constant, with only new cell values assigned to the grid data frame—modified from (ESRI, 2011).

With the NNR approach, the pixel cell centre from the input raster closest to the cell centre for the output processing pixel cell will be used. **Figure 2.4** outlines how the output value (blue dot) is derived by identifying the nearest cell centre (red arrow) on the input raster (black dot) and assigning itself to that value. This iterative process is repeated until all output cell values are given a new value at the desired output cell resolution (Studley and Weber, 2011). For this study, the original LiDAR DEM was resampled using X and Y coordinates of 5, 5; 10, 10; 20, 20; 30, 30; 60, 60 and 90, 90, respectively. NNR techniques can be used on continuous data, but the results can be blocky. So, this method is most suitable for reprojecting a raster surface (without a change in cell size) when preserving the original cell values for later quantitative analysis is essential. A point to note is that NNR may introduce noticeable error and other linear features where pixel realignment may be evident (eXtension, 2008). For that reason, NNR methods are generally considered the least accurate amongst interpolation approaches.

2.2.3.3 Top-to-Raster (Hydrological corrected surface)

This study's final raster generalisation method generates DEM surfaces that are hydrologically conditioned (HCD), used mainly for hydrological studies but commonly and generically used by many practitioners for non-hydrological analyses of soil-landscape properties or DSM. In selecting the HCD as a comparative interpolation method for terrain attribute extraction, I intend to showcase how the HCD has corrected for terrain drainage artefacts compared to other generalisation methods.

These comparisons are critical in regional DSM studies and provide the necessary quantitative evidence to prioritise the proper selection of the best-suited interpolation approach for the specific DSM objective. Additionally, practitioners who unknowingly use the HCD generalisations methods for a broad swath of terrain applications need to know how terrain variables vary between interpolation methods, derived explicitly from surfaces intended for modelling hydrological functions, across different resolutions.

The *topo-to-raster* tool in ArcMap's *3D Analyst* Toolbox is used to generate the HCD based on the ANUDEM program developed by Hutchinson (1988) and Hutchinson (2011). The HCD procedure is designed to take advantage of commonly available input elevation data, contours, and point heights, and the known characteristics of elevation surfaces. This method then uses an iterative finite-difference interpolation technique in a drainage enforcement process is to remove all sink points in the output DEM that have not been identified as sinks in the input sink coverage. The model assumes that all unidentified sinks are errors since sinks are generally rare in natural landscapes (ESRI, 2011). The HCD is optimised to have the computational efficiency of local interpolation methods, such as inverse distance weighted (IDW) interpolation, without losing the surface continuity of global interpolation methods, such as Kriging and Spline (Goodchild & Mark, 1987).

Briefly, the concept of the ANUDEM approach to HCD relies on the user determining the weight factor for all cells that flow into each downslope cell. Cells of undefined flow direction will only receive flow and will not contribute to any downstream flow. A cell is considered to have an undefined flow direction if its value in the flow direction raster is anything other than a series of predefined cell values of 1, 2, 4, 8, 16, 32, 64, or 128. Thus, the accumulated flow is based on the number of cells flowing into each cell in the output raster. Output cells with a high flow accumulation are concentrated flow areas and can identify stream channels. Output cells with a flow accumulation of zero are local topographic highs and can identify ridges. One of the primary considerations to deriving hydrologic characteristics of a surface is determining the direction of flow from every cell in the raster. Therefore, the HCD is derived by estimating the flow accumulation and direction (**Figure 2.5**) (ESRI, 2021).

The ANUDEM model can produce DEMs from digital contour, spot height, cliff, and streamline data and uses a drainage enforcement approach to ensure that catchment drainage integrity is accurately represented (Gesch & Larson, 1996). By applying the drainage enforcement algorithm with an iterative finite-difference interpolation approach, the model can remove spurious sinks from the DEM data, thus producing an elevation surface that more closely represents the actual surface with fewer elevation artefacts present (Gallant & Wilson, 1996). However, the model does not attempt to remove actual sinks in the terrain and is conservative when clearing spurious sinks. In other words, it does not clear spurious sinks that would contradict input elevation data by more than a specified value of tolerance (Hutchinson, 2011). However, Van Niekerk (2012) reported that DEMs produced by ANUDEM often have artefacts such as artificially deep river channels and gorges, particularly in areas of relatively moderate terrain that can have a significant impact on specific terrain applications. Known limitations include applying the HCD approach with densely sampled input elevation data such as LiDAR (ESRI, 2011). In these instances, the application of a surface interpolator may not be consistent with the input dataset. If the input elevation data contains sinks with more points than cells in the output raster, the HCD will fail. Hydrological modelling relies heavily on DEM data, and therefore, these models tend to have built-in hydrological correction factors able to improve the accuracy of selected topographic features (Li & Wong, 2010).

Nonetheless, a proper understanding of how the interpolator works is essential to prevent misapplication. To derive the HCD surface for this study, the LiDAR 1 m data was first converted to the contour intervals of 1, 5, 10, 20, 30, 60 and 90 m, respectively, using ArcMap's *Contour* tool from the *3D Analyst* Toolbox. The "raw" elevation contours were not smoothed or filtered and were used directly to create the HCD at the same output cell extent as the original contour data using the *topo-to-raster* tool with drainage enforcement selected and all other model parameters left as default to mimic the typical application by the ordinary, uniformed and non-GIS-savvy end-user.

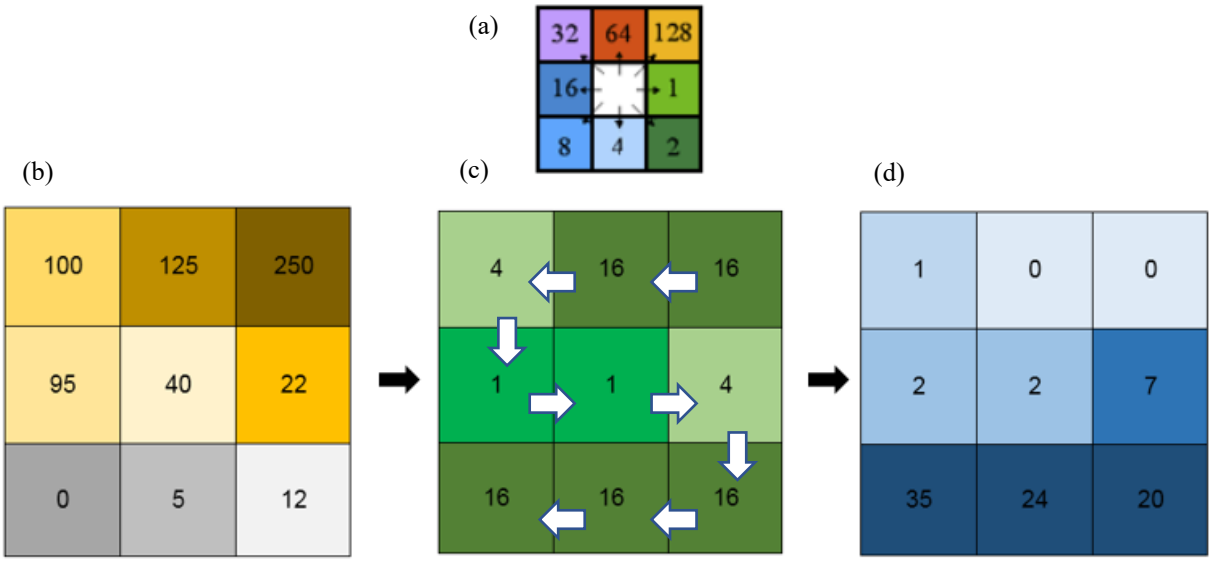


Figure 2.5: Calculation of ANUDEM using the commonly referred to eight-direction (D8) (Jenson & Domingue, 1988) flow model for flow accumulation and flow direction for deriving the HCD. Inset (a) direction coding, (b) elevation surface (m), (c) flow direction, (d) final flow accumulation. Modified from (ESRI, 2011).

2.2.3.4 Non-LiDAR DEM surfaces.

With the second goal of the study, I endeavour to investigate how the utility of using medium resolution satellite non-LiDAR derived (NLD) DEMs, like the SUDEM, SRTM and EarthEnv datasets, influence the extraction and representation of the same terrain variables: *Elev, SlopAsp, SlopGrad*, *ProCurv, PlanCurv, TWI, TST, LS, and TRI,* extracted for the LD products. Specifically, the LiDAR 5 m generalised products would be compared to the 5 m SUDEM while the LiDAR 30 m and 90 m products would be compared to the SRTM 1 and 3 arc-sec and EarthEnv90 datasets, respectively. It is necessary to evaluate the utility of different DEM sources in soil-landscape terrain extraction as DEM source, and production typically contains trades-offs between cost, accuracy, spatial coverage and pixel resolution (Robinson *et al.*, 2014), including the way the data are prepared, corrected and presented (Callow *et al.*, 2007).

LiDAR data is increasingly being used at an operational field level mainly since it offers users a far more accurate DTM and higher spatial resolutions than other DEM sources. However, the products often have limited spatial coverage with higher acquisition and processing costs (Sampson *et al.*, 2012; Baugh *et al.*, 2013) and may contain unnecessary artefacts or features that are not important for modelling soil-landscape relationships (Liu *et al.*, 2007). Nevertheless, coarser-resolution DEMs may still be a more cost-effective option for DSM to represent landscape form and process due to their broader synoptic coverage and free availability. Mashimbye *et al.* (2014) demonstrated that a DEMs resolution, source data and development approach significantly influenced land component delineation and associated applications. Li & Wong (2010) further state that there is enough research evidence to suggest that DEMS from different sources are more likely to produce inconsistent results than DEMs at varying resolutions but from the same source. I applied minimal pre and post-processing techniques to all NLD DEMs since the version and format of the NLD products used in this study had already undergone rigours processing before their release by the agencies responsible for their distribution.

2.2.4 Terrain analysis

All digital terrain and local morphometric data were extracted, analysed and processed using both ArcGIS® and the System for Automated Geoscientific Analyses (SAGA) (Conrad *et al.*, 2015) since both software programs provide all the necessary spatial tools for DEM and terrain property analysis for this study. The study focused on using the second-order finite difference algorithms of Zevenbergen & Thorne (1987) as well as Iwahashi & Pike (2007) as the basis to derive the following terrain indices for each of the 22 DEM datasets: *Elev, SlopGrad, SlopAsp, ProCurv, PlanCurv, TWI, LS, TST and TRI*. These DEM-derived topographic attributes were selected since they are considered the most indicative of the predominant morphological, ecological and hydrological drivers within the soil-landscape environment and considered necessary explanatory variables for predicting a range of soil-related properties (Florinsky, 2016). All derived DEM surfaces were compared based on vertical height and horizontal distance deviation from the 1 m LiDAR surface. Several studies address in detail the description and importance of these terrain attributes mentioned with regards to soil-landscape modelling, and the reader is referred to select critical studies, namely MacMillan *et al.* (2000); Thompson *et al.* (2001); Olaya (2009); Tesfa *et al.* (2009); Florinsky (2016); Maynard & Johnson (2014) and Gillin *et al.* (2015). A brief description of each investigated terrain variable evaluated is provided for this study (**Table 2.2**).

Notwithstanding, it is necessary to elaborate on the soil-terrain linkage of selected extracted variables briefly. In particular, *Elev* is considered the main driver for distinguishing how soil-terrain properties vary between upper and lower portions of the top sequence, i.e., summits and peaks from foot slopes and drainage zones. Similarly, the *Slope* is essential in identifying the spatial correlation between soil properties and steep backslope positions (Smith *et al.*, 2006). The second-order DEM derivatives such as *Curv* best explain the soil property relationships based on concavity or convexity in the landscape. Lastly, *TWI* provides a relative measure of hydrological conditions such as water availability and drainage at a given location in the landscape (Sörensen *et al.*, 2006).

2.2.5 Forestry soils database

Sappi Southern Africa (Sappi Forest) provided the Forest Industry Soils Database (FSD) co-operative gridded survey data for Braemar (**Figure 2.1**). The FSD served as the sampling locations for the terrain data extraction with a total of 1 168 points spatially analysed. Whilst antiquated, The FSD data remains the most extensive compilation of industry-standard soil surveyed data across selected forested areas in South Africa with detailed modal profile data collected at each site (Erasmus, 1998). The modal soil parameters, including *inter alia*, soil depth; wetness; texture; structure, colour, available soil water, and organic carbon, were not considered for this study but may require further analysis in evaluating potential soil-landscape trends in future research studies.

2.2.6 Statistical analysis

To evaluate the accuracy and quantify the level of agreement between the various scale-dependent DEM generalisation and terrain extraction outputs, I performed accuracy assessments to compare the derived values of the LD and NLD datasets (treatments) with the 1 m LiDAR surface (control). In all reviews, the original 1 m LiDAR DEM and all-terrain variables derived were considered to represent the 'actual' geomorphological surface from which the LD and NLD interpolated products were benchmarked. First, I displayed the descriptive statistics resulting from the DEM smoothing within the three generalisation approaches across the entire site. The minimum (*min*), maximum (*max*), mean (*mean*) and standard deviation (*SD*) statistics for each of the DEMs, including their respective DEM terrain derivatives, and presented in **Table 2.2**. The second assessment method involved the comparison of DEM components by evaluating several popular statistical measures: the Mean Absolute Deviation (MAD), Standard Error of Mean (SEM), and finally, the Nash-Sutcliffe efficiency coefficient (Nash & Sutcliffe, 1970). The study will also evaluate the Koppes formula (Maling, 2016), typically used for estimating RMSE vertical elevation error caused by covariation of horizontal error and slope, to instead determine the horizontal displacement (HD) using the vertical error of the predicted DEMs. The Koppes assessment results will highlight the often-overlooked extenuating effects of vertical displacement on horizontal variation, which in turn can have significant implications for sampling for soil-landscape examinations. For this study, I opted to use the MAD in favour of the RMSE as a measure of representing model prediction error and evaluating their performance (Aguilar *et al.*, 2005; Guo *et al.*, 2010). Despite being a relatively simple and universal model evaluation, metric concerns have been raised about the RMSE as an average model accuracy indicator. It has been suggested that the RMSE may be misleading in representing average error (Willmott & Matsuura, 2005) (**Equation 2.1**), particularly when evaluating DEMs and associated products.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(z_i^{predicted} - z_i^{actual}\right)^2}{n}} \tag{2.1}$$

The RMSE is perhaps best suited to data that conforms to a Gaussian or normal distribution of error where the standard deviation of the error is (or assumed to be) zero. Studies have shown that the mean error in surface analyses is, in fact, not equal to zero (Li, 1988; Monckton, 1994; Wechsler & Kroll, 2006) and that RMSE' suitability for representing dataset accuracy in terrain assessment is questionable. As an alternative, particular studies by Desmet (1997), Fisher & Tate (2006) and Thomas *et al.* (2014) have proposed the use of alternative, more complete representations of error. One such option is the use of the mean absolute error (**Equation 2.2**), which can be either negative or positive and thus can report systematic under or overestimation of DEM elevation error (bias):

$$\text{MAD} = \frac{\sum_{i=1}^{n}\left|z_i^{predicted} - z_i^{actual}\right|}{n} \tag{2.2}$$

Table 2.2: Summary description of terrain variables evaluated in this study.

Terrain Attribute	Unit	Description	Reference
Elevation	[m]	Elevation above sea level.	(Tesfa *et al.*, 2009)
Aspect	[deg]	Clockwise direction from North that a slope faces.	(Tesfa *et al.*, 2009)
Slope	[%]	The angle of inclination of the surface between the tangent and horizontal plans.	(Cavazzi *et al.*, 2013)
Profile Curvature	[m m^{-2}]	The curvature of the surface in the direction of maximum slope. Profile curvature affects the acceleration or deceleration of flow across the surface. Negative values indicate that the surface is upwardly convex at that cell. Positive values indicate that the surface is upwardly concave at that cell. A value of zero indicates that the surface is linear.	(Moore *et al.*, 1991)
Planform Curvature	[m m^{-2}]	Curvatures of the surface perpendicular to the direction of the maximum slope. Profile curvature relates to the convergence and divergence of flow across a surface. A positive value indicates that the surface is side-wardly convex at that cell, while negative values indicate that the surface is side-wardly concave. A value of zero indicates that the surface is linear.	(Moore *et al.*, 1991; Moore *et al.*, 1993)
Topographic Wetness Index	[index]	A steady-state wetness index as a function of slope and upstream contributing area per unit width orthogonal to the flow of direction. High values indicate large, gentle upslope areas with high water availability, while small index values indicate steep, small upslope areas with low water availability.	(Gessler *et al.*, 1995)
LS Factor	-	A dimensionless factor where the L-factor indicates the impact of slope length while the S-factor indicates slope steepness.	(Borrelli *et al.*, 2016)
Terrain Surface Texture	-	Emphasises the fine versus coarse expression of topographic spacing or grain. Each grid cell value represents the relative frequency of the number of pits and peaks within a radius of ten cells.	(Iwahashi & Pike, 2007)
Terrain Ruggedness Index	[index]	A quantitative measure of topographic heterogeneity by calculating the sum change in elevation between a grid cell and its eight neighbouring cells.	(Riley *et al.*, 1999)

The SEM (**Equation 2.3**) represents how well the predicted DEM values approximate the actual DEM values with smaller SEM values, indicating higher agreement with the actual DEM values.

$$SEM_{\bar{x}} = \frac{S}{\sqrt{n}} \tag{2.3}$$

The SEM should not be confused with standard deviation (*SD*), although the SEM can estimate *SD* (**Equation 2.4**). Unlike the *SD* that indicates the variability of responses around the mean, the SEM indicates the standard mean reliability.

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{2.4}$$

Likewise, the intra-correlation coefficient (ICC), or *p*, measures the homogeneity of observations within the DEM derived terrain classes relative to the variability of such observations between the DEM classes. The ICC describes how strongly units in the same group resemble each other is a number, usually found to have a value between 0 and 1.

In this study, the ICC refers to correlations within a class of data (for example, correlations within repeated measurements of similar DEM attributes, i.e., *Elev vs Elev*), rather than correlations between two different categories of data (for example, the correlation between *Elev* and *SlopGrad)* (Liljequist *et al.*, 2019). The ICC is an incredibly useful index for indicating the ability of an experimental method to detect and measure systematic differences between subjects. The ICC serves as a quantitative estimate of this aspect of reliability. The ICC is generally calculated as a ratio ICC (*p*)= between cluster variability / (within-cluster variability + between cluster variability). Using scientific notation, the ICC is presented in **Equation 2.5**. If the between variance is equal to or larger than the within variance of interest, the method's reliability is low. Therefore, the ICC will be zero only when the estimated effect of a random factor is zero and will reach unity only when the estimated effect of the error is zero, given that the total variation of the observations is more significant than zero (Hays, 1988).

$$\text{ICC}\,(p) = \frac{S_{between\ cluster\ variance}}{S_{between\ cluster\ variance} + S_{within\ cluster\ variance}} \tag{2.5}$$

The previous assessment approaches mentioned effectively quantify the DEM variation at a site level with all 1 168 (*n)* sample data points contributing to the overall accuracy assessment, i.e., there is no anisotropic structure considered with the "landscape" evaluation of DEM variability. To further test the functionality of the derived products, I assessed the robustness of the model outputs at a "localised" scale by segmenting all the data products into six-equal interval elevation class-bins with DEM and terrain attribute accuracy assessment then further defined for each of the class intervals: 230-350 m, 351-450 m, 451-550 m, 551-650 m, 651-750 m, 751-800 m (**Table 2.4**). The decision to segment the DEM products by elevations intervals is intended to test the following statistical hypothesis: *There is no significant difference between the actual surface model and derived terrain products at a localised scale; that is, the predicted products closely match the actual surface* (H_0:$\mu_1=\mu_2..,$). The null hypothesis is then expected to describe the spatial similarity or agreement between the 1m LiDAR and the derived products across the various resolutions. The alternative hypothesis would evaluate the following: *There is a significant difference between the actual surface model and derived terrain products at a localised scale* (H_1:$\mu_1\neq\mu_2..,$). The null hypothesis is evaluated using the Analysis of Variance (ANOVA) by assessing the *p*-values, *F* values and *F* critical values between the 1 m LiDAR and the generalised surface products (treatments). The ANOVA results will provide a basis for either accepting or rejecting H_0 in favour of H_1; however, the ANOVA will not describe the multiclass variability between products. A pair-wise comparison using the Tukey-Kramer (Kafdar, 1994) posthoc test is applied to the localised data to statistically determine which products differ significantly from the 1 m LiDAR data. The central premise of the Tukey-Kramer test (also referred to as the honestly significant difference, HSD) is to compute a new critical value that can be used to evaluate whether differences between any two pairs of means are statistically significant using a studentised range defined by a *q* distribution (Williams & Abdi, 2010). The *q* statistic gives the exact sampling distribution of the largest difference between a set of mean values originating from the same population.

All pairwise differences are evaluated using the same sampling distribution used for the largest difference. The critical value, therefore, represents the *mean* difference that has to be exceeded to achieve significance. Once the critical value for each treatment is calculated, the difference between all possible pairs of *mean* values is also defined. Each difference is then compared to the HSD critical value. If the difference is larger than the HSD value, the comparison is significant. Several posthoc tests are available for pairwise comparisons of population means under the same conditions and include the Dunnett's test or Bonferroni test (Motulsky & Searle, 1998); however, the Tukey-Kramer HSD test is perhaps the most robust in its application. It provides a good combination of calculation simplicity and prediction accuracy. For further detail on the mathematical derivation and application of the Tukey-Kramer posthoc test, the reader is referred to Somerville, 1993.

The Nash-Sutcliffe efficiency (NSE) coefficient (**Equation 2.6**) is also used as a measure of statistical association, or goodness of fit, between the LD products and the original 1 m LiDAR surface, indicating the percentage of the observed variance that is explained by the derived, lower resolution and generalised models. The NSE has found popularity in environmental and atmospheric based research (Lin *et al.*, 2017). Reasonable NSE values have been categorised as NSE > 0.65 excellent, 0.5 to 0.65 very good, 0.2 to 0.5 as good, and <0.2 as poor (Allen *et al.*, 2007). Values of NSE range from 1.0 (perfect fit) to - ∞ with values below zero indicating that the mean of the observed data (actual) would have been a better predictor than the model (predicted) (Krause *et al.*, 2005). The NSE results should not be confused as a redundant measure of MAD. The NSE has only been applied to measure the degree of DEM (elevation) (dis)association between all observed surfaces since elevation quality and accuracy is the primary variable of concern in most DEM choices. Similarly, MAD has been applied to all derived terrain attributes across all resolutions to measure standard comparison of derived features for the different DEM surfaces.

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Z_i^{actual} - Z_i^{predicted}\right)^2}{\sum_{i=1}^{n}\left(Z_i^{actual} - Z_i^{mean}\right)^2} \quad (2.6)$$

The final data assessment includes a series of visual interpretations (maps) evaluating how well the predicted DEM products, compared to original LiDAR DEM and associated terrain derivatives for each sensor type and resolution. This approach worked incredibly well to visualise and identify errors related to the simulation of directional or geometric properties such as Slope direction or *SlopAsp* for the various DEM datasets. *SlopAsp* is a circular variable: its values range from 0^0 to 360^0, and both values correspond to a north direction. Thus, *SlopAsp* cannot readily be used in linear statistical analysis to examine and compare independent datasets statistically. Computationally, the difference between 1 and 360 degrees is the same as the difference between 1 degree and 2 degrees, making it problematic to plug into a typical statistical model (Jenness, 2007). To analyse the aspect data across all the DEM surfaces, the circular (radian) values were transformed into continuous scalar sine (easterness) and cosine (northerness) values, decomposing them into north-south (x) and east-west (y) components. By doing so, sine values of 1 can be interpreted as the slope being entirely east-facing while a sine of −1 denotes that the slope is entirely west facing. Cosine of 0 implies that the dominant slope is facing either north or south. Equally, a cosine value of 1 indicates a wholly north-facing slope with a cosine value of −1 outlining a south-facing slope and a sine of 0 indicating that the slope is facing either east or west (Penížek & Borůvka, 2006; Piikki *et al.*, 2013). Jenness (2007) further cautions that the decomposition of the *SlopAsp* values to scalar variables does not maintain a constant interval between the transformed units. That is, the magnitude of the values change by a variable amount depending on the direction, such that a change in sine values corresponding to a change of 1 degree = 0.00015 when going from 90 to 91 degrees, with increases of more than two orders of magnitude to 0.017 when moving from 180 to 181 degrees. This may only be relevant in the statistical analysis if the approach assumes that the data are at an interval-level. For this study, the data were represented at a continuous interval, so the applied transformations were applicable. *SlopAsp* results were then represented in a pie chart-type image where the "slices" are shaded according to the direction of the aspect (Jenness, 2007).

Maps were generated for selected terrain products and included representations of calculated statistics such as the *Elev* bias, *Elev* difference, *SlopAsp* difference, *ProfCurv* and *PlanCurv* variability and relative SlopGrad variability. Maps were not generated for the entire 22 datasets, and in most cases, were only generated for visual comparison of the NLD surface results to the 1 m LD. However, the summary statistics table (**Table 2.3**) compares all surface derived products comprehensively. DEM surfaces' horizontal consistency is fundamental to the comparative analyses of their operational applicability to soil-landscape research. Nevertheless, the horizontal component's positional errors are generally larger in magnitude than the vertical element for most sensor platforms (Mora, 2015). Hodgson *et al.* (2005) point out that the elevation error will typically increase as the surface's slope increases for a constant error in the horizontal location of observation. For example, a continuous surface *SlopGrad* of 40° and a 200 cm horizontal elevation displacement may result in a vertical elevation error of up to 72 cm. Earlier accuracy assessment studies by METI on GDEM1 reported horizontal displacement errors of 0.97 pixels (~ 30 m). This is not surprising because the horizontal uncertainty also contributes additional elevation uncertainty in sloped terrain according to Koppe's relationship (**Equation 2.8**). Essentially, Koppe's formula is based on the tangent of the *mean* surface slope (α) and observed HD defined as:

$$\text{Elevation Error} = \text{Tan}(\alpha) \times \text{Horizontal Displacement} \quad (2.7)$$

A final exploratory analysis of the DEM derived products included assessing horizontal displacement by modifying Koppes formulae for conventionally used for elevation error to calculate the absolute horizontal error instead (**Equation 2.7**).

$$\text{Horizontal Displacement} = \frac{ABS(Elevation\ Error)}{Tan\ (\alpha)} \quad (2.8)$$

It must be noted that the present study did not set out to empirically quantify the DEM horizontal displacement, but rather DEM elevation error. Comparisons of vertical error *vs* horizontal error *vs* DEM source were restricted to descriptive and graphical comparisons for the three generalisation approaches (**Figure 2.14**).

2.3 RESULTS AND DISCUSSION

2.3.1 Comparison of LiDAR reference surface with LD and NLD elevation products

Comparisons were performed for the twenty-two resulting DEM surfaces derived from each of the sensor platforms – LiDAR, SUDEM, SRTM and EarthEnv – with the primary aim of comparing the variation in DEM outputs and associated descriptive metrics associated with applying the three generalisation approaches. The results are presented in a summary table (**Table 2.3**) and further independently discussed in detail for each of the seven extracted terrain variables due to the extensive statistical and spatial analyses of the DEM products.

The first set of coupled results include the outcomes of the correlation analysis, conducted in SPSS ® version 13.1, to determine the strength of the relationship between the *mean* elevation values of the original 1 m LiDAR DEM (actual) and the resampled (predicted) LD DEMs (5 m to 90 m) and SRTM, SUDEM and EarthEnv DEMs. By comparing the statistical relationship(s) using all 1 168 sample points, a strong positive linear relationship for vertical accuracy between the 1 m LiDAR surface and sensor platform was observed in the correlogram for the full range of DEM resolutions (**Figure 2.6**). The *prima facie* results suggest that all derived DEM surfaces accurately represent the average terrain elevation of the surface model. These results are further highlighted in **Table 2.3,** where the mean absolute elevation deviation from the 1 m LiDAR is between 0.30 m and 12.0 m for selected DEM surfaces. Generally, the derived DEMs produced a near-perfect goodness-of-fit, measured with the coefficient of determination (R^2) values ranging between 0.98 for the least best performing DEM and 0.99 for the best predicted DEM surfaces. This study's R^2 values are similar to those reported by Shafique *et al.* (2011), comparing similar sensor DEM products across the same resampled pixel resolutions. Correlation coefficients, however, merely demonstrate the robustness of the relationship between the actual and predicted surfaces.

Therefore, these results only suggest that all predicted DEM elevation values vary in the same direction, at the same point in the feature space (the entire study site) and similar magnitude as the actual surface, in this case, the 1 m LiDAR DEM. This highlights the limitation of comparing DEM surfaces purely from a "global" or macro-landscape perspective since much of the elevation variation is masked when evaluated with standard descriptive analysis measures. A more considerate quantitative descriptive statistic worth evaluating is the Nash-Sutcliffe Efficiency (NSE) index. To reiterate, the NSE is a normalised statistic that determines the relative magnitude of residual variance ("noise") compared to measured data variance ("information") with values generally ranging between 0.0 and 1.0 indicating levels of performance (Moriasi *et al.*, 2007). Whereas the R^2 values measure the goodness of fit of a statistical model, the NSE is s used to quantify how well a model simulation can predict the outcome variable. The NSE values for all the LD and NLD DEM surfaces ranged from 0.94 to 0.99, indicating that all simulated DEM surfaces are suitable predictors for the observed 1 m LiDAR surface, suggesting optimal altitude prediction performance across all 22 DEM surface models. However, it must be noted that the *mean* elevation values presented in **Table 2.3-2.4** are merely descriptive summaries for each DEM model. As Gonga-Saholiariliva *et al.* (2011) point out, these simple global metrics may conceal substantial local variability between data sets; and alternative methods may be necessary to provide further insight into the variability between surface products.

For this reason, the MAD, virtually a measure of systematic bias from the actual surface, provides a better representation of the variability between the different DEM scale-dependent products. The MAD values presented in **Table 2.3-2.4,** along with boxplots (**Figure 2.7a-f),** illustrate the distribution of deviation (error) values in DEM elevation among the 22 DEM grids. It is important to note that the box and whisker plot

distributions observed in **Figure 2.7a-f** represent the uncategorised elevation values using all the sample points across the entire study site. A comparative analysis of the *interquartile range, mean, median, minimum and maximum* elevation values of the data set for the predicted LD surfaces derived from the LiDAR DEM and the NLD surfaces products show minor deviations from the reference 1 m LiDAR surface. While the boxplots help display the distribution of the data points around the mean, they are not particularly good at indicating whether the mean values are significantly different from one another. The standard deviation of DEM elevation heights represents a more measurable metric of spread about the mean.

Observing the results for DEM and sensor resolution in the correlogram (**Figure 2.6**) and the box plots (**Figure 2.7a**), it is clear that elevation distributions across pixel resolutions yielded high correlations with the reference 1 m LiDAR. Grohmann & Sawakuchi (2013) found similar results and point out that these results may seem somewhat counterintuitive, as one would expect the elevation to vary significantly with an increase in cell sizes. However, the correlogram results show a slightly better agreement between the high-resolution derived surfaces (5 m-20 m) and the reference surface than the lower resolution products. Gonga-Saholiariliva *et al.* (2011) explain these elevation trends by suggesting that while there is a spatial error present in the various derived DEMs, including outliers, neither data scatter in the linear regression models nor variance in the vertical accuracy constitute sufficient criteria to achieve a full definition of DEM error because the error is expressed globally, i.e., the DEM error is masked when evaluated across the entire study site. A proposed theory on these findings is that due to a lack of spatial association tested within the DEM surfaces, the application of linear regression models may be susceptible to overestimating both the strength of the correlation and precision. Importantly, our findings agree with studies conducted by Hoffman & Krotkov (1990), Barber & Shortridge (2005), Gonga-Saholiariliva *et al.* (2011) and Shi *et al.* (2012). While the variability between the LD DEM surfaces is generally uniform, minor spatial residual trends are noticeable but not considerable between the different generalisation approaches and pixel resolutions. As anticipated, the general trend exhibited an underestimation of the predicted reference surface by the LD derived elevation surfaces as the resolution decreased from 60 m to 90 m regardless of the generalisation method.

Notwithstanding of the three generalisation approaches, the HCD performed the least consistent in predicting the actual surface with MAD vertical error predictions manifesting from as low as 10 m DEM resolution and ranging between 0.4 m to 24 m across all derived pixel-resolutions. **Figure 2.8** provides a useful snapshot from a site perspective of how each generalisation approach can represent the reference surface features. The three main observations here are that firstly all surfaces, regardless of generalisation, between 5 m-20 m, can still represent the primary elevation features present in the reference surface despite the apparent decrease in pixel resolution. Second, even as resolution devolves from 60 m to 90 m in the NNR and MA approach, the general surface features are still aligned to the primary reference surface. Moreover, both the MA and NNR LD surfaces performed considerably better than the NLD surfaces. This result may prove crucial for certain soil-landscape studies given the acquisition costs, processing demands and limited footprint of high-resolution LiDAR products.

Table 2.3: Summary of 1 m LiDAR surface and extracted topographic indices for each generalisation method and pixel resolution including results for non-LiDAR DEMs.

		DEM										SLOPE						LS						ROUGHNESS					
Data Source	*Resolution*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*	*SEM*	*ICC*	*HD*	*NSEG*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*
LiDAR 1 m		576.0	237.0	784.0	121.0							26.0	1.0	136.0	17.0			3.1	0.0	11.5	2.3			0.3	0.0	0.8	0.2		
Aggregated																													
	5	576.1	237.8	784.5	120.9	1.0	0.3	4.2	1.0	0.3	1.0	24.4	1.5	101.4	13.7	0.8	5.8	3.2	0.0	12.7	2.1	0.8	0.8	0.2	0.0	0.7	0.1	0.6	0.2
	10	576.1	238.6	784.7	120.9	1.0	0.6	4.2	1.0	0.3	1.0	24.0	1.2	73.4	12.9	0.8	6.4	3.2	0.0	10.3	2.0	0.8	0.9	0.3	0.0	0.6	0.1	0.1	0.2
	20	576.0	238.8	785.3	121.0	1.0	1.2	4.2	1.0	0.4	1.0	23.3	1.6	59.1	12.1	0.7	7.3	3.2	0.1	8.8	1.9	0.8	0.9	0.4	0.1	0.6	0.1	-0.1	0.2
	30	576.1	241.5	784.8	120.9	1.0	1.8	4.2	1.0	4.1	1.0	22.4	1.7	56.4	11.3	0.7	8.0	3.1	0.1	8.8	1.9	0.7	1.0	0.5	0.2	0.7	0.1	-0.1	0.2
	60	575.8	240.4	783.6	120.9	1.0	3.3	4.2	1.0	4.5	1.0	19.5	1.1	51.8	9.3	0.6	10.3	2.9	0.1	11.9	1.7	0.6	1.3	0.6	0.3	0.8	0.1	-0.1	0.3
	90	575.5	245.7	785.6	120.9	1.0	5.1	4.2	1.0	3.0	1.0	16.7	0.5	43.9	7.8	0.5	12.3	2.5	0.0	8.9	1.5	0.5	1.5	0.6	0.3	0.9	0.1	-0.1	0.3
Hydro																													
	5	576.1	239.1	784.5	120.9	1.0	0.4	4.2	1.0	0.2	1.0	24.1	0.0	66.8	13.0	0.8	6.7	3.1	0.0	11.7	2.0	0.8	0.9	0.2	0.0	0.4	0.1	0.4	0.2
	10	576.0	239.5	784.4	120.8	1.0	1.0	4.2	1.0	0.4	1.0	23.2	0.6	58.1	12.6	0.7	7.4	3.1	0.0	8.2	2.0	0.8	1.0	0.2	0.0	0.4	0.1	0.1	0.2
	20	575.6	240.0	784.3	120.7	1.0	2.7	4.2	1.0	0.5	1.0	20.3	0.1	53.4	11.6	0.6	10.1	2.8	0.0	8.7	1.9	0.6	1.3	0.4	0.1	0.5	0.1	-0.1	0.2
	30	574.5	238.9	781.4	120.7	1.0	5.0	4.2	1.0	3.8	1.0	16.6	0.0	48.7	11.1	0.4	13.1	7.3	0.0	57.3	9.5	0.0	7.4	0.4	0.1	0.7	0.1	-0.1	0.2
	60	571.8	241.0	782.4	120.3	1.0	13.0	4.2	1.0	64.9	1.0	9.2	0.0	34.6	8.8	0.2	18.8	1.4	0.0	7.9	1.6	0.2	2.3	0.5	0.2	0.8	0.1	-0.1	0.2
	90	566.9	263.3	726.5	114.9	1.0	24.0	4.0	1.0	10.9	0.9	5.1	0.0	27.2	6.4	0.1	21.7	0.8	0.0	8.5	1.1	0.1	2.6	0.4	0.1	0.7	0.1	-0.1	0.2
Resample																													
	5	576.1	237.1	784.2	120.9	1.0	0.3	4.2	1.0	0.3	1.0	24.4	1.1	93.6	13.7	0.8	5.6	3.2	0.0	12.6	2.1	0.8	0.8	0.3	0.0	0.7	0.1	0.6	0.1
	10	576.1	237.8	783.8	120.9	1.0	0.6	4.2	1.0	0.3	1.0	24.0	1.5	80.3	13.1	0.8	6.5	3.2	0.1	12.5	2.1	0.8	0.9	0.3	0.0	0.6	0.1	0.2	0.2
	20	576.0	238.0	783.6	120.9	1.0	1.3	4.2	1.0	5.0	1.0	9.2	0.7	22.8	4.2	0.5	17.3	3.2	0.1	9.1	2.0	0.8	0.9	0.4	0.2	0.7	0.1	-0.1	0.2
	30	576.2	241.4	785.1	120.9	1.0	1.9	4.2	1.0	1.9	1.0	22.8	1.0	59.2	11.6	0.7	7.9	3.2	0.0	9.1	1.9	0.7	1.0	0.5	0.2	0.7	0.1	-0.1	0.2
	60	576.0	237.7	783.7	120.8	1.0	3.5	4.2	1.0	0.0	1.0	20.4	1.1	53.4	9.9	0.6	10.0	3.0	0.0	13.3	1.8	0.6	1.3	0.6	0.2	0.8	0.1	-0.1	0.3
	90	575.8	242.3	785.9	120.8	1.0	5.3	4.2	0.9	3.6	1.0	17.9	0.8	49.6	8.5	0.4	12.0	2.7	0.0	14.3	1.7	0.5	1.5	0.6	0.3	0.9	0.1	-0.1	0.3
SUDEM	*5*	583.9	248.2	788.7	119.9	1.0	8.0	4.2	1.0	33.4	1.0	18.6	1.1	64.2	10.0	0.5	11.1	2.5	0.0	10.8	1.6	0.5	1.4	0.2	0.1	0.4	0.1	0.2	0.2
SRTM	*30*	588.1	253.0	794.0	118.6	1.0	12.0	4.2	1.0	3.1	1.0	20.7	0.0	62.1	11.0	0.4	11.4	2.9	0.0	9.5	1.9	0.5	1.6	0.7	0.5	0.8	0.1	0.1	0.3
SRTM	*90*	586.9	260.0	793.0	118.3	1.0	13.0	4.1	0.9	34.2	1.0	16.1	0.6	49.0	8.5	0.3	13.6	2.3	0.0	8.0	1.5	0.4	1.7	0.7	0.4	0.9	0.1	0.0	0.4
EarthEnv	*90*	586.1	262.0	789.0	118.4	1.0	12.0	4.2	0.9	6.1	1.0	14.6	0.6	52.2	7.7	0.2	14.6	2.1	0.0	9.2	1.4	0.3	1.8	0.7	0.4	0.9	0.1	0.0	0.4

Table 2.4: (cont.) summary of extracted topographic indices for each generalisation method and pixel resolution/ filter combinations.

		TRI						TWI						PLANCURV					PROCURV				
Data Source	*Resolution*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*	*Mean*	*Min*	*Max*	*SD*	*Corr*	*MAD*	*Mean*	*Min*	*Max*	*SD*	*MAD*	*Mean*	*Min*	*Max*	*SD*	*MAD*
Lidar 1 m		0.2	0.0	1.0	0.1			4.8	0.3	16.1	1.7			0.4	-315.0	561.6	53.7		0.3	-3.1	5.6	0.4	
Aggregated																							
	5	0.9	0.1	3.7	0.5	0.8	0.7	5.4	2.3	17.1	1.4	0.5	1.0	0.4	-74.7	75.3	7.8	27.4	0.1	-6.7	7.5	1.4	19.7
	10	1.7	0.2	5.3	0.9	0.8	1.5	5.5	3.2	16.6	1.4	0.3	1.2	0.4	-24.9	55.7	4.3	27.3	0.1	-3.6	3.0	0.6	19.7
	20	3.4	0.3	8.4	1.7	0.7	3.2	5.8	3.7	15.5	1.5	0.3	1.3	0.3	-33.0	16.3	2.7	27.2	0.1	-1.4	1.9	0.4	19.7
	30	5.0	0.5	12.0	2.4	0.7	4.8	5.9	3.9	15.2	1.4	0.3	1.4	0.5	-18.0	40.9	2.9	27.3	0.1	-1.2	1.3	0.3	19.7
	60	9.4	1.0	23.2	3.8	0.6	9.2	6.4	4.6	15.1	1.5	0.2	1.8	0.3	-11.4	16.7	2.0	27.2	0.0	-0.8	0.8	0.2	19.7
	90	12.7	1.8	28.3	4.7	0.5	12.6	6.8	0.0	15.2	1.5	0.1	2.3	0.3	-7.1	13.2	1.7	27.1	0.0	-0.5	0.5	0.2	19.7
Hydro																							
	5	0.9	0.0	2.4	0.5	0.8	0.7	5.4	3.0	15.7	1.3	0.3	1.1	0.8	-70.3	327.9	12.8	27.9	0.1	-4.0	2.7	0.6	19.8
	10	1.7	0.0	4.1	0.9	0.7	1.5	5.6	3.5	14.9	1.5	0.4	1.2	0.2	-39.1	14.6	3.6	27.4	0.0	-2.3	1.3	0.4	19.8
	20	2.9	0.1	7.6	1.6	0.6	2.7	6.1	3.9	16.3	1.6	0.2	1.6	0.4	-20.8	35.3	2.9	27.3	0.0	-1.1	0.9	0.3	19.7
	30	3.7	0.0	10.3	2.3	0.4	3.5	7.8	-1.5	18.4	3.9	0.0	4.2	0.3	-13.6	12.5	1.5	27.2	0.0	-0.3	0.4	0.1	19.7
	60	4.1	0.0	14.7	3.8	0.2	3.9	8.4	4.7	19.6	2.3	0.0	3.8	0.2	-8.9	42.0	2.0	27.1	0.0	-0.3	0.3	0.1	19.7
	90	3.4	0.0	17.4	4.1	0.1	3.3	9.9	6.0	18.1	2.2	0.1	5.1	-1.4	-1174.2	6.1	41.2	28.6	0.0	-0.1	0.1	0.0	19.7
Resample																							
	5	0.9	0.1	3.4	0.5	0.8	0.7	5.4	1.4	17.9	1.5	0.5	1.0	0.0	-200.3	137.3	14.6	27.6	0.1	-10.8	13.6	2.0	19.7
	10	1.8	0.1	5.8	0.9	0.8	1.6	5.5	3.4	17.6	1.5	0.4	1.1	-0.1	-157.6	51.4	8.1	27.4	0.0	-4.8	3.5	0.8	19.7
	20	3.5	0.4	8.8	1.8	0.7	3.3	5.7	3.7	15.5	1.5	0.3	1.3	-0.3	-13.1	13.6	1.7	27.3	0.0	-0.5	0.4	0.1	19.7
	30	5.2	0.5	12.6	2.5	0.7	5.0	5.8	3.7	15.3	1.4	0.3	1.4	0.6	-27.0	80.1	4.3	27.4	0.1	-1.5	1.3	0.3	19.7
	60	10.0	0.9	24.6	4.2	0.6	9.8	6.4	4.4	15.3	1.5	0.3	1.8	0.4	-15.8	31.4	2.4	27.3	0.0	-0.9	1.1	0.2	19.7
	90	14.1	2.0	32.3	5.5	0.5	13.9	6.7	0.0	15.5	1.7	0.1	2.2	0.3	-12.9	12.8	1.9	27.2	0.0	-0.6	0.8	0.2	19.7
SUDEM	*5*	0.7	0.1	2.5	0.4	0.5	0.5	6.0	2.9	17.9	1.4	0.2	1.6	0.4	-31.2	48.4	5.4	27.2	0.1	-6.9	7.3	1.2	19.8
SRTM	*30*	4.5	0.0	13.0	2.2	0.5	4.3	6.1	4.0	17.0	1.6	0.1	1.8	0.5	-11.1	34.4	2.9	27.3	0.1	-1.5	1.6	0.4	19.7
SRTM	*90*	11.9	1.3	33.2	5.0	0.4	11.7	7.0	4.9	15.7	1.7	0.1	2.4	0.4	-15.7	36.9	2.8	27.3	0.0	-0.5	0.5	0.2	19.7
EarthEnv	*90*	10.1	1.5	34.0	4.6	0.3	9.9	7.2	5.1	16.8	1.6	0.0	2.6	0.3	-7.7	18.4	1.5	27.2	0.0	-0.5	0.5	0.1	19.7

SD=Standard Deviation, Corr=Correlation, MAD=Mean Absolute Deviation, SEM=Standard Error of Mean, ICC=Inter Correlation Coefficient, HD=Horizontal Deviation, NSEG= Nash-Sutcliffe Efficiency Criteria

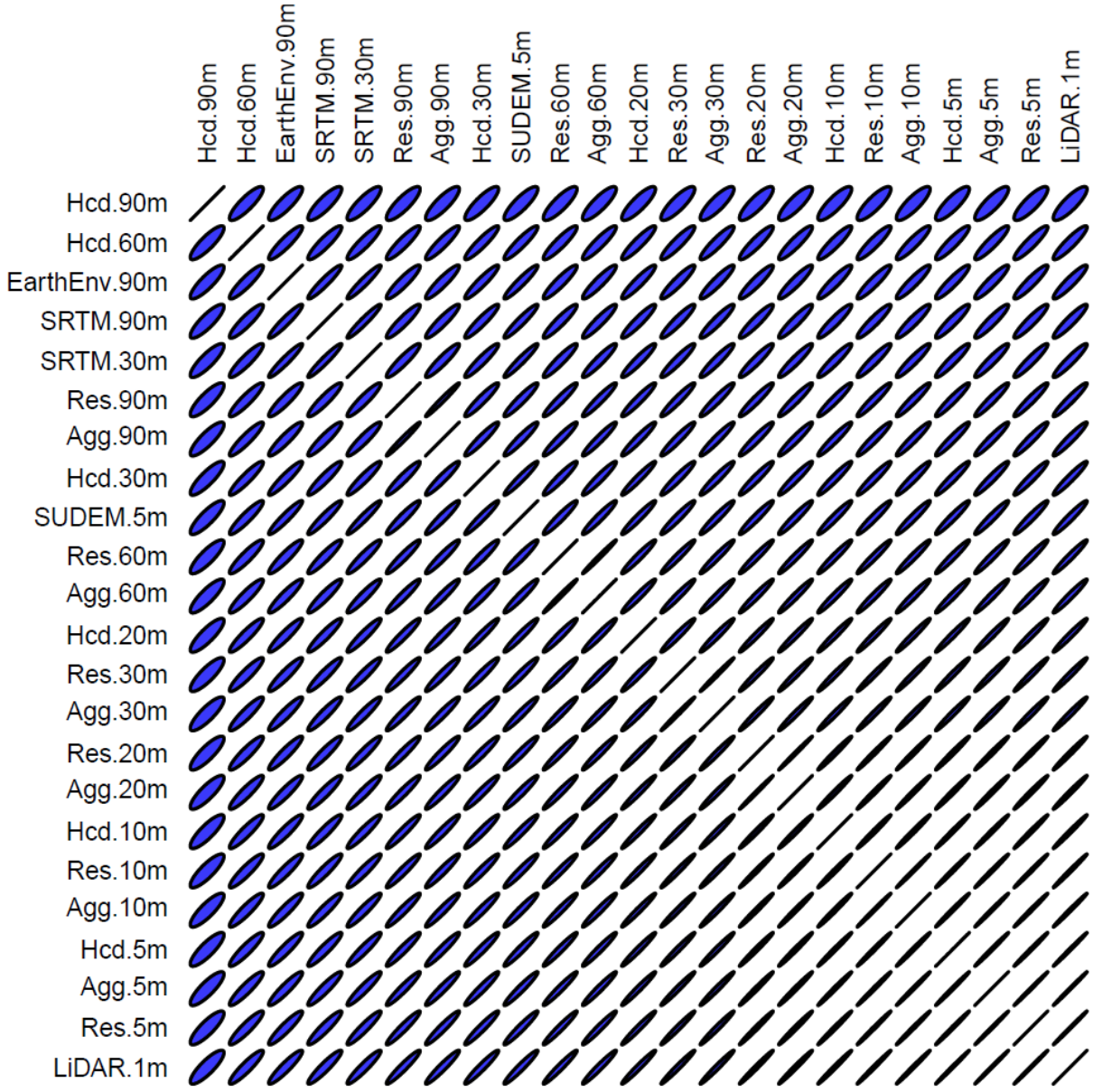


Figure 2.6: Correlation summary graph for the 22 DEM surface datasets. The ellipse's shape represents the relative correlation between DEM surfaces with high correlations described as a diagonal straight line. Note the high level of positive agreement between the derived surfaces and actual 1 m LiDAR surface model—note *HCD: Hydrologically correct DEM, Agg: Aggretagted DEM, Res: Resampled DEM.*

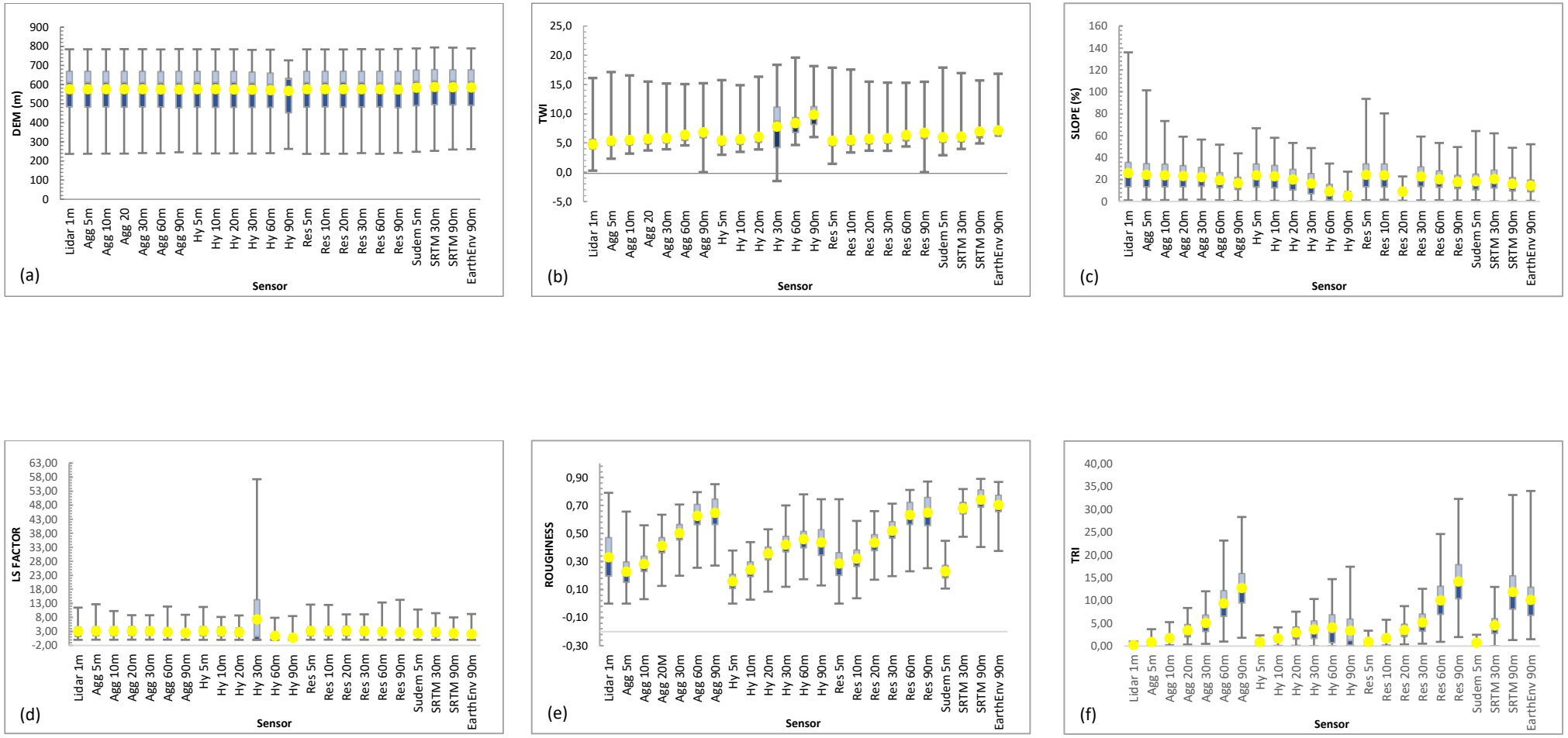


Figure 2.7: Box and whisker plots showing the distribution of six topographic matrices for all 22 datasets derived from LiDAR and non-LiDAR platforms. The box's lower boundary indicates the 25th percentile while the box's upper boundary indicates the 75th percentile. The lower and upper whiskers indicate the 5th and the 95th percentile. The black horizontal lines within the box indicate the median and the yellow dot the mean values for each variable.

Therefore, the ability to utilise LiDAR-derived DEM surfaces at lower resolutions that retain a high accuracy level to the reference surface significantly benefits data reduction and processing operations. However, despite the coarse resolution DEM surfaces capturing the broad terrain-feature outlines useful for visual representations, there is a steady loss of clarity with the coarse resolution surfaces. These products may offer limited functionality for selected DSM and soil-landscape analyses, given that specific geomorphology processes that drive pedological functions are not always detected at coarser spatial resolutions (Thomas *et al.*, 2017). Still, the glaring observation is perhaps the HCD 60 m and 90 m DEMs results, which show a complete breakdown of surface features from the reference surface. It must be noted that these variations are since the HCD model is intended to condition the DEM surface for hydrological purposes. So deviations are expected, but much of the loss of terrain information is not always appreciated by the end-user and can significantly impact terrain covariate predictions. These results are well aligned to similar studies by Chen *et al.* (2012), who reported that the ANUDEM method's surfaces always had the most significant RMSE values for the generalised DEMs generalised from a 3 m DEM. An interesting finding is that unlike the LD DEM surfaces that consistently under predicted the reference surface elevation, the NLD DEMs, in contrast, are more prone to over predict the reference 1 m surface model.

Given that the four NLD DEMs all contain, in varying degrees, some influence of the SRTM dataset(s) as well as the ASTER GDEM2 in the case of the EarthEnv, these findings are in line with other similar studies which showed that SRTM and ASTER derived surface products tend to over predict surface elevation values. It has been suggested that these phenomena may be related to the different data acquisition or processing methods applied to both the SRTM and ASTER datasets (Jarihani *et al.*, 2015). Another possible explanation for these results is that unlike the LD derived products representing an accurate surface model, the NLD DEM products detect reflective surfaces and represent the height of imaged surfaces, not the bare surface height. Therefore, if obtrusive surface features dominate an area, the above-ground features will likely contribute to a positive elevation bias (Gesch *et al.*, 2012). This is somewhat expected given that the study site is dominated by a variety of mixed-age commercial tree species. With that being said, the vertical error related to the SRTM DEM surfaces reported in this study are still well within the performance requirements for the absolute and relative vertical errors of 16 and 10 m, respectively and are typically associated with the SRTM datasets (Rodriguez *et al.*, 2005).

The results in **Figure 2.9** provide a different perspective of the variation in elevation surface, specifically between the satellite-derived DEMs and the 1 m LiDAR surface, by applying a DEM-differencing approach using ArcMap's *Raster Calculator*. The difference, or visual MAD, between the NLD DEMs and the reference surface highlight regions of over and under-estimation of elevation (**Figure 2.9a to Figure 2.9d**) and slope (**Figure 2.9e to Figure 2.9h**) values. The performance of the 5 m SUDEM is of importance for future application to soil-landscape analyses in South Africa. Given that the SUDEM still relies on SRTM surface data to supplement the DEM mosaic, the pre-processing of the products and inclusion of localised height variables such as contours and spot heights provides a high-quality surface product at a fraction of the expenditure for similar-resolution products.

Moreover, when considering the slope prediction derived from the SUDEM, the *SlopGrad* is rich with detail, albeit still containing the same "banding" that plagues certain scenes in the SRTM imagery. A significant contribution of these findings is comparing the statistical versus the visual interpretations of the DEM surfaces' findings. For instance, when evaluating the *SlopGrad* results graphically in **Figure 2.9e to Figure 2.9h,** it may suggest that the coarse resolution *SlopGrad* are more variable than the 5 m SUDEM due to "salt and pepper" pixelation of the 90 m DEM surfaces. These experimental results are similar to those observed by Cavazzi *et*

al. (2013), who was also able to show a reduction in statistical variability with an increase in both window and pixel size. Indeed much research has highlighted that coarsening of DEM resolution has a more considerable impact on slope values than on the actual elevation due to the neighbourhood size (Papadimitriou, 2009; Shafique *et al.*, 2011). However, the statistical analysis results presented in **Table 2.3** highlight that the *SlopGrad SD* values of the higher resolution 5 m SUDEM are higher than the *SD* values of the NLD 90 m slope surfaces. Therefore, users need to be aware that while coarse resolution DEMs may not be able to capture the relative detail of their high-resolution counterparts, they do offer a more generalised surface devoid of much of the residual "noise" and micro-terrain variability. This due to the error in a DEM smoothed out to some extent by a larger search neighbourhood (Florinsky, 1998a).

This makes these products particularly relevant for regional soil-landscape predictions of local terrain descriptors. This brings into question the broader concept of "accuracy" that is primordially reference-dependent – with people using widely different references in the "accuracy" evaluation of DEM surfaces (Shi *et al.*, 2012). Since no exact value of terrain attributes, like *SlopGrad*, can genuinely be defined for a real-world land surface, the quality or usefulness of the calculated values of such details is then entirely application and situation-specific. The evaluation has to be based on what is used in that specific application as the reference value. In this instance, the 1 m LiDAR surface.

An exciting finding is that while the vertical MAD may not show any significant variation at a global scale for the respective DEMs when disaggregated to a localised scale, the DEMs exhibit drastically more variation between the different resolution surfaces. **Figure 2.10** graphically represents the variation in DEM elevation values when the landscape is divided into user-defined altitudinal gradients, i.e., elevations values are now compared within smaller "binned" elevation regions. The data are presented on a multi-axis (3D) graph where the y-axis represents the altitude values for the entire site; the x-axis indicates the respective DEM surfaces, and the z-axis represents the MAD upper, and MAD lower values (grey regions) around the 1 m LiDAR *mean* elevation value of 576 m (global) characterised by the red dashed lined. The first observation worth noting is that all the NLD surfaces can either under or over predict the reference surface elevation values by as much as 35 m in the low-mid altitude regions. The 90 m DEMs appear to be over predicting by as much as 50 m in some instances, particularly at higher altitudes above 680 m. the second main observation is the constant increase in MAD values from high-resolution DEM surfaces to the low-resolution 90 m DEMs. This trend is consistently manifested in all six elevation classes but is more defined in the higher altitude regions. As pixel resolution decreases for both the global and local *mean* elevation values, there is a corresponding increase in elevation MAD*,* suggesting that the derived surfaces are less accurate. Additionally, the global *mean* value appears to be most influenced by extreme deviations, most notably at the higher elevation ranges. Thus, it is challenging to determine which product is superior to the other in terms of overall performance since each DEM error is unique and inconsistent in its elevation prediction within the different altitude ranges.

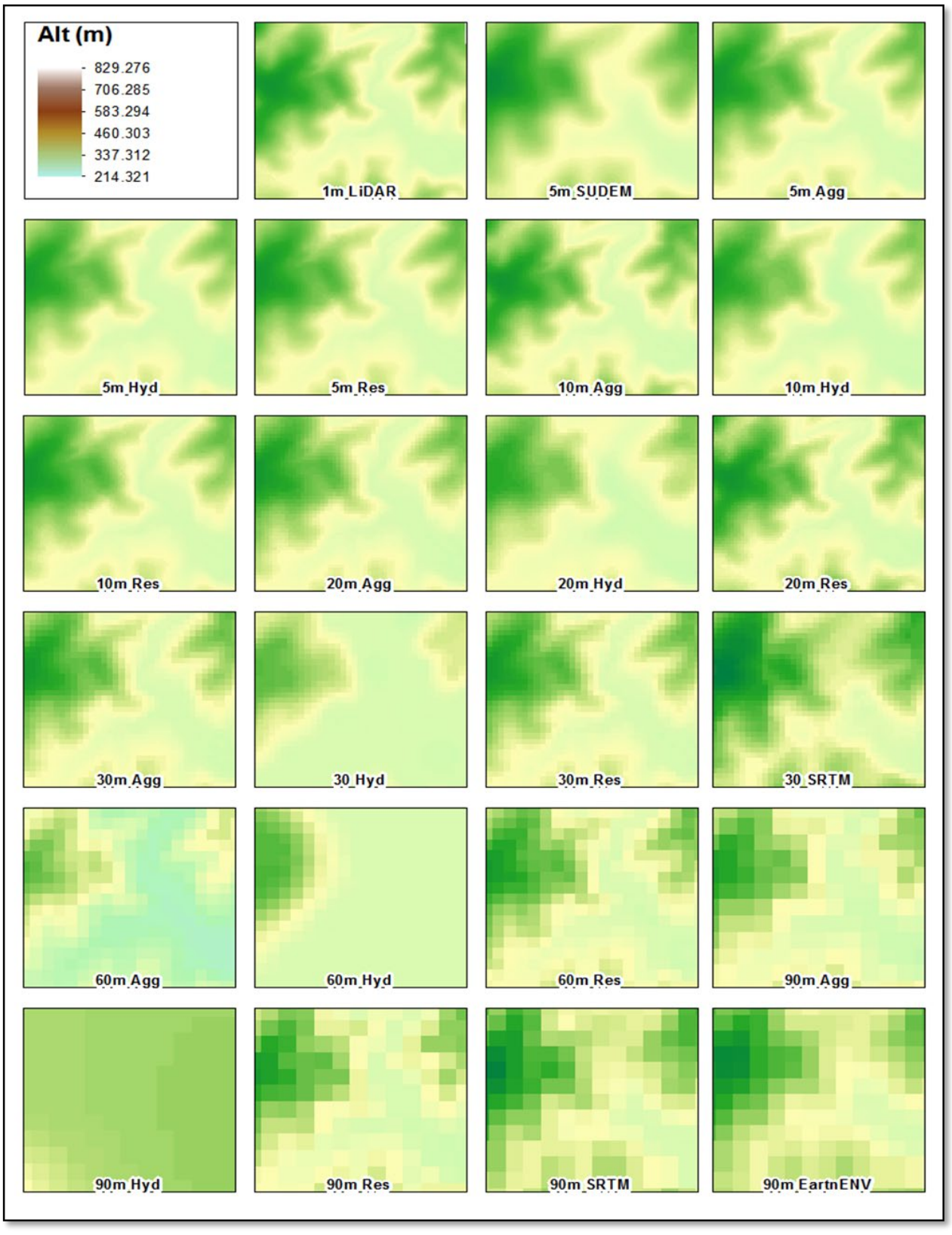


Figure 2.8: DEM surface comparison for a low-elevation region in the study site. Each inset map represents a generalisation approach and one of the six derived resolutions.

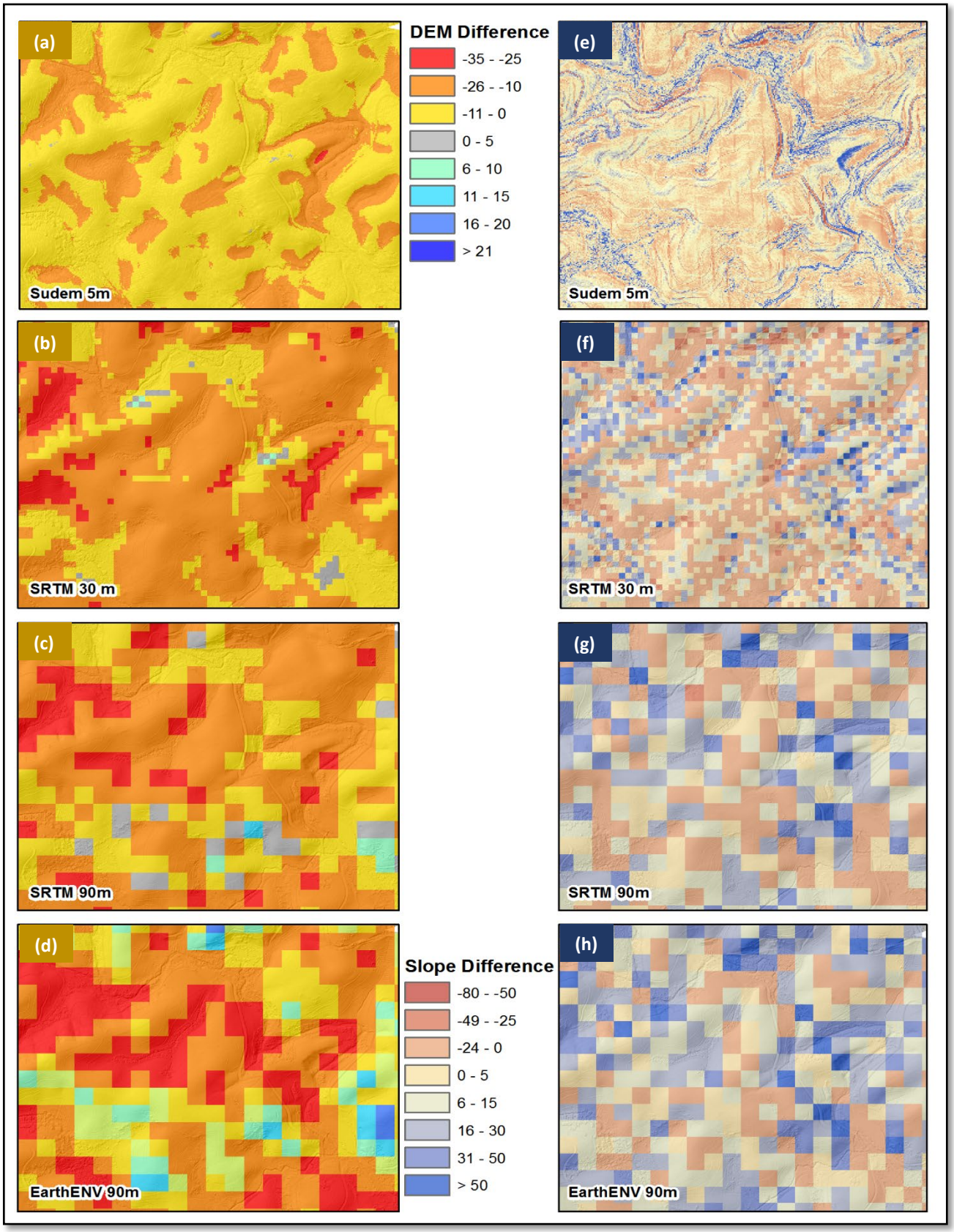


Figure 2.9: (a-d) Elevation error comparisons between LiDAR and non-LiDAR DEMs. (e-h) Percent slope difference between LiDAR and non-LiDAR surfaces. The colour ramp indicates the magnitude of the difference between the LiDAR and non-LiDAR DEM products.

Users' choices of DEM surface models should be based on scientific research and landscape considerations, with DEM selection determined by performance at a global scale, local scale, or a combination of both. Finally, it is possible to observe the HCD DEMs' performance, notably the 90 m HCD surface, which exhibits significant variation at altitudes greater than 690 m. Similarly Chen *et al.*, (2012) reported that *SlopGrad* and surface roughness maps produced using the ANUDEM method or HCD methods tend to omit slopes and obscure much of the terrain features. These findings directly apply to DSM, not only for the present study site but also for the entire region. The maximum elevation of the study site is recorded to be 820 m. The sub-optimal performance of the HCD model at altitudes far below 800 m suggests that the application of HCD models in the KwaZulu-Natal midlands where elevation values exceed 1200 m would not be a pragmatic solution for DSM related analysis. This consideration is essential for Provincial and National strategic planning purposes since the KwaZulu-Natal midlands region is the “bread-basket” of the Province where most of the highest values productive land and highest rainfall region is located. Many soil surveying approaches adopted by seasoned specialists still rely on a mixed-modal approach to soil surveying, relying on GIS systems and traditional methods for scoping, planning, surveying and non-predictive mappings soil resources. Therefore, these findings provide a necessary contribution towards the consideration of relevant and accurate digital data sources by users who aim to produce provincial-based decisions support tools using DSM methods.

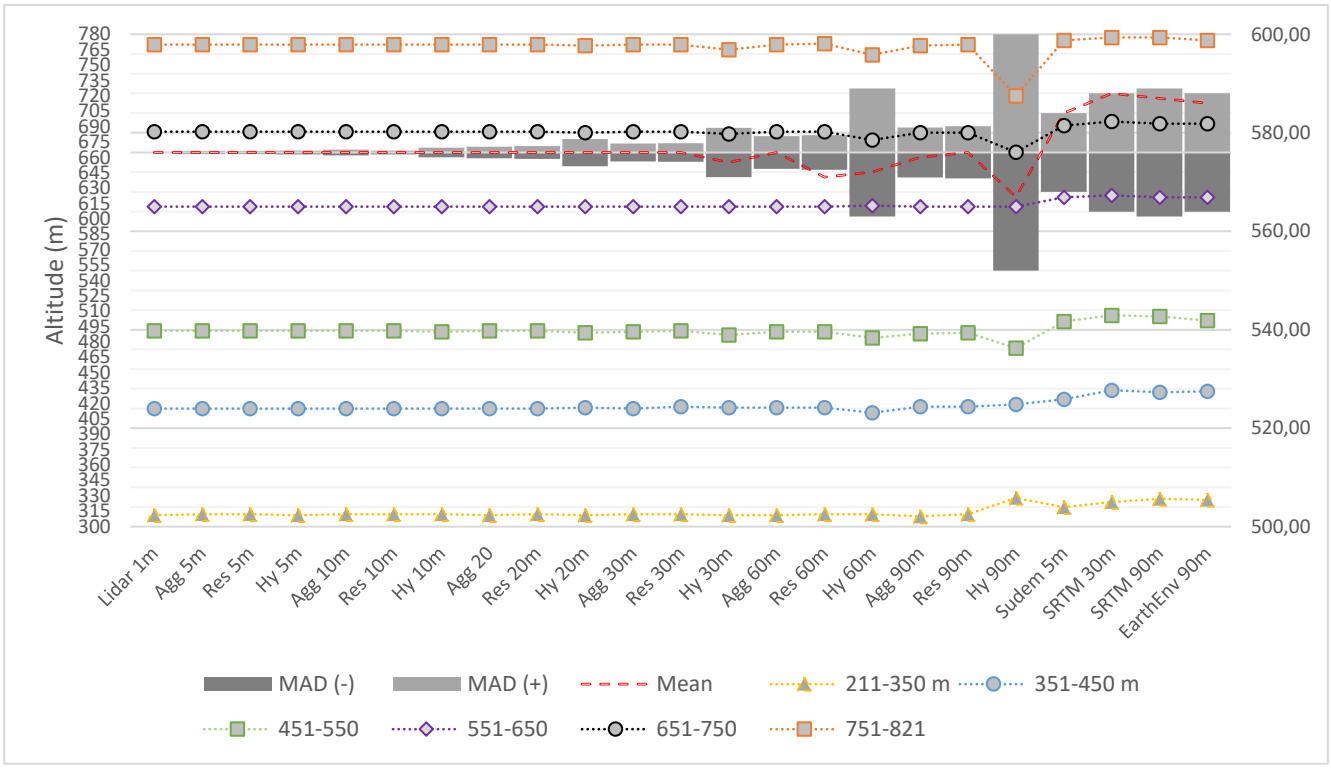


Figure 2.10: Graph comparing regional and localised mean elevation values and influence on MAD for LiDAR and non-LiDAR surface DEM products.

The central conclusion of these results also indicates that at the macro-scale, regardless of the interpolation approach, DEM data source and even pixel resolution does not produce significant errors in surface elevation to exclude their utility in soil-landscape analysis completely. An alternative consideration of these findings, expressly the results of the NLD DEMs, tentatively suggest that, given the excellent approximation of elevation values to the reference 1 m surface, high-resolution DEMs such as the 1 m LiDAR may not be necessary for generating applicable soil-landscape models. To reiterate, this is entirely dependent on the nature and objective of the analysis. In fact, similar to Gonga-Saholiariliva *et al.* (2011), elevation values for the NLD surfaces are comparatively identical to the reference LiDAR surface and thus perform almost as well as the most accurate reference grids despite their comparatively lower initial resolution. This is promising since DEM resolution, precision, and sensor-type issues are increasingly being pushed to the limit for their application in soil-landscape level models, especially where detailed topographic field surveys are not feasible (Thompson *et al.*, 2001). From a soil-landscape modelling perspective, the critical question is whether this conclusion prevails when the feature space's sampling frequency is reduced to a mesoscale level and categorised into more functional elevation classes across the study site. Previous studies have shown that certain terrain variables are sensitive to elevation differences and that the influence of elevation on relief characteristics can influence the prediction of soil-landscape processes (MacMillan *et al.*, 2000; Albani *et al.*, 2004; Wechsler & Kroll, 2006). **Figure 2.10** highlights this point by showing how the *mean* elevation values per elevation class naturally differ from the *mean* value of the reference surface for the entire site and which elevation products contribute most to the deviation for both the global and local elevation values.

A hypothesis was formulated to investigate further this theory of elevation influence on topographic property extraction: H_0 = *There is no significant difference between the actual surface model and derived terrain products at a localised scale*; that is, the predicted products closely match the reference surface. Alternately, H_1 = *There is a significant difference between the actual surface model and derived terrain products at a localised scale*. This hypothesis would statistically describe the spatial similarity or agreement within each altitude class between the 1 m LiDAR and LD and NLD pixel resolutions.

2.3.2 Evaluating the influence of altitude range on the extraction of terrain variables at various pixel resolutions and DEM generalisation methods using ANOVA

The hypothesis analysis results using the single-factor (one-way) pair-wise ANOVA are presented in **Table 2.6**. The following steps outline the ANOVA approach:

1. I set a threshold *P*: $\alpha < 0.05$ based on the relative consequences of missing an actual difference or falsely finding a difference.

2. Define the null hypothesis. Since I compare the *mean* values of extracted terrain attributes, the null hypothesis is that they have the same mean.

3. I then apply the ANOVA test to compute the *P*-value.

4. I compare the *P*-values to the preset threshold value.

5. If the *P*-value is less than the threshold, I then "reject the null hypothesis", and the difference is "statistically significant".

Table 2.5 provides an example of the ANOVA method applied to DEM elevation values within the 230-350 m altitude range. Here altitude is understood to imply the class breaks of altitude range across the study site and must not be confused with digital elevation values extracted from the DEM surfaces. The ANOVA analysis considered all 22 DEM surfaces and included six extracted terrain attributes within specific altitudinal classes. All the respective terrain attribute values for each DEM surface were grouped by altitude class. The ANOVA was calculated as the variance of *mean* attribute values between DEMs within the individual altitude range.

What is important to note is that the null hypothesis, H_0, will be rejected under two conditions: firstly, if the *P*-value is greater than the *alpha level (a),* i.e., 0.05, and secondly, if the *F*-factor is greater than the *F*-critical value. More importantly, if the *P*-value is less than the *F*-critical value, I will fail to reject the null hypothesis and conclude that there is no significant difference between the *mean* values of the different DEM datasets at that particular altitudinal class. Alternatively, should the null hypothesis be rejected favouring the alternative, it would indicate that the mean DEM values, vis a vie the extracted terrain attributes, are different within a particular altitude range. The *P*-value is a fraction, and the term significant is a seductive one. So, it must be stressed that statistical significance does not mean "functional" significance. A low *P*-value indicating statistical significance does not by default imply a consequence of operational importance. In fact, a result that is not statistically significant may yet turn out to be operationally significant (Motulsky & Searle, 1998). Following the description by Motulsky & Searle (1998), the interpretation of significance results hereafter follow two possible explanations:

- The extracted terrain variables are identical to the mean reference surface values, so there is no difference at 5% of the experimentation
- The variables are different, so the H_1 conclusion is correct.

Therefore, the ANOVA results should not be interpreted as contradictory or even compared to the results obtained from the other statistical analyses presented earlier (**Table 2.3**). Earlier results focus on analysing the descriptive statistics, such as the MAD, on investigating operational "significance" between the DEM surfaces, i.e., DEM surfaces that display high MAD values beyond an acceptable threshold, say 100 m, may be significant for applications for terrain roughness analysis.

The DEM ANOVA statistics in **Table 2.6** demonstrate that the DEM products exhibited a significant difference in elevation values for all altitude ranges, with only the first class (230-350 m) with a *P-value* of 4.0% showing minor agreement across all DEM surfaces. The results highlight the altitude classes where the terrain attributes are statistically significant, i.e., exhibit a difference in *mean* values that is not due to chance intended to identify inequalities between the different elevation datasets. As the altitude class values increase, certain DEM products are consistently recognised as significantly different from the actual surface models.

Table 2.5: Variance results for all 22 DEM models for an elevation range of 230-350 m. ANOVA results are included here to indicate the calculation of critical values. Note n=49 means that 49 samples were extracted within the 250-350 m altitude class.

Elevation 230-350 m

Groups	*Point Count*	*Sum*	*Average Alt*	*Variance*
Res90	49	15274.16	311.72	974.96
Res60	49	15270.41	311.64	996.59
Res30	49	15310.33	312.46	981.02
Res20	49	15265.28	311.54	987.13
Res10	49	15275.41	311.74	1001.29
Res5	49	15271.18	311.66	1016.46
Agg90	49	15214.24	310.49	963.15
Agg60	49	15215.04	310.51	1003.52
Agg30	49	15268.93	311.61	984.61
Agg20	49	15249.65	311.22	992.49
Agg10	49	15265.40	311.54	1007.74
Agg5	49	15268.59	311.60	1013.54
Hy90	49	16055.38	327.66	1354.30
Hy60	49	15275.01	311.73	1352.04
Hy30	49	15221.54	310.64	971.76
Hy20	49	15242.36	311.07	963.09
Hy5	49	15262.64	311.48	1005.81
Hy10	49	15268.10	311.59	1001.62
SUDEM	49	15619.59	318.77	1002.89
EarthEnv90	49	15996.00	326.45	1047.71
SRTM90	49	16003.00	326.59	1019.62
SRTM30	49	15883.00	324.14	1059.83

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	36569.62	22	1662.26	1.61	0.04	1.55
Within Groups	1138597.60	1104	1031.34			
Total	1175167.215	1126				

As altitude class-values increase, the *P*-values become much more significant with values reported to be < 0.05, and the *F-statistic* values are also higher than the *F-critical* values suggesting that there are statistical differences (differences in *mean* values) between the actual surface model and LD and NLD terrain products at a localised scale (H_0: rejected). Put differently: the derived products exhibit considerable differences to the actual 1 m LiDAR surface model at predominantly all altitudes across the study site. Therefore, I reject H_0 = *There is no significant difference between the actual surface model and derived terrain products at a localised scale.* The ANOVA analysis is considered a supervised filter technique for linear or univariate data, providing more in-depth insight into the driving forces of prediction (Behrens et al., 2010). Hengl & Reuter (2011) and

Hodgson *et al.* (2005), albeit with different parameterisation criteria, found similar statistical significance between other DEMs across different elevation classes despite different parameterisation criteria.

2.3.3 a posteriori testing to determine pairwise comparison significance of LiDAR-derived and non-LiDAR derived DEMs from 1 m LiDAR reference surface

It must be noted that the summary ANOVA results presented in **Table 2.6** do not provide an indication of which DEMs differ significantly from the reference surface, only that there is a difference. While the ANOVA may identify differences between datasets, it cannot specify which datasets contribute to this difference. A Tukey-Kramer (HSD) *a posteriori*, or posthoc, test was therefore applied to all DEM surfaces across the altitude classes to calculate a new critical value that can be used to evaluate if there is a significant difference in *mean* elevation values specifically between two sets of DEM surfaces, i.e., NLD DEMs compared to 1 m reference DEM and LD DEM compared to 1 m reference DEM. The critical value, known as the studentised range statistic (q), provides the benchmark for which the mean differences between the DEM surfaces must be exceeded to achieve significance (**Equation 2.10**). I have selected a risk value of 5% to determine the critical value q, which is used to compare the means' standardised difference.

$$q = \frac{Mean_1 - Mean_2}{\sqrt{\frac{Mean\ square\ error}{n}}} \tag{2.10}$$

The difference between all possible pairs of elevation *mean* values are then compared to the HSD critical value. If the difference is larger than the HSD value, the comparison is significant (Wickens & Keppel, 2004). With that in mind, the HSD posthoc test is conventionally applied to simultaneously detect significance between all possible pair-wise comparisons across all data *means*, i.e., this would also require comparing differences between LD and the NLD DEM products. I am not particularly interested in the pairwise comparison between LD and NLD DEMs and their respective extracted attributes with the HSD approach. Instead, the objective is to determine how similar the derived characteristics from these two DEMs are to the reference surface and which DEMs are statistically significant at a 95% confidence interval. Therefore, I consider only the difference between LD and NLD products from the 1 m LiDAR surface. This approach conveniently requires fewer overall HSD comparisons to be performed. Typically, HSD pairwise comparisons are imputed on all datasets within the acceptable range of analysis, i.e., each DEM surface is compared to the reference surface and its "competing" DEM counterparts **(Figure 2.11a**). For this study, I modified the Tukey-Kramer HSD to adapt the pairwise comparisons of only DEM surfaces to the 1 m reference surface (**Figure 2.11b**).

Table 2.6: ANOVA results for the six terrain attributes calculated for all 22 Dem surfaces at six different altitude ranges.

DEM

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	36570	22	1662.3	1.6	0.04	1.55
351-450	75271	22	3421.4	4.1	0.00	1.55
451-550	144338	22	6560.8	10.4	0.00	1.55
551-650	74239	22	3374.5	4.8	0.00	1.54
651-750	173741	22	7897.3	12.1	0.00	1.54
751-800	77063	22	3502.9	38.1	0.00	1.56

TRI

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	24194	22	1099.7	141.9	0.0	1.55
351-450	47499	22	2159.0	311.0	0.0	1.55
451-550	68987	22	3135.8	471.2	0.0	1.55
551-650	87357	22	3970.8	437.7	0.0	1.54
651-750	92972	22	4226.0	469.6	0.0	1.54
751-800	7873	22	357.9	58.2	0.0	1.56

Slope

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	69505	22	3159.3	26.7	0.00	1.55
351-450	113683	22	5167.4	43.9	0.00	1.55
451-550	126518	22	5750.8	59.0	0.00	1.55
551-650	117239	22	5329.0	42.5	0.00	1.54
651-750	159832	22	7265.1	57.7	0.00	1.54
751-800	7856	22	357.1	4.2	0.00	1.56

Texture

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	36	22	1.6	180.8	0.0	1.55
351-450	71	22	3.2	365.3	0.0	1.55
451-550	94	22	4.3	506.4	0.0	1.55
551-650	171	22	7.8	822.4	0.0	1.54
651-750	161	22	7.3	726.9	0.0	1.54
751-800	17	22	0.8	82.4	0.0	1.56

TWI

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	2133	22	96.9	28.7	0.00	1.55
351-450	3726	22	169.4	43.6	0.00	1.55
451-550	5119	22	232.7	107.8	0.00	1.55
551-650	5845	22	265.7	73.8	0.00	1.54
651-750	6008	22	273.1	104.3	0.00	1.54
751-800	979	22	44.5	34.3	0.00	1.56

LS

Alt Range (m)	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
230-350	1593	22	72.4	8.5	0.0	1.55
351-450	3019	22	137.2	19.5	0.0	1.55
451-550	2953	22	134.2	24.4	0.0	1.55
551-650	10291	22	467.8	55.6	0.0	1.54
651-750	5857	22	266.2	40.4	0.0	1.54
751-800	497	22	22.6	4.7	0.0	1.56

SS= Sum of squares within, MS=Mean square within-group variability, df= Degrees of freedom

The results of the HSD pairwise comparisons between the LD and NLD surfaces and the 1 m reference surface are presented graphically in **Table 2.7** in a lookup format. This should allow rapid assessment and comparisons of DEM source and generalisation approach across terrain attribute and altitudinal range.

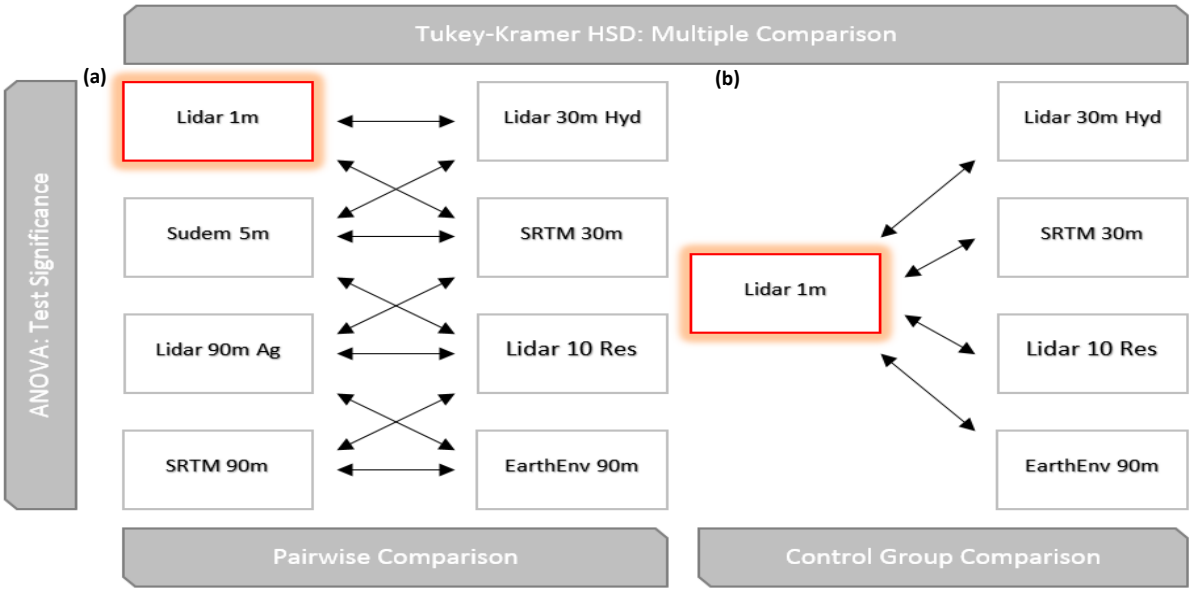


Figure 2.11: Process-flow schematic of HSD modification for one-way pairwise comparisons between LD and NLD surfaces and the reference 1 m LiDAR surface.

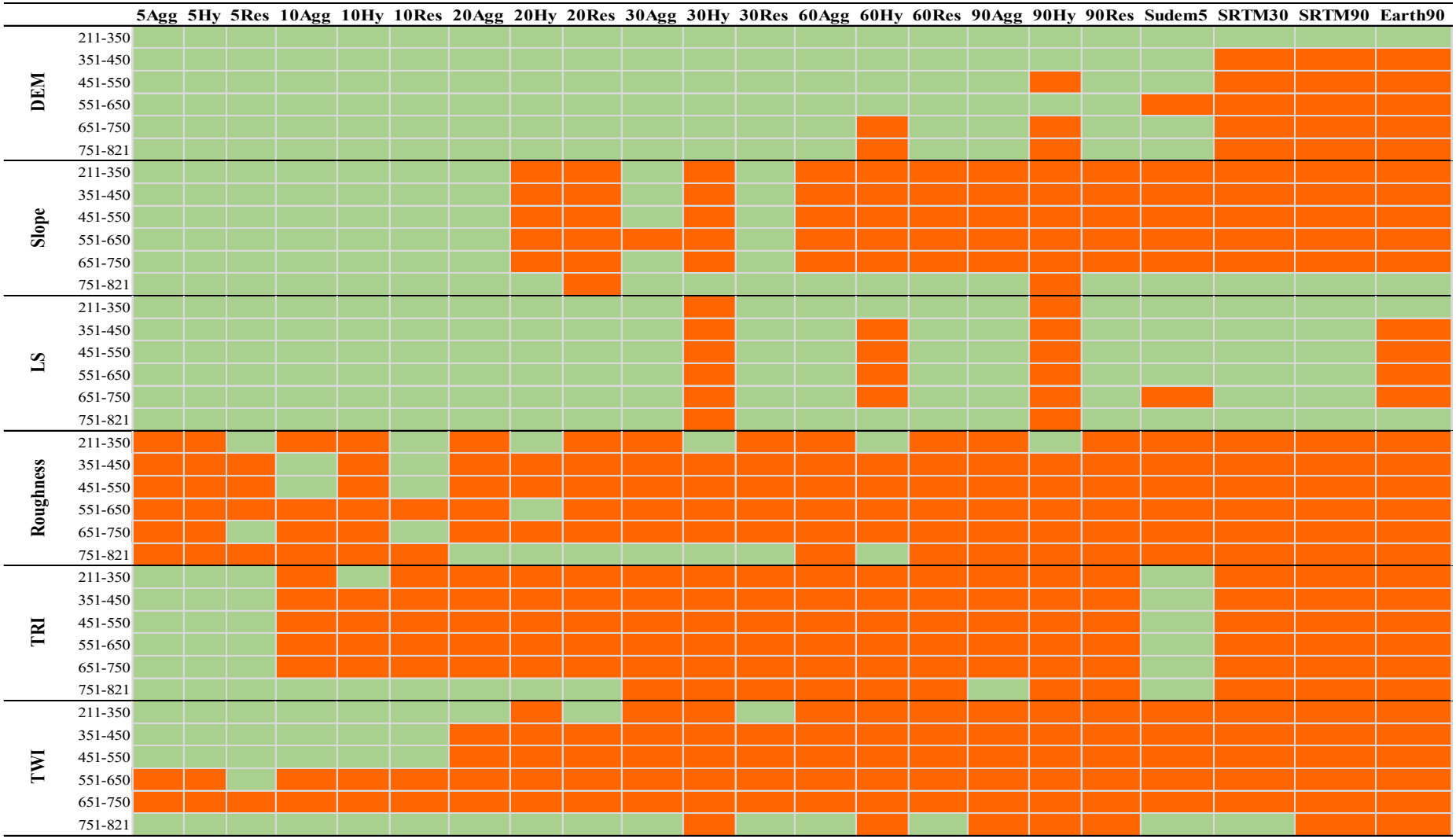
It must be pointed out that the ANOVA results may differ from the results of the HSD posthoc analysis. This is because ANOVA aims to detect whether a factor significantly affects a dependent variable globally. However, multiple pairwise comparisons, such as the HSD tests, involve the computation of a *P*-value for each pair of the compared group. As the number of pairwise comparisons increases and the number of P-values increases, it becomes more likely to detect significant effects due to chance in reality. Put differently, given a significance *alpha level* of 5%, I would likely find 5 significant *P*-values by chance over 100 significant *P*-values (McHugh, 2011). Multiple pairwise comparison tests involve P-value corrections to deal with this problem: P-values are penalised (their value increases) as the number of comparisons increases. Therefore, it becomes less likely to draw erroneous inferences (Sauder & DeMars, 2019). Moreover, the *P*-value penalisation procedure differs from one posthoc test to another. According to XLStat (2020), suggestions for why posthoc tests may vary from the ANOVA global effect include, without limitation:

- a lack of statistical power. For example, when groups have small sizes. When pairwise comparison tests are not statistically robust, it is less likely to detect significant differences.
- A high number of factor levels can also result in varied outputs. The more the pairwise comparisons, the more the *P*-values get penalised. Thus, the risk of rejecting null hypotheses decreases.
- A weakly significant global effect (*P*-value of the ANOVA is equal or close to the significant level).
- A conservative multiple comparisons test. The more conservative the test, the more likely to reject significant differences between *mean* values that, in reality, are meaningful.

Notwithstanding, the variation between ANOVA (*F*-test) and HSD is due to the two tests' nature. For instance, the HSD test is based on the studentised *q*-range that is calculated differently from *F*-test. Moreover, posthoc tests seek to control Type I errors, so the fact that I detected significant results with the global ANOVA suggest that I could also expect statistically significant HSD results. For instance, concerning elevation, the HSD test results in **Table 2.7** indicate that almost all the LD DEMs and the SUDEM 5 m surface are statistically similar to the actual LiDAR surface. In contrast, most NLD surfaces are statistically different across most altitude ranges, except for the lowest and highest altitudes, respectively. Zhang *et al.* (2008) have shown similar results comparing DEM resolution with soil modelling using the WEPP model (Flanagan & Nearing, 1995). There appears to be a general pattern regarding optimal resolution and generalisation approach across elevation ranges, with the NNR and MA methods proving to be the most robust. The HCD approach is again proving to be more inaccurate from 60 to 90 m. These results are encouraging since they align well with the descriptive statistics presented in **Table 2.3**, drawing similar conclusions.

From a data modelling perspective, the results from both the HSD and descriptive tests are promising. Both methods can consider the deviation in the DEM elevation from the reference surface as a significant factor and so highlight the general inability of NLD products to accurately represent elevation values from altitudes of 351-750 m. While the HSD approach uniformly identifies the best performing DEMs within the spectrum of terrain attributes and generalisation approaches, it shows much variability between the altitude class values. For instance, the SUDEM, which I would expect to perform relatively well throughout the altitude ranges given its similar surface representation to the 1 m LiDAR, was reported as being dissimilar to the actual surface elevation at 551-650-m. This result is possibly due to spurious issues of the DEM surface or the effect of elevation outliers on the SUDEM elevation values explicitly within this altitude range. A particular finding worth acknowledging relates directly to the influence of generalisation methods on terrain attribute extraction. The 60 m and 90 m HCDs DEMs could not capture the actual surface model's detail from elevations between 651-821 m. However, the NLD 90 m products were statistically similar to the reference surface at these same altitudes. The question then is whether or not there is an existential explanation for why the HCDs DEMs perform less than expected – specifically relating to elevation - than the NLD within the same pixel resolution? Since the altitude class breaks are user-defined, the influence of sampling point frequency of elevation values at these higher altitudes may seem to be a plausible avenue to explore this variation since low point frequency for the HCD model to generalise the terrain could result in low surface predictions. However, after further analysing the sampling frequency for each altitude range in **Figure 2.12,** we see that the 651-750 m elevation range contains almost 250 sample points using the supervised user-defined elevation range. The data further confirms that even if the elevation ranges for the altitude categories had not been selected using the user-defined interval but rather automated Natural Jenks or Equal Interval categorisation in ArcGIS®, we would have obtained similar sampling point densities regardless. Notably, the user-defined classification approach was consistently applied across all altitude ranges, and LD and NLD DEM surfaces showed similar relationships.

Table 2.7: HSD results for LD and NLD DEMs across the three generalisation methods and six altitude ranges. **Note**: red squares denote datasets where $P > 0.05$, that is I reject H_0 and conclude that there is a significant difference between the derived and reference surface. The green squares denote datasets where $P<0.05$ and I fail to reject H_0 and conclude that the derived and reference surface are statistically the same.



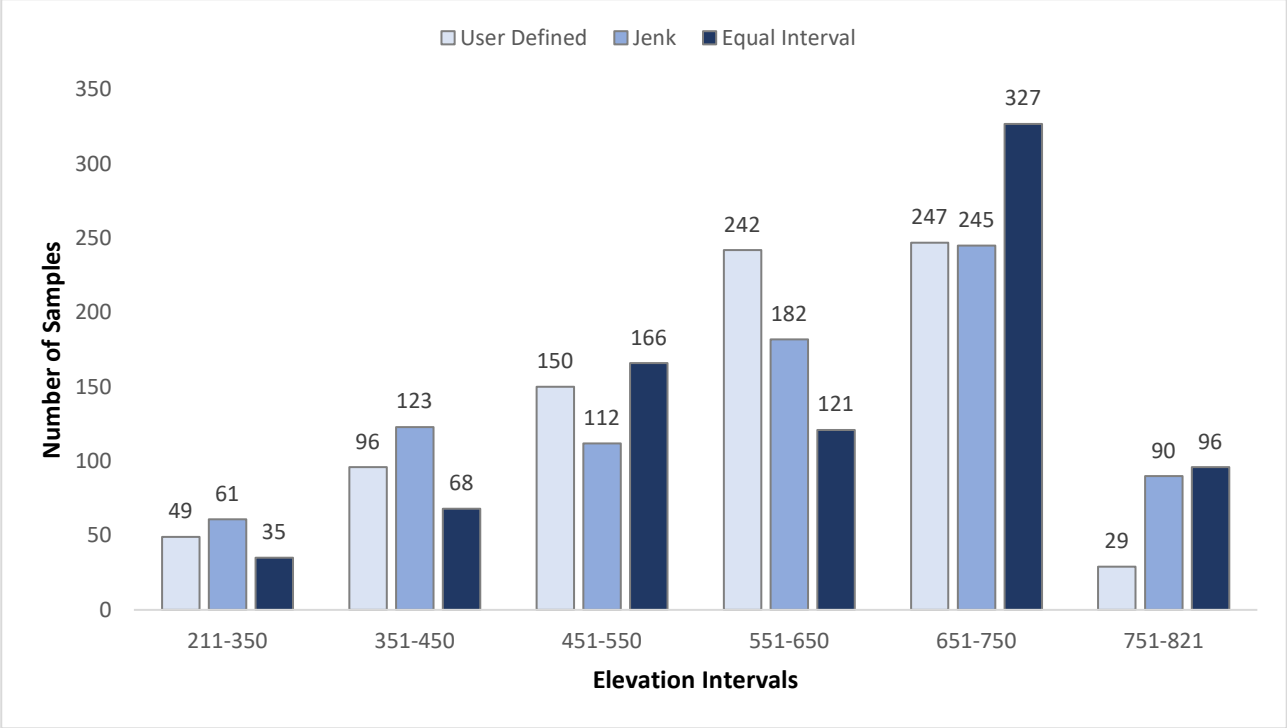


Figure 2.12: Distribution of sample frequency for extraction of terrain variables per elevation class.

By comparing the three categorisation approaches, it is evident that the only real possible explanation for the variation could be from applying the equal interval approach, resulting in more sampling points, 80 more samples to be exact, for the 651-750 m range. Whether more samples per elevation range may significantly influence the analyses' elevation results is not the focus of this research and has not been investigated in detail further. However, consider that the altitude range could be regarded as a distal factor for relative relief and terrain morphological unit (ARC, 2003) occurrence with lower altitudes characterised by less terrain variation while higher altitudes are characterised by more broken terrain features (in terms of slope gradient) (Iwahashi & Pike, 2007). Therefore, an increase in sampling frequency across higher elevation ranges characterised by a higher relief variation, if anything, would explain the distorted performance of the HCD approach accounting for the poor results at higher altitudes. Secondly, Braemar is a *Euclayptus grandis* commercial forestry plantation. The canopy-effect of the vegetation at that particular altitude may have influenced the number of LiDAR pulses penetrating the top-of-canopy and reaching the terrain surface. Cowen *et al.* (2000) reported that forest cover greater than 80% could result in only 10% of laser pulses reaching the ground. This low density of second-return pulses can contribute to higher occurrences of data voids within the surface model. **Figure 2.13** depicts how canopy cover could present a problem for predicting a DEM surface with the LD HCD generalisation method using data acquired in a densely vegetated timber plantation.

Figure 2.13a represents the original 1 m LiDAR-derived terrain surface's general concept showing a high level of terrain detail. To illustrate how the HCD generalisation method hydrologically conditions the DEM surface and thereby alter the elevation of the surface, I compare the HCD DEM (**Figure 2.13c**) with the NLD DEMs (**Figure 2.13b**). The NLD DEMs typically record the top of surface features and yet are able to outperform the LD HCD products and better represent the reference surface based on the generalisation method. We cannot ignore the possibility that the terrain within that particular elevation range may be characterised by a surface that contains depressions or spurious sinks that the LiDAR DEM may not adequately represent.

By applying the HCD approach, these surface artefacts and channel depressions would computationally be removed or filled to represent the hydrological surface better. Overland flow routing through DEM grid cells requires a DEM without disruptions. The by-product of removing the naturally occurring barrier features to overland flow is an overly generalised surface model. This is the most plausible scenario as most hydrologically conditioned DEM surfaces aim to model flow patterns and not actual elevation data. Therefore, surface characterisation is of secondary importance to flow direction (Djokic, 2011).

While selected studies have reported on the variable influence of canopy density on LiDAR accuracy of surface terrain mapping in a forest environment (Hofton *et al.*, 2002; Andersen *et al.*, 2005; James *et al.*, 2007; Lang *et al.*, 2020), the results presented here do not point to limitations of the LiDAR extraction but rather the inadequacy of the HCD method for representing the entire spectrum of terrain features. Furthermore, the canopy cover would not account for the variation in results between the LD DEMS since both MA and NNR methods could represent the terrain surface better than the HCD method. These findings do tentatively suggest that sensor type and pixel resolution may be less critical than the generalisation approach when using DEMs for topographic analysis at specific altitudes.

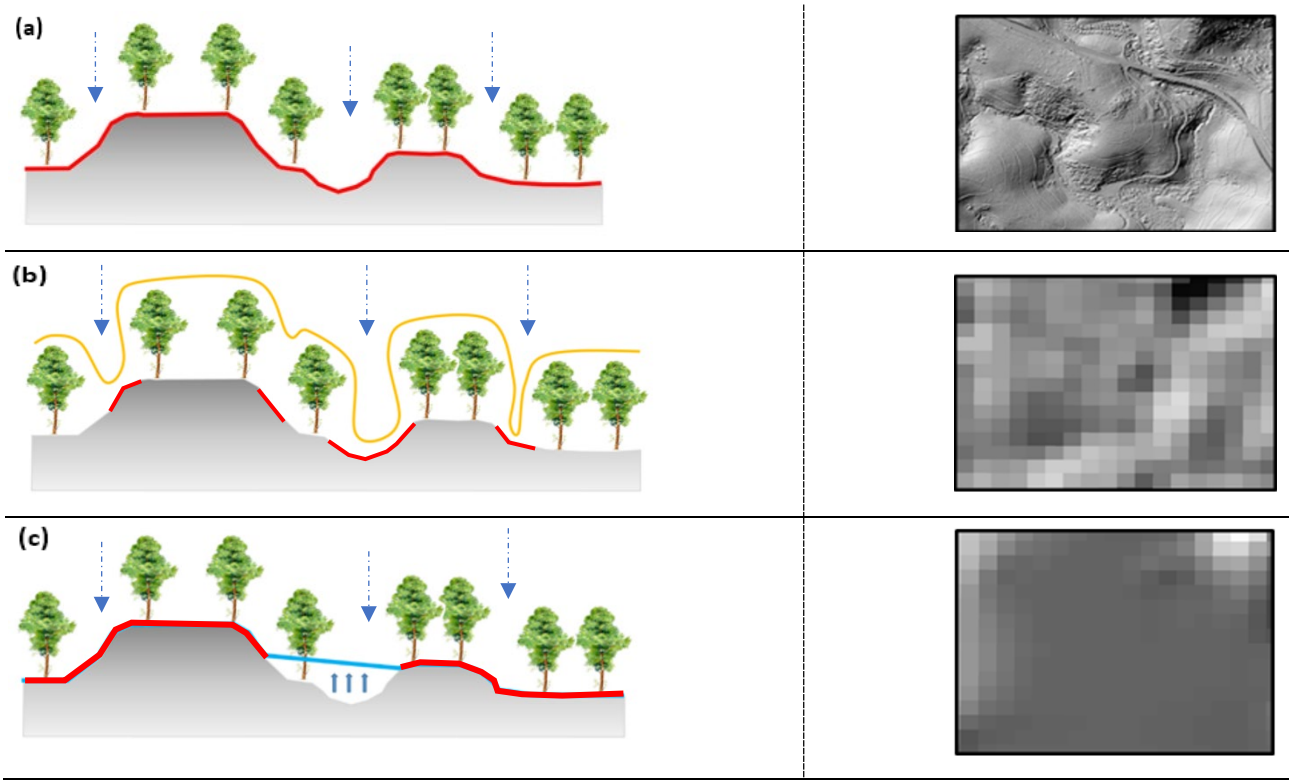


Figure 2.13: Diagram showing the general representation of how terrain surface is represented between the LD and NLD sensors with *hillshade* inset for indicating the influence on terrain detail. (a) Detailed 1 m LD bare surface terrain DEM (b) 30 m NLD top-of-canopy surface DEM where elevated features dominate elevation heights and where signals can penetrate canopies then represent terrain, (c) LD HCD representing terrain surface. Still, where the surface contains sinks, the ANUDEM model fills the voids (blue region) for overland flow modelling. Note the poor terrain representation depicted in the hillshthade inset.

There are no specific or clear answers for the observed performances or inconsistencies of the HCD method across various altitude ranges tested in the study region. Moreover, to further pursue any detailed empirical explanation is beyond the scope of this research. However, the investigation into the spatial heterogeneity of errors associated with the quality and preparation of DEMs affecting model performance has been addressed in previous critical studies by Florinsky *et al.* (2002); Kienzle (2004); Su & Bork (2006); Hengl (2006); Reuter *et al.* (2009) and Jarihani *et al.* (2015). However, similar anomalies, as observed in this present study, have been discussed in the literature. Le Coz *et al.* (2009) point out that with regards to hydrological simulations, model efficiency generally decreases with DEM aggregation, but results are case-specific. Moreover, generalisation approaches modify the topographic basin properties and the model efficiency in a specific way depending on the selected method applied and even software used. Cho & Lee (2001) and Kienzle (2004) concluded that although higher resolution DEMs generated with hydrological approaches produced higher simulated runoff predictions, lower resolution DEMs ultimately had smoother surfaces.

Furthermore, Wu *et al.* (2008) showed that certain terrain features such as slope varied by resampling a 10 m DEM to lower resolutions (up to 200 m) and deriving topographic parameters from these coarse DEM products consistently with lower resolutions. Still, other characteristics, such as the flow path, did not exhibit consistent changing patterns. Studies in terrain analysis are increasingly focused on assessing the mathematical derivatives of elevation and DEM resolution.

These derivations are recognised as providing a more practical test of error location and propagation than elevation itself because of their greater sensitivity to error, thus offering a better measure for DEM quality assessment (Wise, 2000; Wechsler, 2003). Furthermore, critical evaluation of primary surface derivatives like slope, aspect and terrain curvature provide the basis for detailed landform characterisation and interpretation (Evans, 1998; Wilson & Gallant, 2000). This similarly provides a more comprehensive repository of environmental variables for DSM and soil-landscape covariate selection.

2.3.4 Evaluating the horizontal displacement across DEM surfaces based on DEM vertical error

The vertical positional errors associated with applying various surface generalisation approaches and pixel resolutions may have significant implications regarding landscape feature alignment, contiguousness, size and shape (Datta & Schack-Kirchner, 2010). Significant horizontal differences in DEM surfaces can contribute to geometry errors in particular terrain attributes that may rely on the accurate estimation of surface area in their calculation, i.e., *ProCurv*, *PlanCurv, SlopAsp* and *TST*. Like vertical error, horizontal error in DEM surfaces can be generated from a variety of factors ranging from the inherent properties of the sensor used to the processing of the raw data to the type of topography surveyed and physical properties of the surface (Jarvis *et al.*, 2004; Falorni *et al.*, 2005; Brenner *et al.*, 2007). The evaluation of horizontal displacement (HD) and the relationship between vertical accuracy in this study is by no means as systematic as similar studies by Reinoso (2010) and Reinoso *et al.* (2016). However, the results provide a graphical representation of the functional implications of HD performance on sensor selection, pixel resolution and generalisation approach on DEM surface interpolation and terrain feature extraction. For a detailed description with focused applications and analysis of vertical and horizontal error assessment, the reader is referred to Florinsky (1998b); Kinsey-Henderson (2007); Yang *et al.* (2014) and Gillin *et al.* (2015).

Generally, Koppes' formula (Maling, 2016) is used for calculating vertical displacement error from DEM surface altitude since the equation (**Equation 2.8**) includes a reasonably rudimentary, mathematical assessment of HD. By modifying Koppe's original formula, I compute the DEM HD rather than the vertical error of each LD or NLD DEM surface. For a constant error in the horizontal vector direction of observation, the vertical elevation error is expected to increase as the slope gradient of the terrain surface increases (Hodgson *et al.*, 2005). The same assumption is made for HD since Koppe's formula is fundamentally based on the terrain surface slope's tangent (α). It must be noted that the HD values reported in **Table 2.3** are not only a result of horizontal misalignment error but are also most likely exaggerated due to variations in DEM pixel resolution (Wang *et al.*, 2013). While horizontal displacement errors related to DEM surfaces' sensor predictions do influence surface feature identification, DEM pixel resolution ultimately determines the identification of surface features. The overriding consideration here is that pixel resolutions must detect surface features with a width equal to or larger than the sensor's DEM resolution (Stott, 2010).

The DEM HD results presented in **Table 2.3** remained relatively low for the MA and NNR generalisation methods up to 10 m DEM resolutions but then increased by a factor of 10 as the pixel resolution decreased to 90 m. The HCD surfaces exhibited the highest standard deviation in HD concerning pixel resolution showing increases in HD in the order of a 300 from the reference 1 m LiDAR surface for the 60 m HCD surface compared to 0.23 m for the 5 m HCD surface. Given how well the SUDEM has performed both in the statistical and descriptive analysis, a surprising result is that the SUDEM surface exhibited an HD of nearly 34 m from the 1 m LiDAR surface. In contrast, the HCD DEM surfaces showing significant DEM surface generalisation variation could limit the HD at 5 m resolution to only 0.23 m. Furthermore, the 5 m SUDEM HD result is comparable with the coarse resolution SRTM 90 m surface with the EarthEnv 90 m DEM, even outperforming the high-resolution DEM with an HD error of only 6.05 m. Recall, however, that *SlopGrad* greatly influences the estimation of surface error. Therefore, I interpret these results to suggest that while the SUDEM may outperform the coarse-resolution satellite products in vertical accuracy, its high-detail and representation of terrain features, such as *SlopGrad*, may be too detailed and contain high amounts of data "noise" contributing to the high HD error observed. This is further confirmed if one considers the higher *SlopGrad* SD and MAD values of the SUDEM surface than the other NLD DEM surfaces. The practical implications of these HD values for both the LD and NLD DEM surfaces are depicted in **Figure 2.14**. The intention is to graphically represent HD error's influence between the DEM surfaces by evaluating the variation in surface area estimation, based on the HD values in **Table 2.3**, from a standard 100 x100 m digital sample plot within the study site.

In **Figure 2.14**, the grey square represents a hypothetical 100 m x100 m sample plot within the study site covering an area of 10 000 m^2 (1 ha). It is expected that the 1 m LiDAR surface well represents the terrain data extracted within the 1 ha plot. After that, the immediate green boundary represents the surface area error associated with the 60 m MA surface. Although only registering an HD of 4.45 m, this still translates into a surface area of 10 909 m^2 and a surface error of 0.09 ha (900 $m^{2)}$. The most significant misalignment is the 60 m resolution HCD surface with a staggering 65 m HD error and surface area of 2.71 ha – almost double the original 1 ha plot area. The combined graph of vertical error vs horizontal error vs surface area represented in **Figure 2.15** further highlights that the generalisation approach significantly influences inter-class DEM accuracy, more so than pixel resolution in many instances. The SRTM 90 m and the EarthEnv DEM surfaces recorded lower vertical and surface error values than the 60 m HCD method. A real surprise is the lower than expected performance of the 5 m SUDEM, registering a surface area of 1.78 ha. This translates into a surface area error of almost 1 ha. The SUDEM is derived from a combination of high-resolution products (SRTM DEM and contour data). Therefore, I would expect better accuracy, considering it is one of the few high-resolution NLD based DTMs (digital terrain model) available for South Africa.

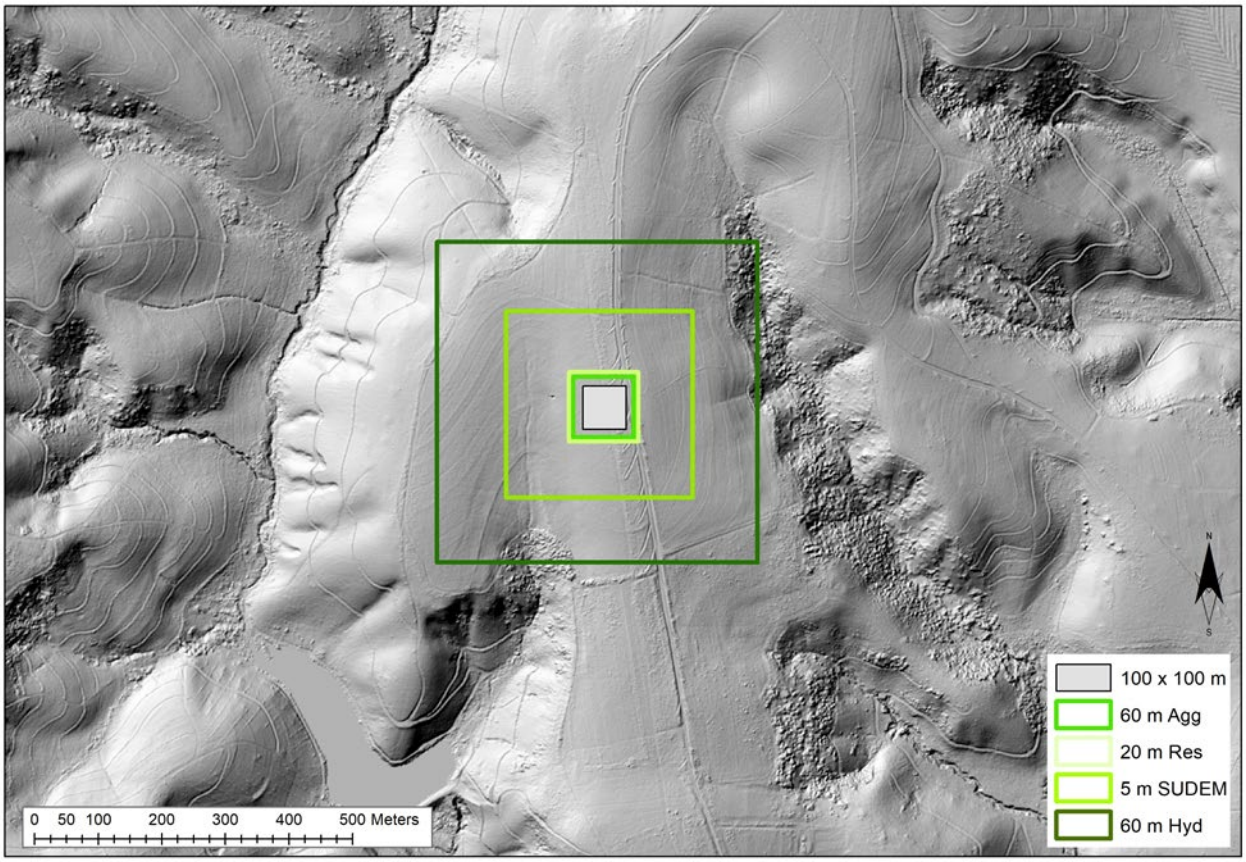
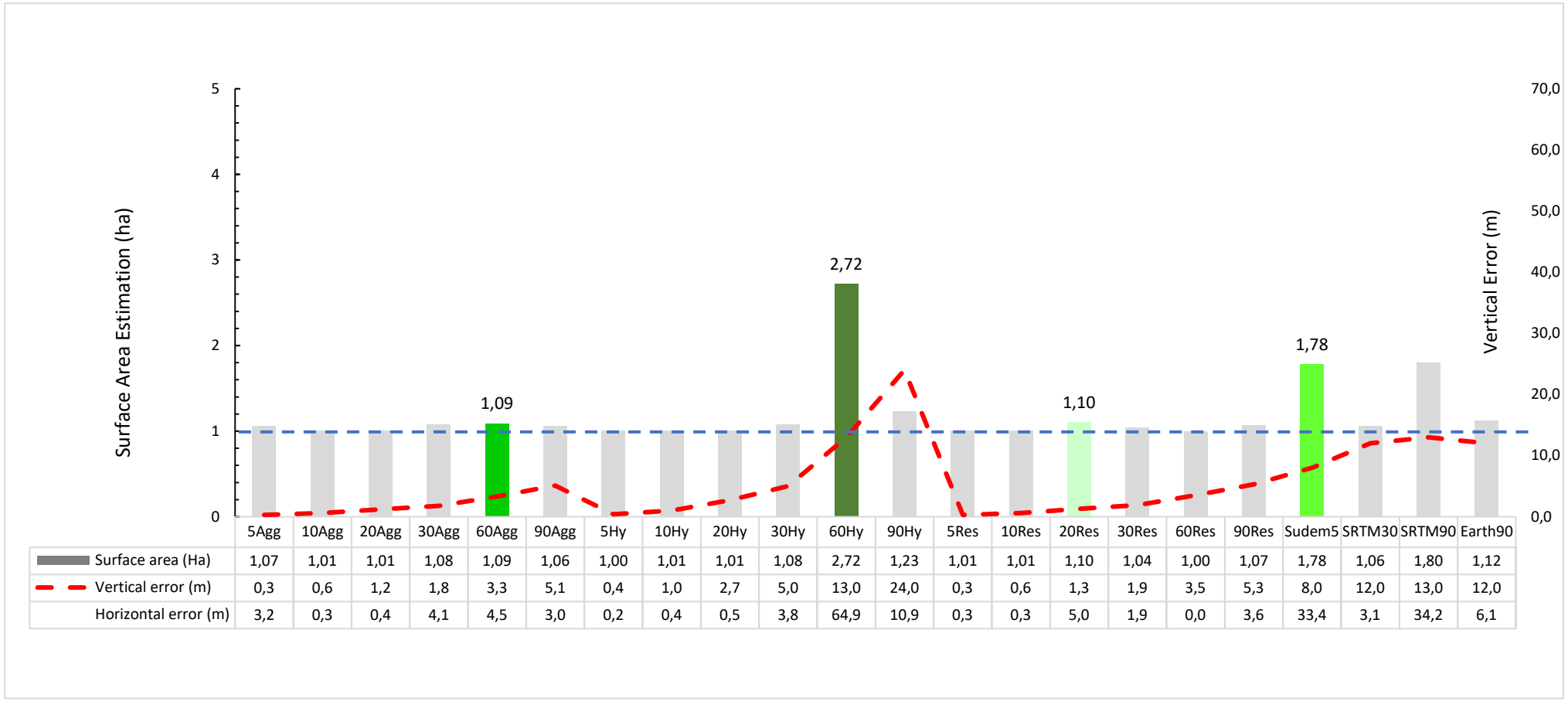


Figure 2.14: Representation of the HD DEM error, as a function of vertical DEM error, on terrain surface area measurements. The grey plot represents a 100 x100m sample plot of interest and contains the required detail of interest. As DEM surface resolution and generalisation methods vary, the effect of increasing surface area error is observed with the 60 m HCD surface showing the highest magnitude of surface area error of all the DEM surfaces.

Van Niekerk (2016) confirms that the SUDEM DEM suite is continually undergoing *adhoc* attribute error correction of surface artefacts. However, he admits that using discrete features such as contours to represent a continuous surface (terrain) should be approached with caution when representing terrain data. Interpolation from contours will produce surfaces characterised by artefacts such as banding, tiger stripes, and wave effects (**Figure 2.8e & Figure 2.16b)**. It is likely that these artefacts are residual to the present SUDEM dataset in this study and possibly account for the low surface area accuracy. The geometric implications on linear and non-linear terrain parameter calculations with the HCD approach, specifically the 60 m product, require serious consideration regarding soil-landscape process modelling. These findings are, however, not uncommon. Using a Monte Carlo simulation of LiDAR DEM uncertainty, Wechsler & Kroll (2006) were able to show that on average upslope area is expected to deviate from the actual surface value by approximately 700 per cent (R-Bias) with an absolute deviation of 720 per cent.

Additionally, Gillin *et al.* (2015) reported the effect of DEM generalisation and surface area error on the delineation of watershed boundaries compared with manually delineated boundaries. The four DEM surfaces generated to evaluate watershed boundary delineation all resulted in different catchment boundaries and watershed areas ranging from -1 ha to 1 ha variation from the field-surveyed boundaries. Their results validate the importance of the present study on quantifying, or at minimum acknowledging, HD's effect on surface feature area estimation. The most significant implications of these HD errors on future soil-landscape analysis

are the possible over and underestimating feature boundaries, e.g. farm boundaries or survey plots. These influence the evaluation of *inter alia,* topographic complexity, terrain relief, and drainage characteristics across the study site. Therefore, it is incumbent upon DSM, landscape, and geospatial practitioners to factor these errors into their landscape and terrain evaluation and determine the optimal tradeoff of pixel resolution, generalisation method and acceptable horizontal resolutions necessary to extract adequately, model and represent the topographic variables for field or -landscape-scale soil mapping. In many instances, this may require a compromise in sensor resolution, vertical accuracy, or horizontal displacement, depending on the study's specific objectives. Based on the HD and MAD results that relate DEM resolution to vertical accuracy, to adequately characterise local topography variability, the vertical precision should increase as the horizontal resolution increases so that the vertical precision remains greater than the average difference in elevation between surface features (Thompson *et al.*, 2012).



	5Agg	10Agg	20Agg	30Agg	60Agg	90Agg	5Hy	10Hy	20Hy	30Hy	60Hy	90Hy	5Res	10Res	20Res	30Res	60Res	90Res	Sudem5	SRTM30	SRTM90	Earth90
Surface area (Ha)	1,07	1,01	1,01	1,08	1,09	1,06	1,00	1,01	1,01	1,08	2,72	1,23	1,01	1,01	1,10	1,04	1,00	1,07	1,78	1,06	1,80	1,12
Vertical error (m)	0,3	0,6	1,2	1,8	3,3	5,1	0,4	1,0	2,7	5,0	13,0	24,0	0,3	0,6	1,3	1,9	3,5	5,3	8,0	12,0	13,0	12,0
Horizontal error (m)	3,2	0,3	0,4	4,1	4,5	3,0	0,2	0,4	0,5	3,8	64,9	10,9	0,3	0,3	5,0	1,9	0,0	3,6	33,4	3,1	34,2	6,1

Figure 2.15: Multi-component graph showing the relationship between vertical error (MAD), absolute horizontal displacement and surface area estimation. The blue dashed line represents the area of the reference plot set to 1 ha. The area of each bar that exceeds this reference line is considered the surface area error. The bar graphs represent the surface area estimation based on the vertical error and HD values for each DEM. The dashed red line represents the vertical error for each DEM surface. **Note**: The coloured bars coincide with the corresponding DEM plot squares illustrated in **Figure 2.14**.

2.3.5 Analyses of terrain parameter results derived from LD and NLD DEMs

2.3.5.1 Slope gradient

The *slope gradient (SlopGrad)* results presented here explain the findings for each DEM surface and the generalisation approach for the respective pixel resolutions. Unlike the previous sections where *SlopGrad* results were only intermittently described concerning DEM elevation, here I provide a thorough review, with graphical comparisons, of the derivation of *SlopGrad gradient* as the second derivative for each DEM surface. Firstly, the calculated descriptive statistics for *slope* presented in **Table 2.3** values show higher variability levels, represented by the *SD* and MAD values, between the generalisation approach than those observed for the elevation *SD* and MAD predictions across the same spatial resolutions. Moreover, there is an apparent inverse relationship between *slope* variability and pixel resolution across all generalisation approaches. Observe, with all three generalisation approaches, the *SlopGrad SD* decreases as the pixel resolution decreases, with the 90 m products of all three generalised surfaces consistently showing the lowest variability. It can be argued that this is to be expected given that the coarser-resolution products represent a more generalised surface with less surface terrain detail and, therefore, less variability. However, the lower variability associated with the coarse-resolution DEMs does not explicitly imply a better *SlopGrad* surface product. Users are advised to consider the descriptive results holistically and interpret the findings collectively to better understand DEM surface integrity. Namely, the MAD results for the same coarse resolution DEMs show a reverse trend of decreasing vertical accuracy with the reduced resolution with the 90 m products, clearly showing errors in slope percentage by factors of 4 compared to the higher resolution 5 & 10 m products. Therefore, it is essential to note that while SD measures provide a necessary metric to evaluate feature variability statistically, the metrics such as MAD provide a pragmatic interpretation of feature error.

Lindsay & Creed (2005) were able to show that depression removal significantly altered the spatial and statistical distributions of derived terrain attributes, including *SlopGrad*. This explains why the HCD 90 m DEM displays the highest SlopGrad deviation of approximately 22% from the reference surface of the three generalisation approaches. Even more concerning is that when considered with the HSD significant test results presented in **Table 2.7**, it is clear that for the 90 m HCD approach, this error is manifested at all altitude ranges, including the lowest, most flat regions. This suggests that even on a wholly zeroed *SlopGrad* surface, we could expect the 90 m HCD surface to represent the *SlopGrad* as having values between 1-22%. For DSM studies intended to evaluate land capability reliant on slope gradient estimations, these error magnitudes can be catastrophic for decision-support based mapping. On the other hand, the lowest *SlopGrad* error is attributed to the 5 m NNR approach showing only a 6% slope deviation from the 1m LiDAR surface. Of the NLD products, the SUDEM and SRTM 30 m DEM surfaces both exhibit MAD values of approximately 11%, with NLD MAD values increasing between 13-15% and a coarsening in spatial resolution to 90 m in the SRTM and EarthEnv *SlopGrad* products, respectively.

What is surprising to note is that all the predicted *SlopGrad* values, regardless of pixel resolution or sensor platform, on average, exhibited a positive bias or MAD. Although this may be expected for the HCD *SlopGrad* products, given the topographical elimination of depression artefacts or sinks across the study site (Wechsler & Kroll, 2006), I did not expect the results to be so uniform across all generalisation approaches. However, the low variability observed between the DEM surfaces and the study site's undulating terrain may offer some perspective on these findings.

Wechsler & Sciences (2007) reported that by sink filling DEMs for hydrological pre-processing, the elevation of cells where sinks are found increased, resulting in a considerable positive bias for DEM elevation. The increased height consequently leads to increased *SlopGrad* estimations in these areas, resulting in positive slope bias. Furthermore, their study reported a higher predominance of depressions in *flat* areas than in higher-relief regions, suggesting that these errors are not limited to the high-terrain areas typically more topographically variable than *flat* areas. Still, Wechsler & Kroll (2006) found similar trends of *SlopGrad* bias when comparing values to upslope contributing area estimators and further reported an overall MAD of 8% across different slope estimations, concluding that the slope values were not significantly different from each other or the measured surface *SlopGrad*. Notwithstanding, this study's HCD approach clearly shows lower *SlopGrad* detail with pixel coarsening from 30 m to 90 m (**Figure 2.16**). The HCD method almost entirely over generalises the *SlopGrad* to such an extent that even the major terrain features expected to be observed at 90 m are not visible.

In contrast, the MA and NNR products, except for the 20 m NNR *SlopGrad*, appear to capture the primary terrain features of the actual surface *SlopGrad* up to DEM resolutions of 60 m. Compared to the 1 m reference LiDAR surface, it is evident that the coarse resolution 90 m DEM products cannot represent the actual *SlopGrad* trends in the landscape. Despite the crude generalisation of *SlopGrad,* predominantly of micro-relief variation, the NLD 90 m products still outperform the HCD *SlopGrad* predictions. Moreover, the DSM appeal of the NLD 90 m *SlopGrad* surfaces is that despite the coarse resolution, the DEM products continually provide the necessary level of terrain detail for large-scale regional mapping where the need for high-detailed topographic surfaces is impractical. Incidentally, the *SlopGrad* results of the posthoc HSD analysis presented in **Table 2.6** confirm that as DEM pixel resolution decreases, the generalised *SlopGrad* is statistically dissimilar to the actual 1 m LiDAR surface model at most altitude ranges. These results show further statistical significance for the HCD method even at DEM resolutions as high as 20 m across most altitude ranges.

As with the DEM surfaces' elevation results that showed high similarity with the reference surface at the highest altitudes (751-821 m), the *SlopGrad* surfaces exhibit similar comparisons with the reference surface across most DEM products. It may be that the higher altitude regions of the study site have a filtering effect on *SlopGrad* since neither sensor platforms, pixel resolution, or even generalisation seem to influence slope gradient predictions significantly. While this may seem inconsistent with previous findings showing how decreased DEM resolution affects the interpretation of terrain detail, these results appear to correspond to Wechsler & Kroll (2006). Their study reported that *SlopGrad* estimators were quite sensitive to DEM surface errors with lower altitude valley areas, especially susceptible to DEM error.

In contrast, upslope area error values were inherently larger in valleys as these locations incorporated the influence of upslope uncertainty. It is also important to add that while the altitude range at Braemar is nearly 1 000 m a.s.l, the study site is characterised by varied undulating terrain with smooth slope transitions (**Figure 2.14**). So, no sharp changes in topography over short distances, such as cliffs, escarpments or artificial features, are expected. Gillin *et al.* (2015) were equally able to show that, in general, the difference between measured field and modelled DEM slope values decreased with DEM coarsening. It is anticipated that *SlopGrad* values will become more intermediate (lower maximum values and higher minimum values) as DEM coarsening averages neighbourhoods of finer-resolution grid cells into larger cells.

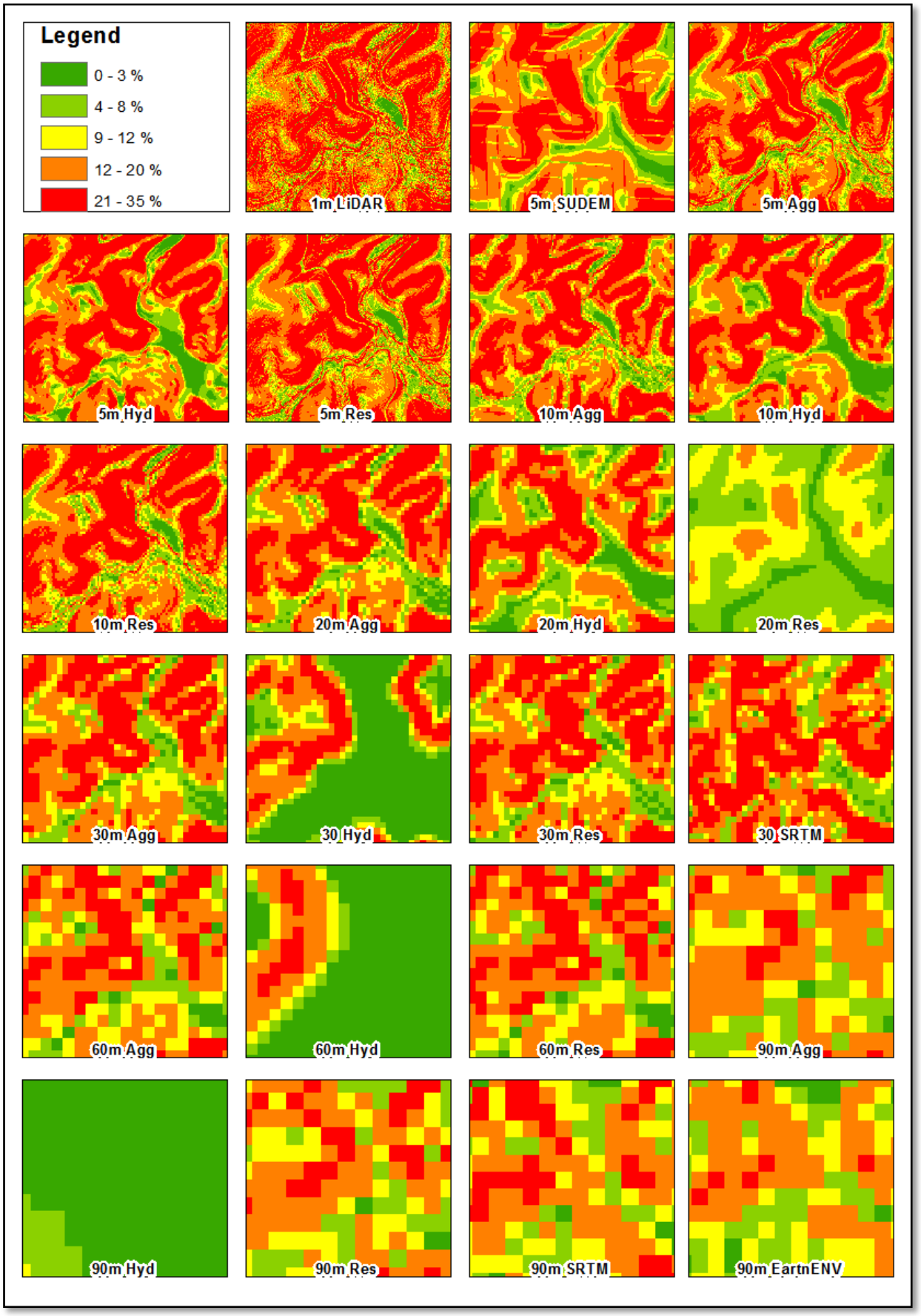


Figure 2.16: *Slope gradient* comparison between LD and NLD DEM surfaces highlighting slope detail with a varying pixel resolution and DEM generalisation method.

2.3.5.2 Slope Aspect

Results of the transformed *Slope Aspect (SlopAsp)*, or the slope orientation relative to the sun's movement, are presented in **Figure 2.17** and have been limited to comparing the LiDAR 1 m reference surface and the NLD DEM products. The LD DEM surfaces were not included in the *SlopAsp* analysis due to the high inter-correlation of surface features between the derived and reference 1 m LiDAR surface. Additionally, while variations in *SlopAsp* would be expected as a result of generalisation variations, the primary focus of the analysis is the comparison of satellite-based surface estimation of *SlopAsp* as a function of the NLD sensor platform and DEM resolution. *SlopAsp* has a significant influence on the distribution of vegetation, biodiversity and agricultural productivity since solar radiation received at a location on the terrain depends on the aspect and shadows cast by terrain (Thomas *et al.*, 2014). The accurate representation of *SlopAsp* values is particularly significant for soil-landscape modelling. In the southern hemisphere, north-facing slopes receive more sunlight than south-facing slopes, and these insolation differences account for variation in soil temperature, moisture and soil thickness (Singh, 2009). The calculation and interpretation of *SlopAsp* values are relatively straightforward (Wilson, 2012). Various studies have explored ways to quantify best the variance associated with aspect estimations across different DEM surfaces Chang & Tsai (1991); Desmet (1997); Raaflaub & Collins (2006) and Stage & Salas (2007).

The present study has, however, applied a similar approach as Goulden *et al.* (2016) in reporting the quantitative decomposed clock-wise north-south and east-west *SlopAsp* results by adopting a simple method of graphically representing the frequency of DEM pixel orientation across the study site using a radar graph (Jenness, 2007; Sharma *et al.*, 2010). The radar chart represents the cardinal direction (Degrees) of extracted sample points (n=1 168) for each DEM surface. The *SlopAsp* radar graph then displays the dominant scalar *SlopAsp* direction (blue region) as a function of sample point frequency of occurrence (yellow line graph) for the respective DEM *SlopAsp* direction across the study site. There are two key findings worth noting with the graphical comparison of the LiDAR reference surface and NLD *SlopAsp* results. First, there is a definite shift in *SlopAsp* representation as pixel resolution decreases with the 1 m LiDAR showing a dominance of *SlopAsp* in the south-east. Similarly, the SRTM and EarthEnv DEMs reported a 45° shift in direction to a dominant southerly direction. Secondly, the results confirm an east/south *SlopAsp* dominance across the study site since all DEM surfaces represented similar direction trends, albeit with varying degrees of frequency, in the same orientation.

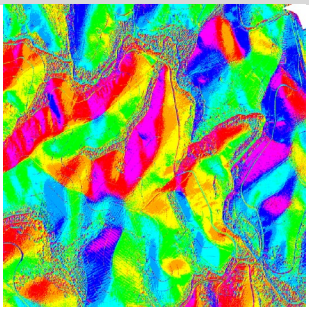
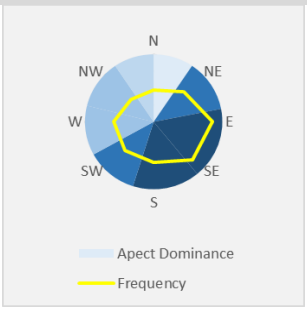
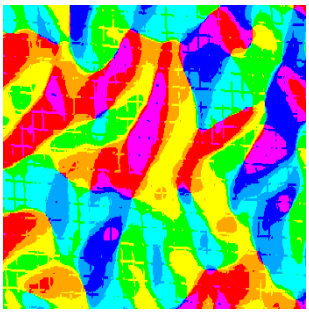
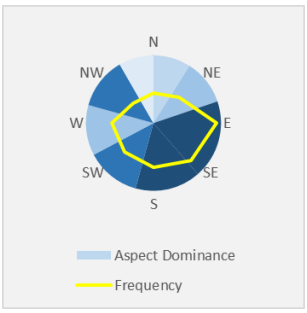
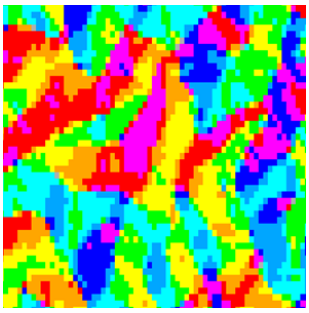
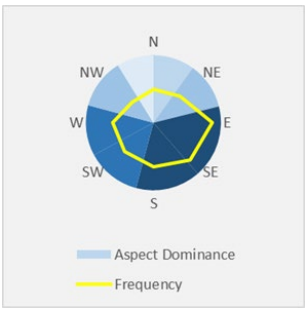
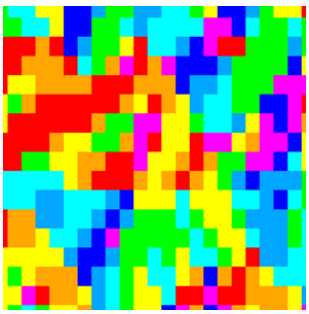
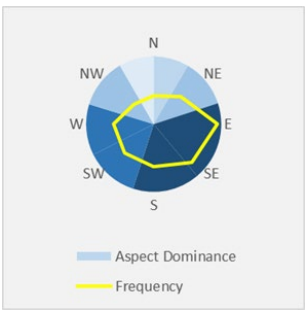
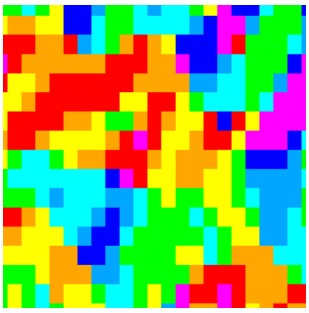
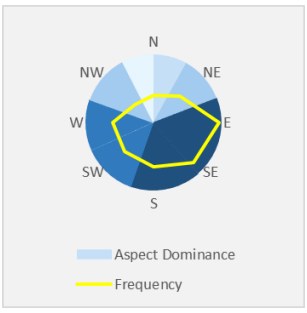
DEM Slope Aspect	Radar graph of *SlopAsp* directional frequency
	
	
	
	
	

Figure 2.17: Aspect representation for a subset of the study site. The Radar graph highlights the orientation of the surface features (Aspect dominance) and the frequency of samples per aspect class (Frequency).

The results confirm that the 1 m LiDAR provides an extremely high level of terrain detail, perhaps more than is necessary for most pedometric applications with possible limitations related to data redundancy, processing time and scale-related issues of modelling soil-landscape relationships. The 5 m SUDEM provides an excellent alternative to the 1 m LiDAR DEM; however, the endogenous "banding" error observed in both the *Elev* and *SlopGrad* surfaces is manifested here in the *SlopAsp* as well, which appear to influence the *SlopAsp* results in the north-western direction. The 30 m SRTM DEM offers users a pragmatic DEM data source that compares well with the reference surface overall *SlopAsp* results. The exception is the inability to detect the micro-relief terrain aspect. The SRTM and EarthEnv DEMs perform relatively well in representing the generalised *SlopAsp* variation of the actual surface model despite having a coarser pixel resolution. While the variation in orientation between the 1 m and 90 m DEMs must be considered, the fact that the coarse resolution DEM surfaces can still adequately represent the major terrain features and orientation is encouraging for broad-scale landscape and DSM studies. These results are supported by similar studies that compared *SlopAsp* maps generated from DEM TOPO-20 m, ASTER-30 m, SRTM-90 m and GMET-250 m (Global Multi-resolution Terrain Elevation Data), reporting almost identical *mean SlopAsp* values modelled for two river basin catchments (Thomas *et al.*, 2014).

I further considered the orientation in *SlopAsp* values for both the LD and NLD DEM surfaces by evaluating the linear correlation with the 1 m LiDAR based on the scalar north/ south and east/west directional values presented in **Figure 2.18.** The line graphs represent the correlation coefficients between the 1 m LiDAR surface and the LD and NLD DEMs for all cardinal directions. The first precise observation is that all DEM surfaces, particularly up to a DEM resolution of 60 m, are positively correlated to the reference surface SlopAsp in the north/ south and east/ west orientation. Second, there appears to be a better correlation with surfaces in an east/west direction than the north/ south direction, with the former showing slightly better correlation coefficients. Finally, there is a gradual decrease in correlation in all directions with decreased DEM resolution with the 90 m HCD*,* SRTM 90 and EarthEnv 90 m DEMs reporting coefficients of approximately 30, 55 and 60%, respectively. Notably, the graph highlights a general decreasing trend (shaded blue) in *SlopAsp* agreement with the reference surface with decreasing resolution and sensor platform. Like Shafique *et al.* (2011), the topographic attributes, such as *SlopGrad*, *SlopAsp* and even *Curv,* are derived from DEMs, employing different generalisation approaches. Consequently, the spatial extent over which these values are computed also varies with the DEM resolution, causing all topographic attributes calculated to change accordingly.

Moreover, these topographic derivatives are so variable over the natural landscape that predicted or modelled values cannot preserve the total *SlopAsp* variability and therefore merely represent the average estimation, or *mean* values, over a particular area (Albani *et al.*, 2004). Studies that have tried to explicitly evaluate the calculation and accuracy associated with *SlopAsp* and DEM relationships are limited and have predominantly focused on comparisons between simulated surfaces and reference values of these variables. However, Florinsky (1998a) argue that there are grounds to think that the accuracy assessment of *SlopAsp* data cannot entirely be determined by comparison of calculated and "reference" values. Mainly since the determination of the appropriate resolution of a DEM is generally a compromise between fully honouring the actual surface and the concern over practical limits of the density and accuracy of the source data (Zhang *et al.*, 2008).

The general impression is that there are still no consistent guidelines for checking the accuracy of land-surface parameters in the field. Therefore, most DEM evaluations and their land-surface parameters are still made visually (Reuter *et al.*, 2009). This could explain why certain studies have found significant disagreements on the spatial distribution of errors of derived *SlopGrad* and *SlopAsp* within different landscape positions.

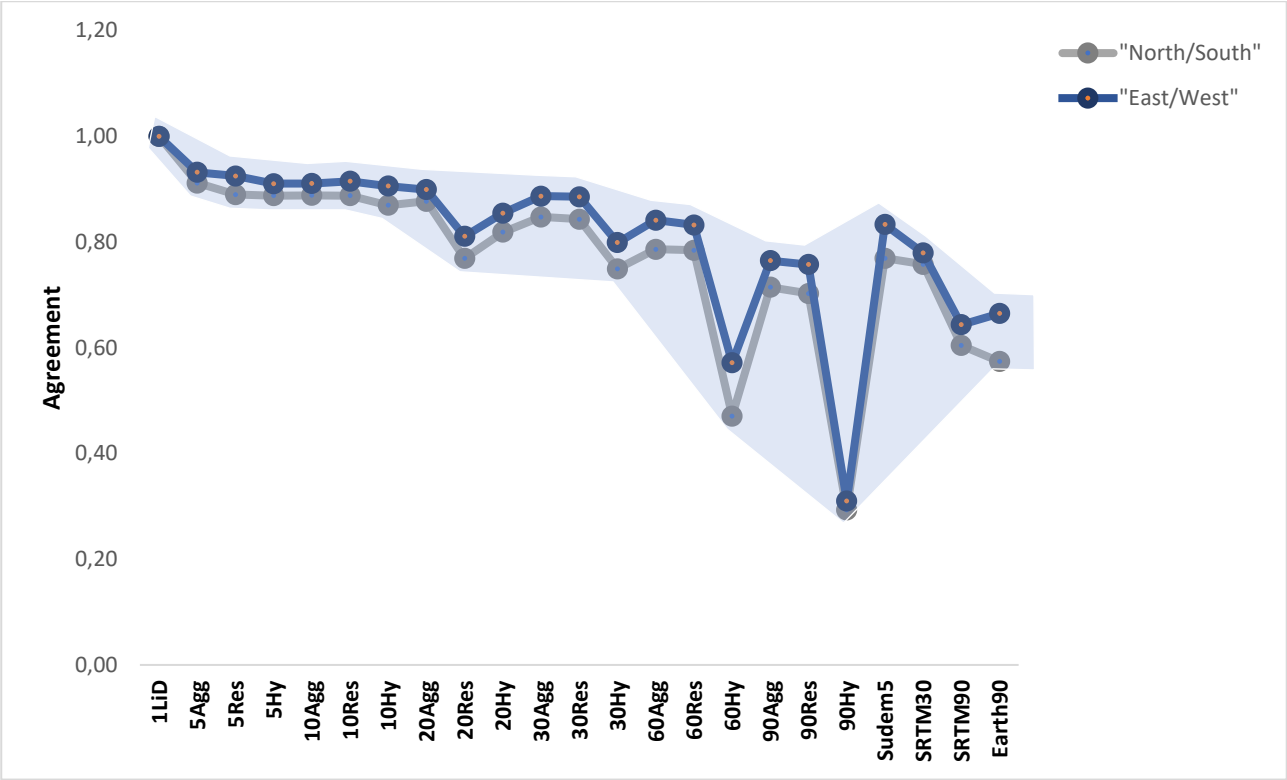


Figure 2.18: Graph showing the relationship between east/west and north/south aspect for each LiDAR and non-LiDAR product. Results approaching 1 show full correlation with reference surface or real surface model. Note also the increase in error (shaded region difference between the maximum and minimum aspect values) as DEM resolution and generalisation methods change.

For a given DEM *SlopAsp*, the absolute value and the standard deviation of elevation errors increase almost linearly with the *SlopGrad* values. Shary *et al.* (2002) provide a detailed theoretical assessment of grid resolution effects on land-surface parameters and objects. Their findings demonstrate that local variables such as *SlopGrad, SlopAsp* and *Curv* are especially sensitive to grid resolution, with second derivatives (*Curv*) more sensitive than first derivatives (*SlopGrad* and *SlopAsp*). Like Florinsky (1998a), I agree that there is no single true or fixed value for local land-surface parameters such as *SlopGrad* or *Curv* at a point but rather a whole range of values that are dependant upon the horizontal and vertical resolution (and neighbourhood size) of the grid used to compute the land-surface parameters and objects. This includes the choice of the algorithm used to calculate them. This is true for the *SlopAsp* results obtained from the different generalisation methods. **Figure 2.18** highlights a definite trend in *SlopAsp* accuracy regarding the interpolation approach, with the MA generalisation approach outperforming the NNR and HCD methods across most pixel resolutions. One possible explanation for the MA performance is that it is computationally similar to the bilinear interpolation approach that uses the four nearest input cells' centre-value to interpolate the weight on the new output raster value. The new value and raster surface for the output cell is then the weighted average of these four values, adjusted to account for their distance from the centre of the output cell (ESRI, 2021). The work by Wise (2011) therefore supports the finding on the utility of the MA generalisation method. Since he too was able to show that when compared to other local interpolation approaches, such as inverse distance weighting, radial basis function and local polynomial, the bilinear interpolation outperformed other methods in calculating *SlopAsp* across a range of DEM surfaces and pixel resolutions. Previous studies have shown some regularities regarding the calculated combinations of *SlopAsp* and DEM properties. Firstly, the maximum error of *SlopAsp* estimation is more likely to occur within flat areas or concentrated within rapid surface change zones such as ridges and ravines (Skidmore, 1989; Davis & Dozier, 1990; Chang & Tsai, 1991; Carter, 1992; Florinsky, 1998b).

Oksanen & Sarjakoski (2005) have also demonstrated the recurring characteristics related to the use of specific error propagation models. Their results highlight that variance in *SlopAsp* calculation appears to be higher with Gaussian than with exponential spatial autocorrelation models, with a critical factor being the correct pairing of cell size and spatial autocorrelation range.

Given that *SlopAsp* and *SlopGrad* are positively related topographic variables (Zhou & Liu, 2004), it is safe to assume the same statistical significance for the different altitude ranges observed for *SlopGrad* in **Table 2.6** may apply to *SlopAsp* as well (Haneberg, 2006). Hence, similar applicability and limitations of sensor platform and pixel combinations would apply. This may explain why the HCD approach with the removal of surface artefacts and "flattening "of the surface has resulted in lower *SlopAsp* correlation with the reference surface since the computation accuracy of HCD land surface parameters has been shown to be inversely proportional to DEM resolution.

2.3.5.3 Slope Length factor (LS)

The *LS* factor (Troeh *et al.*, 1991), a product of slope length (L-factor) and slope steepness (S-factor) is one of the input parameters in many empirical soil erosion models, e.g. the Universal Soil Loss Equation (USLE), including its modified and revised versions. The *LS* factor is dimensionless, having values equal to and greater than 0. It is intended to incorporate topography's influence on soil loss (Wischmeier & Smith, 1978). Previous studies have evaluated the estimation and application of *LS* values within catchments and the impacts on soil erosion (Hickey, 2000; Van Remortel *et al.*, 2001; Van Remortel *et al.*, 2004), with significant differences in application resulting from how overland flow is influenced by the surface terrain (Wischmeier, 1965).

Hickey (2000) further points out that while calculating slope length and slope steepness (slope angle) from a DEM is relatively simple, care must be taken when selecting a suitable algorithm and GIS program for the final calculation. Commonly used GIS software packages such as the ArcGIS® Suite use the *Slope* command that relies on the quadratic surface algorithm developed by Srinivasan & Engel (1991). Other geospatial software used to monitor and model earth systems, such as TerrSet (Clark, 2020), derive *LS* values based on maximum slope gradient. Currently, available algorithms include the neighbourhood method and the best-fit plane method (Srinivasan & Engel, 1991) that calculate an average slope based upon a 3 x 3 neighbourhood in four cardinal directions (N S, E, W). Various limitations in the above methods, such as averages across a 3 x 3 neighbourhood, detecting maximum values, and only using cardinal directions, must be noted. For this study, the *LS* factor was calculated in SAGA, for computational reasons, rather than ArcGIS using a similar approach to Pilesjö & Hasan (2014) by incorporating a multiple-flow algorithm from the SAGA *hydrology* suite of modules and estimating flow accumulation directly from the DEM properties. The added advantage of using SAGA is that it is a cross-platform GIS, and so the algorithms can be loaded as a plugin to QGIS as so-called module libraries. This way, the flexibility and usability of QGIS are coupled with SAGA GIS's power in a completely open-source platform.

The *mean LS* factor values for the 1 m LiDAR surface are relatively high, registering 3.08. These high values are expected given the study site's undulating topography with the lower *LS* values indicative of flatter, lower-lying areas (Panagos *et al.*, 2015). For that matter, the lowest *mean LS* factor values were obtained for the 90 m HCD surface with an *LS* value of 0.79 and MAD value of 2.61 and an *SD* value of 1.14, respectively. This suggests that while the surface may, on average, be underpredicting the existing surface, the 90 m HCD surface is highly variable. Incidentally, the 30 m HCD surface exhibits even more anomalous and exaggerated results

with the *mean LS* factor value of 7.29, a MAD of 7.43 and an astonishing *max LS* value of 57.28, respectively (**Table 2.3)**. The ANOVA analysis also has identified the 30 m and the 90 m HCD DEMs as consistently being statistically different to the 1 m LiDAR surface across all altitude ranges. While the mean LS factor values report a reasonably consistent association with the 1 m reference surface, the high MAD and SD from their respective mean values suggest that the data are highly variable for the remaining DEM surfaces. It must be stated upfront that in the case of *LS* prediction, runoff values similarly vary with varying DEM resolution. With a coarsening of pixel resolution and smoothing by generalisation, we would *a priori* expect lower resolution DEM surfaces to be more prone to under-predicting surface runoff or *LS* factor values. Runoff is higher, and infiltration is lower in areas of steeper topography. More importantly, the higher *LS* values are associated with the surface products that capture the most terrain complexity, with lower values associated with the products that have been "overly "smoothed through DEM generalisation, i.e., HCD. Higher *LS* values indicate surfaces characterised by steeper and longer slopes contributing to a higher surface runoff in the landscape.

Consequently, *LS* values' accuracy is only as reliable as the method used to predict the actual DEM surface. The fact that HCD products, mainly the 5, 10 and 20 m DEM surfaces, have been able to model the reference surface runoff characteristics with negligible variance is uncharacteristic. Theoretically, the HCD approach should less accurately describe the landscape slope steepness and slope angle due to terrain sink or artefact removal, resulting in *LS* values less likely to approach actual reference values (Hickey, 2000). Given that *LS* values appear to be represented by the selected HCD DEM resolutions is promising since much of the products resulting from the HCD generalisation have proven to be sub-optimal in representing the landscape's actual terrain characteristics. However, here I offer two plausible explanations for this optimistic yet typically unconventional *LS* observed result. Firstly, the computational nature of the HCD approach results in a loss of terrain detail and subsequent over generalisation of the surface, leading to an increased representation of slope length and slope gradient and possibly even catchment area (Evans *et al.*, 2009).

These overestimations, in turn, could contribute to increased soil loss per unit area due to the progressive accumulation of runoff in the downslope direction (Prasannakumar *et al.*, 2011), ultimately leading to higher *LS* values as pixel resolution increase beyond a point. In this study that threshold appears to be 20 m. Secondly, the computational utility of the HCD approach will only manifest where DEM surfaces contain apparent sinks that need to be removed to condition the DEM hydrologically. It stands to reason then, if there are no sinks present at a particular location in the landscape that does not need to be filled, the HCD surface should approximate the entire surface at the relevant resolution just as well as the other non-HCD interpolation approaches. This may be the case for the high-resolution HCD DEM surfaces. Whether these assumptions are correct or not would need further investigation beyond this study.

Nonetheless, what the results do show are that given the right pixel/generalisation combination, even coarse-resolution products (60, 90 m), from both the LD and NLD surfaces – except for the *EarthEnv* DEM – through a wide range of altitude ranges can represent the actual surface *LS* values statistically (**Table 2.7)**. These findings may be of regional importance to present and future soil erosion modelling applications, especially since many soil erosion applications are equally concerned with modelling surface runoff and so rely on the application of HCD DEM surfaces. These results are contrary to previous studies by Wu *et al.* (2005), who concluded that a 30 m DEM resolution was adequate for soil erosion assessments while a 100 m should be handled with care due to increased contributing area and decreased slopes. The results have shown that the 90 m LD and NLD, particularly the SRTM, are more than capable of accurately representing the landscape-scale *LS* estimation. The critical consideration is using the appropriate generalisation method rather than merely pixel resolution as a limiting factor. This can be seen with the NLD DEM products that can adequately represent *LS* values, with the SRTM products showing improved runoff representation from the 90 m to 30 m

products. Previous studies have shown similar results to this study by applying SRTM data to calculate (R)USLE topographic factors (Kinsey-Henderson, 2007; Datta & Schack-Kirchner, 2010; Oliveira *et al.*, 2013). The results of the NLD DEM resolution and *LS* factor analysis is encouraging given the reliability and ease of availability of these satellite products to the public, and further studies into the applicability of these, as well as other near-global surface products, for erosion modelling, as well as additional soil-landscape analysis needs further regional investigation.

2.3.5.4 Terrain surface texture

The simplest definition of terrain surface texture (*TST)* is that it represents the standard deviation of residual topographic relief of a local elevation surface (Bennett & Mattsson, 1999; Cavalli *et al.*, 2008). So this implies *strictu sensu* that terrain texture is defined by relief (*Z*) and horizontal distance (*X*, *Y*). Previously referred to as "frequency of ridges and valleys" and "surface roughness" (Iwahashi, 1994; Iwahashi *et al.*, 2001), the metric is now accepted to represent terrain "texture" to emphasise the variability in "grain", or fine-versus-coarse expression of topographic spacing. This is typically achieved by extracting the grid cells (such as "pits" and "peaks") that outline the distribution of valleys and ridges in the DEM (Wood, 1960; Iwahashi & Pike, 2007). Notwithstanding, while surface texture remains as the most accepted term, selected studies have focused on applying different interpretations, including ruggedness (Washtell *et al.*, 2009), rugosity (Wilson & Gallant, 2000; Jenness, 2007), microrelief (Stone & Dugundji, 1965), or microtopography (Herzfeld *et al.*, 2000). For this study, I have intentionally isolated *TST* evaluation from terrain ruggedness (described in detail in §*2.3.5.5*). SAGA considers the two variables to be separate since *TST* is the definition of both relief (feature frequency) and spacing in the horizontal direction where each grid cell value represents the relative frequency (in per cent) of the number of pits and peaks within a radius of ten cells (Iwahashi & Pike, 2007). The *TST* algorithm is refined using a *median* filter, a statistical operation that removes high frequency "noise" by replacing original grid cell values with median values. The filtered DEM is then subtracted from the original DEM, with the final grid containing cells with differences in elevation, located at the surface pits and peaks. In contrast, terrain ruggedness index (*TRI)* is an objective quantitative measure of topographic heterogeneity, calculated as the sum change in elevation between a grid cell and its four neighbour grid cells. Furthermore, *TRI* is represented as absolute values by squaring the differences between the target cell and neighbour cells and taking the square root (Riley *et al.*, 1999).

Several authors have proposed methods to measure the *TST* by LiDAR data, e.g. (McKean & Roering, 2004) (Haneberg, 2006; Frankel & Dolan, 2007; Yang *et al.*, 2014) as well as satellite-derived data (Grohmann *et al.*, 2010; Tran *et al.*, 2014). For this study, *TST* is calculated as the standard deviation of elevation using a neighbourhood analysis within a moving window for both the LD and NLD DEM surfaces using SAGA GIS. The use of SAGA provides an effective and efficient approach for terrain analysis by providing a comprehensive compilation of metrics, requiring only a single sequence for data import, and allowing the user to focus on the individual metrics of interest. Two fundamental shortcomings in the calculation of *TST* estimations is that firstly it is not an intrinsic property of the terrain surface, but depends on the direction of measurement and surface orientation (Hoffman & Krotkov, 1990). Secondly, the complexity of natural terrain surfaces far exceeds the sensor data's dimensionality, rendering roughness estimates non-unique (Mushkin & Gillespie, 2005). Therefore, the actual concept of *TST* is somewhat cognitive and depends not only on the scale of the natural features in the landscape but also on the resolution of the DEM, surface generalisation approach employed, and accuracy of the resulting DEM surface (Rexer & Hirt, 2014). For instance, surfaces that exhibit a high amplitude of terrain features would intuitively have high terrain variability and are considered to have a complex or "rough" terrain surface.

Similarly, highly complex terrain surfaces may have low amplitude of terrain features but have a higher frequency of these features. Likewise, surface features may have low frequency and low amplitude but exhibit irregular spaced features and are still considered highly complex and highly "textured". It is critical to clearly define the scale or level at which *TST* is being considered and the unit which is being used to measure across a "window" of scale, as different results will be relevant at different scales. For this reason, this study evaluates *TST* as a topographic parameter that emphasises the contrast of fine and coarse topographic spacing of features encompassing both relief frequency and spacing within the horizontal plane (Iwahashi & Pike, 2007). *TST* evaluation needs to be considered with other terrain attributes such as relative relief measures (Smith, 2014). Previous studies have defined surface roughness in terms of the variability of elevation, generally expressed as the absolute standard deviation of all values within a window or as the deviation from the best-fit plane (e.g. the 1 m LiDAR surface) (Frankel & Dolan, 2007; Evans *et al.*, 2009).

TST values will therefore be positively correlated with DEM *SlopGrad* that denotes the rate of change of elevation, and *ProCurv* that measures the rate of change of slope. Although *TST* is considered a geomorphometric variable, like Olaya (2009) and Evans (1972), I treat *TST* as a geomorphometric parameter. A variable is a measurable property of a phenomenon, such as *SlopGrad* angle. In contrast, a parameter summarises a population's characteristics, such as *mean SlopGrad.* This allows the findings for *TST* to be reported as summary statistic allowing rapid comparisons within and between DEM surfaces (Grohmann *et al.*, 2010).

Shepard *et al.* (2001) argue that there are no standard methods for quantitatively characterising the anisotropic nature of *TST.* A variety of measurement techniques are available to practitioners. While many measurement techniques differ depending on the scale, sensor technologies, and diverse array of user needs, a set of universal considerations for comparing *TST* measurements has been adopted. These include, *inter alia*, the statistical analyses of terrain properties such as RMSE of terrain height (*Elev*) or RMSE of *SlopGrad* and even *mean* and *absolute* deviation of surface area (Grohmann & Sawakuchi, 2013). Similarly, other studies have exploited variances in surface *ProCurv* (Korzeniowska & Korup, 2016) as proxy indicators for analysing and reporting *TST* results of roughness measurements so that direct inter-comparisons can be made between data sets (Shepard *et al.*, 2001).

Notwithstanding, certain studies have shown that applying specific central tendency statistics may not necessarily be sufficient to diminish the robustness of *TST* measurements across DEM resolutions (Iwahashi, 1994; Zhang & Montgomery, 1994; Guth, 2003). Weeks *et al.* (1996) showed that the one-dimensional universal characterisation of *TST* is inadequate on its own and that as many as five independent parameters may be required for such a task. It is clear then that the complete and comprehensive *TST* assessment across natural surfaces and varying DEM resolutions is beyond this study's scope. Therefore, the assessment methods used in this study for TST assessment have been limited to the most commonly reported for other topographic parameters (Drăguţ *et al.*, 2013). They are meant to be used as a meaningful descriptive comparative within similar scaled products (Shepard *et al.*, 2001) to highlight the spatial patterns of *TST*, rather than a purely mathematical analysis of the *TST* variation across varying sensor platforms and pixel resolutions.

When evaluating the descriptive statistics and ANOVA results presented in **Table 2.3** and **Table 2.6**, it is clear how delicate the interpretation of TST data is within the context of DEM generalisation studies. For instance, small roughness values (< 0.50) would be associated with slight variations within local neighbourhoods, typically coincident with smooth topography. In contrast, a high *SD* typically indicates high variability in the local terrain surface (McKean & Roering, 2004). The findings show a positive relationship of high *TST* representation with decreasing pixel resolution, particularly for the LD DEM surfaces, i.e., the coarse

resolution products represent the terrain as being rougher than the fine resolution products (**Figure 2.7e**). If we collectively consider similar trends in *TST SD* values and *mean SlopGrad* results for both the LD and NLD DEMs, we also observe that increased terrain variability is associated with decreased pixel resolution, as suggested by Shepard *et al.* (2001). Now studies by Mushkin & Gillespie (2005) and Bretar *et al.* (2013), in particular, suggest that the observed trends need further interrogation given that we would expect to see higher *TST* values associated with higher resolution DEM surfaces.

Furthermore, the ANOVA results in **Table 2.6** suggest that *TST* is complicated to replicate across DEM resolution and generalisations since most LD and NLD DEMs show minor similarity with the 1 m LiDAR reference surface. These results should, however, not be dismissed as outright erroneous. Instead, the results confirm the shortcomings in *TST* calculations discussed earlier, based on the assessment of DEM generalisation and pixel resolution methods applied in this study. For instance, while it is understood that *TST* variations are expected to occur over small/medium (local) scales, the frequency of detection of irregular features across a landscape could be higher with reduced pixel resolution (60 to 90 m) since larger areas would allow more features to be detected, albeit at lower detail. This may contribute to the high *TST (*high amplitude of terrain features) and *SD* predictions, particularly in an area such as Braemar, where there are numerous undulating low-altitude peaks and valleys across the landscape. This concept is best described by Snapir *et al.* (2014), who suggest that large-scale roughness features can be sampled appropriately without the need for detailed, high-resolution images. More importantly, their results may indicate merit in assuming that increased spatial extent (not only in cells but in ground units), as a distal factor of increased AOI or surface area, can contribute to higher landscape *TST* detection. Olaya (2009) cautions that while the simplest way to compute *TST* is to use the *SD* of the height in a given analysis window's cells, it is also not a very precise method. Relying solely on *SD* can produce incorrect results, such as assigning high values (rough terrain) to cells constituting a flat *SlopGrad* within an area. McKean and Roering (2004) observed similar results to this study, over a 9 m ground distance extent, and pointed out that even minor topographic variations on very flat areas can cause local changes in *SlopAsp* values of up to 180^{o} between neighbouring pixels. This may also explain why the results for the coarse resolution DEM surfaces are represented as having relatively "rougher" surfaces than the fine resolution products: since the coarse resolution, DEMs tend to generalise the terrain to "smooth" planar surfaces.

Figure 2.20 and **Figure 2.21** detail how the difference in DEM resolution inevitably influences the representation of terrain detail and influences *TST* expression. As per Olaya (2009), we observe first-hand why *SD* values for *TST* at coarser resolutions may report high variability due to the over-estimation of *SD* estimates in planar representations of the terrain surface**. Figure 2.20a** represents the 1 m LiDAR DEM showing the high-level terrain texture over a short transect in Braemar. The corresponding terrain profile graph equally represents the expected high terrain variability (**Figure 2.20b**).

In contrast, **Figure 2.21a** depicts the 90 m SRTM DEM detail, along with the less-detailed terrain profile graph (**Figure 2.21b**). It is evident that the 90 m SRTM DEM oversimplifies the terrain surface and creates an artificial planar surface, especially for peak features, which may cause the higher than expected *TST* values for the course resolution DEM surfaces. Ultimately, DEM resolution variability may be the primary driver in distinguishing between surface-level roughness and relief level roughness (Iwahashi & Pike, 2007). The key to this concept may lie in the findings of Brubaker *et al.* (2013) and Kraus & Pfeifer (1998). They surprisingly suggest that LiDAR-derived DEMs – despite their high accuracy – may contribute to terrain interpretation difficulties and found that high-resolution derived DEMs may over-smooth terrain representation and minimise surface roughness and result in less topographic complexity. In fact, Hoffman & Krotkov (1990) were equally able to prove, using a step-wise function on a synthetic terrain displaying constant Gaussian noise, that the effect of window size influences the extraction of specific *TST* parameters.

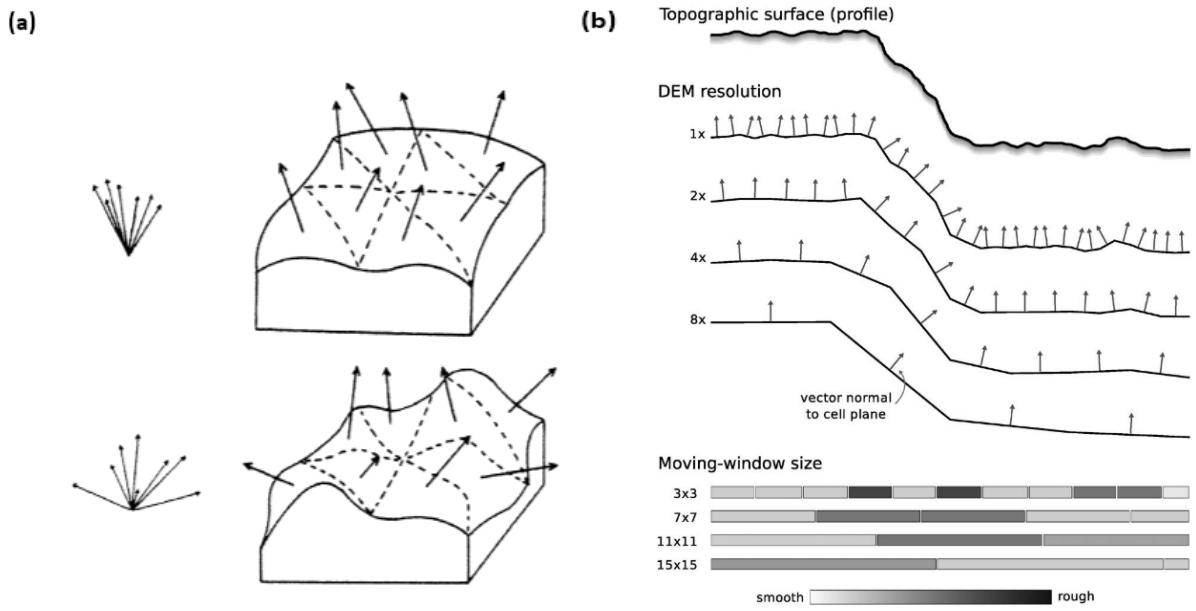


Figure 2.19: (a) Represents *TST* in a hypothetical DEM is represented by unit direction vectors. In smooth topography (top drawing), the local vectors have similar orientations. In rough terrain, the local vector orientations are highly variable. (b) Schematic representation of *TST* variation according to the spatial resolution of the DEM and moving-window size. Taken from (Grohmann *et al.*, 2010).

Figure 2.19 provides a simplified overview of how the increase in the window size used (decrease in pixel resolution) below a threshold results in a lower error in roughness measure. The local vector variables are highly scattered in rough terrain, resulting in lower reported *TST* values (**Figure 2.19a**). Grohmann *et al.* (2010) further emphasised this vector concept by modelling the expected relationship between the DEM spatial resolution and moving window size using a hypothetical topographic profile (**Figure 2.19b**). Iwahashi & Pike (2007) further summarise this concept by stating that coarse-textured relief is highly distance-dependent, in fact, as far as 24 km across the landscape. Fine-textured relief displays a weaker function of distance throughout the spatial domain. Finally, Hobson (1972) explored the representation of *TST* by using directional vectors to explain local (surface) and neighbourhood (Landscape) topographic variability. He reported that the local vectors have similar orientations in smooth topography, resulting in higher *TST* representation. It was reported that only the slope (local relief) breaks are marked with high roughness values with small window sizes. As the window size increases, more extensive relief features (such as scarps) tend to be included in the high-value areas. Notably, there is a point that where the window size becomes bigger than the landforms of interest in the area, it will not be possible to discriminate the response of these features.

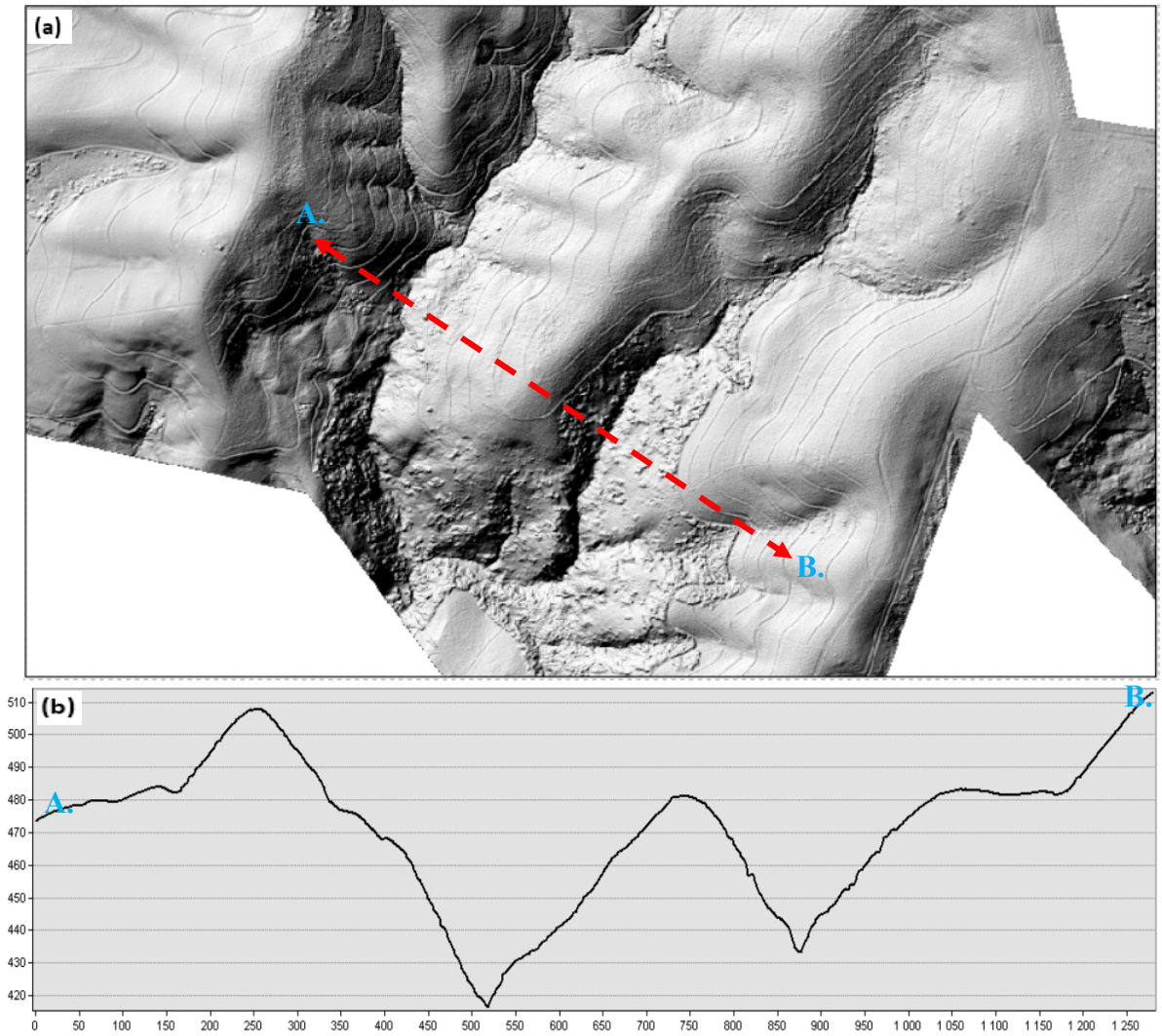


Figure 2.20: (a) 1 m LiDAR DEM surface showing high-detail and high visible *TST*. Note the detection of contour roads detected beneath the canopy cover. The red dashed line represents the transect for the terrain profile graph, (b) Terrain profile graph showing peaks and troughs through transect across the DEM surface.

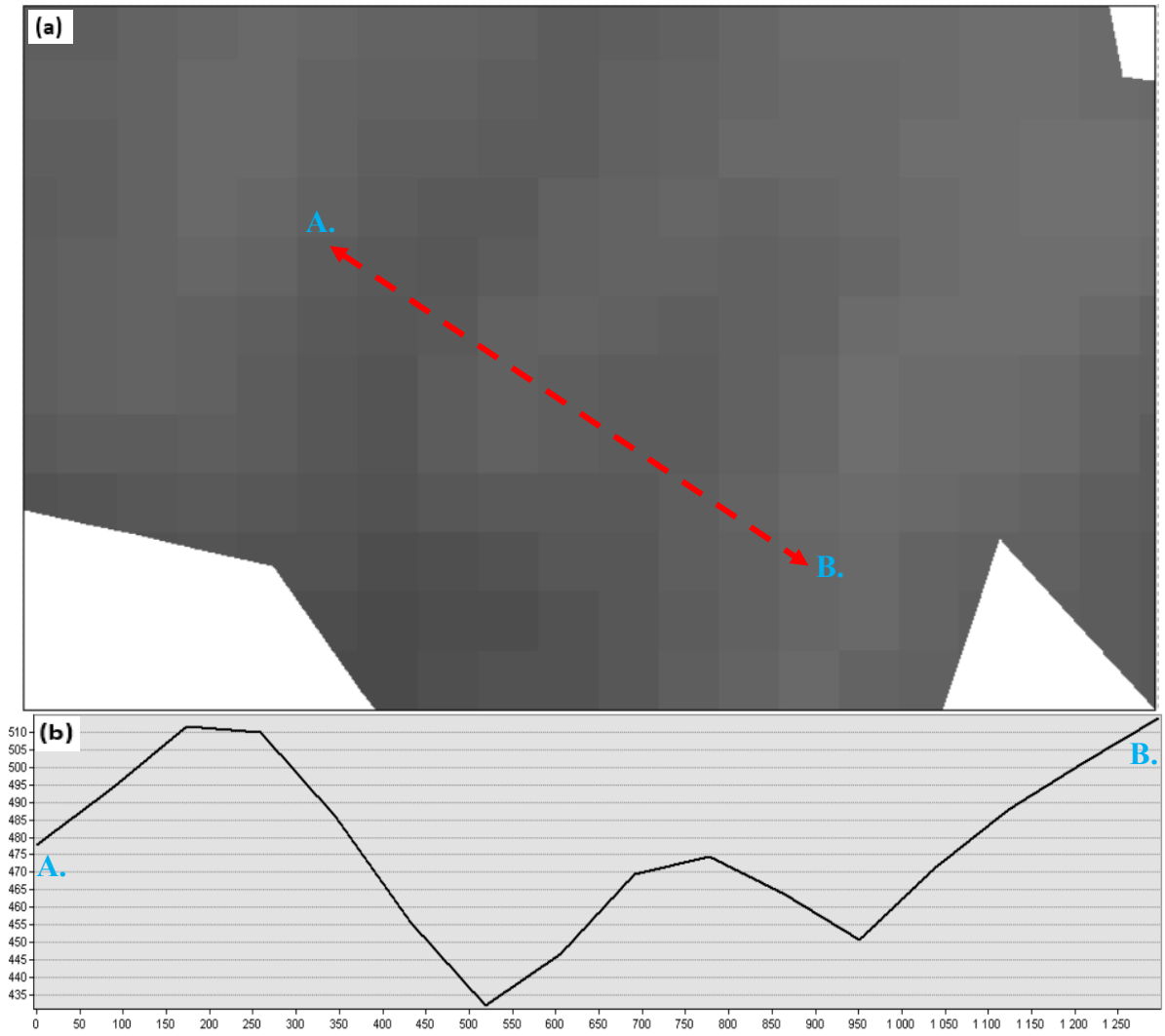


Figure 2.21: (a) 90 m SRTM DEM surface showing low-detail and generalised surface features. Note the generalised flattened peaks and midslope regions. The red dashed line represents the transect for the terrain profile graph, (b) the Terrain profile graph showing peaks and troughs through the transect across the DEM surface.

The findings comparing DEM surface profiles are promising since there appears to be some similarity between the LD and NLD *TST* values between the same DEM resolutions, e.g. the 5, 10, 20 DEM surfaces. These DEM resolutions all have similar *mean TST* values with low corresponding *SD* values. These results may suggest that *TST* can be detected across generalisation methods, albeit at selected DEM resolutions. Finally, it must be acknowledged that *TST* calculations are heavily reliant on *SlopGrad* estimations. I have addressed the relationship of DEM resolution and generalisation method on the *SlopGrad* estimates and further quantified the reduced surface accuracy with decreased pixel resolution. Therefore, the *TST* values displaying the coarse-resolution surfaces as more variable is expected and is a factor of the quality of the DEMs rather than an accurate representation of the terrain surface roughness, i.e., as DEM resolution decreases, the general accuracy of derived topographic attributes also decreases. Terrain surface roughness is often an overlooked topographic parameter despite studies suggesting it could be as effective as *SlopGrad* in classifying local landforms from DEMs (Iwahashi & Pike, 2007). *TST* is, therefore, an essential parameter for DSM and soil-landscape mapping since it can also provide a representation of soil surface roughness which can be a vital factor for a range of modelling applications at various scales, such as surface water flow and sediment/nutrient transport (Turner *et al.*, 2014). Undoubtedly, further detailed analysis is necessary to explore the scale-dependency of surface and *TST,* especially its influence on soil-landscape modelling.

This is necessary since the results have shown that the application of coarse DEMs to *TST* studies could result in measurement errors in the order of three times that of a "reference" surface. Considering these equivocal results for *TST,* it is clear that no meaningful conclusion can be established by comparing the *TST* texture across the various products. Each terrain analysis should consequently assess the utility of the DEM products in extracting and representing either micro-surface or landscape-scale *TST* values. Future approaches may need to examine the efficacy of third-directional cosine analysis, transform the data to fractal surfaces (Andrle & Abrahams, 1989), or even identify eigenvectors parallel to micro-and macro surfaces (McKean & Roering, 2004) to allow easier comparisons of surface products. Furthermore, given that the study area is a forestry plantation, the effect of vegetation in general on *TST* detection may require further independent research regarding local level calibration. The vegetation cover results on TST measurements' influence are inconclusive, with some sensors less affected than others (Mushkin & Gillespie, 2005). Additionally, using object-based classification methods to maximise internal homogeneity and external difference to partition the landscape surface may offer improved surface roughness classification results.

2.3.5.5 Terrain Ruggedness Index

The terrain ruggedness index (*TRI*) (Riley *et al.*, 1999) expresses terrain heterogeneity, quantitatively indicating how undulated, broken, rugged or dissected the landscape is and is an alternative indicator for representing land surface ruggedness. Typically, conventional descriptions of terrain character involve qualitative terms such as undulating, broken, rugged or dissected. *TRI*, therefore, provides users with a simple, yet effective, quantitative measure of terrain character by calculating the *mean* of the absolute differences in elevation between a focal cell and its eight surrounding cells and then quantifies the total elevation change across a 3×3 cell range (Riley *et al.*, 1999). However, *TRI* can be defined differently depending not only on the calculation approach used, i.e., based on an *SD* of *SlopGrad*, *SD* of *Elev*, *ProCurv* and *PlanCurv* (contour curvature), or some other measure of topographic texture. Regardless of how *TRI* is estimated, since the parameter refers to the variability in elevation within a defined radius, like *TST*, its application is susceptible to the selected resolution of the DEMs used, and the scale (ground distance) of landscape features intended to characterise. For instance, Thomas *et al.* (2014) found that *mean TRI* values of space-borne DEMs were far below values derived from higher resolution DEMs, except *TRI* values derived from SRTM products, which were similar to higher resolution DEM *TRI* values.

Moreover, Nellemann & Fry (1995) deduced that fine-scale *TRI* values best-described niche habitat characteristics, while coarse-scale *TRI* values best described large scale surface patterns. This would suggest the need for *a posteriori* approach of segmenting the landscape into general landforms (or altitude ranges) before conducting the *TRI* analysis may yield better surface ruggedness results. This should better regulate the *TRI* across areas with significant differences in contour densities (plateaus *vs* valley bottoms *vs* ridgeline areas), or even products that represent differences in spatial resolution, as is the case with this study. This will limit the likelihood of anomalous differences in surface ruggedness (Beasom *et al.*, 1983). That is, comparisons need to be limited to the respective pixel resolutions and compared across the range of generalised products and not only compared with the 1 m reference surface. In doing so, we see that the outcomes at each different DEM resolution are positively correlated between the generalisation methods with the minor difference between ruggedness values and sensor platforms (**Table 2.7**). Still, the DEM resolution and terrain ruggedness results follow a similar counterintuitive pattern as that of DEM resolution and the *TST* results, i.e., increased ruggedness index with decreased pixel resolution (**Figure 2.22**).

Although *TRI* and *TST* are independent terrain "character" indicators, they are also highly correlated terrain rugosity measures (Olaya, 2009; Jenness, 2013). For this reason, it is anticipated that the same scale-dependent limitations will define *TRI* estimates as *TST* describing the landscape heterogeneity for Braemar. Comparing the LD and NLD DEMs results for *TRI* and *TST* in **Table 2.3,** we see that this is true. The results clearly show an increase in *TRI* values as the DEM resolution decreases from 5 m to 90 m. Moreover, the Box and Whisker plots for *TRI* and *TST* (**Figure 2.7f**) exhibit a similar cyclic pattern for the same LD and NLD DEM products. The *TRI* values differ from *TST* because the 1 m LIDAR *TRI* value is uncharacteristically low. Even the LD-derived high-resolution (5-20 m) products across all generalisation methods have not approximated the "true" surface ruggedness. It is accepted that scale is pivotal in any estimation and representation of ruggedness and roughness; however, what has not been discussed is the influence of sampling density (inspection density) on the extraction of topographic properties. For a comprehensive review of the effect of sampling density on the accuracy of DEM derived terrain attributes, readers are referred to the studies by Li *et al.* (2004), Aguilar *et al.* (2005) and Hengl & Evans (2009).

Considering this study's results, two key factors may explain the observed trends obtained for *TRI* and *TST.* First is the influence of the inspection density and sample location of the FSD soil sampling data within the study site. The FSD data conforms to a specific set of survey criteria; namely, that soil samples are collected on a 150 m x 150 m sampling grid basis (approx. 1 point/ ha) and are limited to areas that are either already under timber or are identified as being suitable for future afforestation. Furthermore, the reference LiDAR DEM has a field of view of 1 m^2, yet the support size (Hengl & Evans, 2009), typically a fixed area or volume of land sampled, of the FSD data, is 1 point per 10 000 m^2 which is far greater than the 1 m LiDAR DEM pixel resolution. Stated differently, the FSD point sampling density is primed for evaluating general relief feature variation, while the high-resolution LiDAR DEM is better suited for detecting local topographic variation. Thus the limitation of using the FSD sample point locations to extract terrain parameters is the influence of point density on the accurate representation of terrain features using each FSD sample point location, especially when combined with higher resolution DEM surfaces.

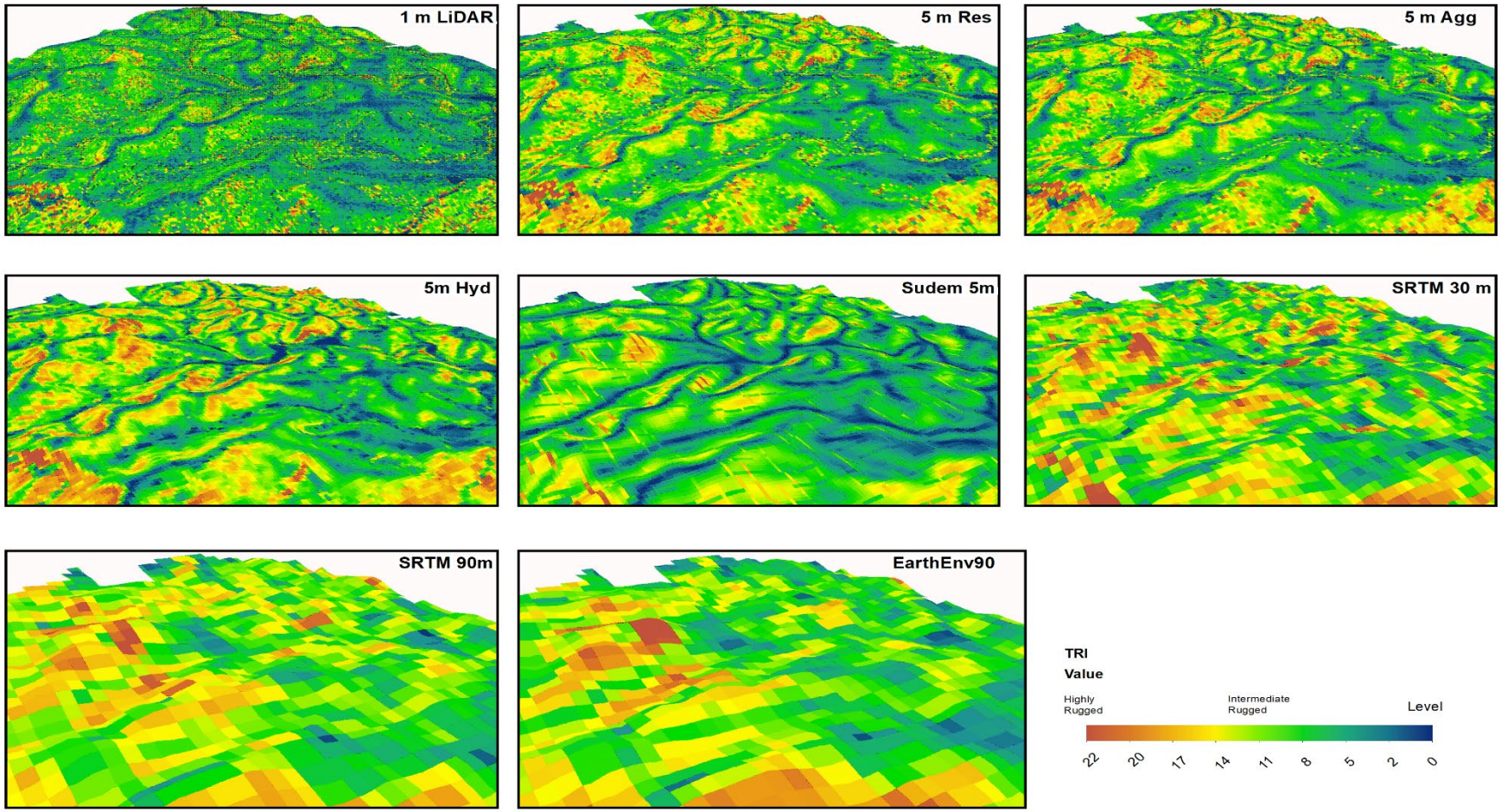


Figure 2.22: Representation of *TRI* calculated for Braemar. Note the high-detailed 1 m LiDAR and 5 m SUDEM *TRI* surfaces compared to the 30 and 90 m TRI surfaces' smooth and generalised surface.

Raaflaub & Collins (2006) reported that by using a fixed support size, gridded DEMs suffer from sampling problems. The terrain is usually under-sampled in rough (rapidly fluctuating elevation) areas and oversampled in flat areas. The user must find a compromise between the sampling interval and other requirements. Typically, the consensus is that the smaller the terrain variation, the smaller the sample size needed to achieve a given degree of accuracy required for accurate estimates. The dissimilarity between the support and cell size is less significant for DEMs' visualisation, which explains why it is often harder to detect significant variations between DEM surfaces when presented graphically (**Figure 2.22**). Still, it can be vital for validating simulation models or assessing uncertainty in elevation measurements. For practical reasons, users should always select a DEM cell size that matches the sample dataset's support size and avoid using a smaller cell size than the support size (Hengl & Evans, 2009). However, with LiDAR data, this is often impractical when analysing landscape-scale terrain variability. For this study, then, the high-resolution LD DEMs are considered to represent local relief features rather than global topographic features.

Hengl (2006) provides a pragmatic approach for selecting the correct pixel size based on the inherent properties of the input DEM data. He further defines and provides mathematical solutions for deriving three standard grid resolutions for DEM surface:

- *Coarsest legible grid resolution*: represents the largest resolution that users should use given a specific scale, positional accuracy, size of objects or/and complexity of actual terrain.
- *Finest legible grid resolution*: the smallest grid resolution that represents 95% of spatial objects or topography.
- *Recommended grid resolution* is a user-defined compromising resolution, usually the intermediate range between the coarsest and finest resolutions.

Secondly, as the distribution of data points is limited to planted areas, i.e., flat to moderately steep slopes, it is impossible to represent the complete hierarchy of terrain features or landforms within the study site. So, a complete representation of terrain ruggedness is limited. This has direct implications for thoroughly comparing the effectiveness of predicting *TRI* across generalisation methods and pixel resolutions. Unfortunately, both *TRI* and *TST* have not provided reliable information for concluding the effectiveness of sensor platform, DEM resolution and generalisation approach to deriving terrain indices of ruggedness. The scale dependency of *TRI* coupled with the fixed support space of the FSD soil data complicates the ability to establish generalised rules or guidelines for the tacit contribution of *TRI* to soil-landscape modelling. As previously discussed, optimising soil-landscape models requires identifying and adjusting the scale of terrain attributes to match the modelled soil-landscape process scale. This idea is a significant caveat to the present study and influences the interpretations of all the derived terrain attributes investigated in Braemar. While the FSD soil data is sampled at 1 per 10 000 m^2, which may seem coarse when compared to the 1 m^2 or even 25 m^2 of the 1 and 5 m LiDAR products, this is considered an industry standard for high-density soil surveying, suitable even for irrigation suitability analysis (Smith, 2006).

A higher sampling frequency of the FSD data for non-precision agricultural purposes would be impractical and costly for field-based soil survey evaluations. By benchmarking the LD and NLD DEMs' performance against the high resolution 1 m reference LiDAR, I have again assumed that the “reference” surface is the *defacto* gold standard for modelling and representing the terrain attributes investigated in this study. Put differently, the assumption, until proven otherwise, is that the better a digital surface can characterise the actual

ground surface, the better it is to predict soil-landscape processes. Maynard & Johnson (2014) points out that while this may be the case for specific landscape attributes, e.g., forest metrics or micro-topographic variability for precision agriculture, it remains uncertain whether the high horizontal resolution achievable with LiDAR is sufficiently beneficial in soil-landscape modelling. A broad range of studies has examined DEM resolution effects on soil-landscape model predictions in a range of ecosystem environments such as coastal dunes (Kim & Zheng, 2011), agricultural farmlands, and grazed hillslope environments (Park *et al.*, 2009). These studies have found that coarse resolution DEMs (between 30-50 m) provided the most robust predictions of soil properties.

Interestingly, recent studies have also shown that changes to terrain attributes due to changing DEM resolution is primarily a result of changes to neighbourhood extent rather than modifications to the DEM grid resolution (Smith *et al.*, 2006; Thompson *et al.*, 2012). It is not surprising then that in landscapes with a high degree of topographic complexity, and the smaller the grid resolution/neighbourhood extent, the less accurate the DEM can characterise larger scale terrain patterns, thus resulting in weak correlations. Similarly, coarsening the DEM resolution to match the spatial scale at which soil properties may well filter out micro-topographic variation not related to the property being modelled can also unintentionally disturb the accuracy of terrain information. According to Thompson *et al.* (2001), to adequately characterise local topography, the vertical precision must increase as the horizontal displacement increases such that the average change in elevation between grid points is more significant than that of their *vertical* accuracy. While this may provide a pragmatic solution for better representing DEM resolution and *TRI* properties, the expertise required for such calibration may limit its complete application.

For this reason, users may be more inclined to explore alternative measures of landscape classification, such as the topographic position index (*TPI*). *TPI* is calculated as the difference between the elevation at a cell and the average elevation in a neighbourhood surrounding that cell (Guisan *et al.*, 1999; Jenness, 2013). *TPI* is not a new concept but is a simplification of the landscape position index described by Fels & Zobel (1995) and later refined in detail by Weiss (2001). The *TPI* values provide a simple and powerful means to classify the landscape into morphological classes. The derivation involves evaluating elevation differences between a focal point and the mean elevation of the surrounding cells within a user-defined rectangle, annulus or circle (Prasannakumar *et al.*, 2011). The benefit is that users can classify the landscape into both topographic positions (i.e., ridge top, valley bottom, mid-slope, etc.) and landform categories (i.e., steep, narrow canyons, gentle valleys, plains, open slopes, mesas, etc.) (Jenness, 2007). Users can specify neighbourhood size and shape to capture the detail of terrain features in the landscape adequately. To detect small terrain features, users would select small neighbourhood sizes, and for extensive features, users would opt for larger neighbourhood radii. Notably, it is possible to perform *TPI* analysis in many GIS software packages with minimal parameterisation. Since *TPI* can be used as an indicator for terrain complexity, it also provides a fair assessment of terrain ruggedness. Recent studies by Flynn *et al.* (2019a) have again proven the usefulness of *TPI* measures for the semi-automatic disaggregation of national resource inventory datasets into farm-scale soil depth class maps.

2.3.5.6 Topographic wetness index

Beven & Kirkby (1979) were the first to attempt to integrate topographical information to capture hydrological variation at varying scales by introducing a framework known as the topographic wetness index (*TWI*). *TWI,* also known as the compound topographic index (Tagil & Jenness, 2008), is a second-order derivative of DEM elevation and can be computed as ***ln(a/tanβ)*** where ***a*** considers the specific upslope catchment drainage area

through a certain point per unit slope gradient **(*tanβ*)** (Sörensen *et al.*, 2006). Larger upslope drainage areas and shallower slopes will produce larger *TWI* values and are expected to have relatively higher water availability than locations with a small upslope and steep slopes, which would receive a low *TWI* indicating high trends of surface runoff (Quinn *et al.*, 1991). Consequently, higher positive values represent wetter areas in the landscape, while lower negative values represent drier *TWI* regions.

Thus, *TWI* has become a familiar and widely used indicator for the spatial distribution of hydrological conditions within the landscape, i.e., the potential occurrence of shallow groundwater or even soil moisture. This makes *TWI* estimates particularly attractive for a variety of DSM and soil-landscape analyses. *TWI* is usually calculated from gridded elevation data with the main differences in estimates related to how the accumulated upslope area is routed downwards, how valley-bottoms are represented, and which measure of *SlopGrad* is used (Sörensen *et al.*, 2006). The *TWI* is a well-studied indicator of soil pedogenesis and soil moisture distribution at different landscape positions (Xue-Rong *et al.*, 2002; Debella-Gilo & Etzelmüller, 2009) and even used in some study areas to predict the solum depth (Gessler *et al.*, 1995).

It is important to note that *TWI* is based on several assumptions: Firstly, *TWI* is assumed to be static and assumes that local *SlopGrad*, is an adequate proxy for sufficient downslope hydraulic gradient, which may not always be accurate in low relief terrain. Grabs *et al.* (2009) point out that the local *SlopGrad* may overestimate the downslope hydraulic gradient in flat terrain because of downslope water tables. Moreover, the *TWI* concept is less suited for representing hydrological properties in flat areas due to relatively undefined flow directions, which are more likely to change over time. Thirdly, the *TWI* indices assume that the *SlopGrad* of the ground surface represents the *SlopGrad* of the groundwater table. Which Grabs *et al.* (2009) have shown not to be the case. At the same time, soil hydraulic conductivity and precipitation are expected to be uniform over the studied landscape (Sorensen & Seibert, 2007). Therefore, *TWI* provides a relative and not absolute measure of soil moisture across the landscape and within a given pixel (Buchanan *et al.*, 2014).

For the *TWI* analysis, I used the standard SAGA wetness index (**Figure 2.22**) since the application allows for the adjustment of width and convergence of uni- to multidirectional hydraulic flow (Menezes *et al.*, 2013). I evaluated *TWI* differences between DEM resolution and generalisation methods using *mean TWI* and MAD values (**Table 2.3**) as suitable descriptors for comparing soil moisture results between various LD and NLD *TWI* products (Saulnier *et al.*, 1997; Buchanan *et al.*, 2014). Overall, the results show that the *mean TWI* values for the LD and NLD DEM surface products approximate the real surface moisture accumulation between 5 and 30 m DEM resolutions. Beyond 30 m DEM resolution, the *mean TWI* values for both LD and NLD DEMs increase almost twofold for 60 and 90 m resolutions for both the HCD and EarthEnv DEM. The increase in *mean TWI* with a decrease in DEM resolution suggests that the terrain wetness is being overestimated at coarse resolutions on average. Despite this trend, this study's *mean TWI* values are not very high and are in line with other studies across similar landscapes (Tagil & Jenness, 2008). Considering only the *mean TWI* values, it would appear that Braemar is well-to-moderately drained, e.g. well-drained < 7.7 < poorly drained (Giasson *et al.*, 2006). The fact that the results define the area as well-to-moderately drained comes as no surprise considering that FSD sample points used for this study were limited to survey locations restricted to higher altitude areas more suited for forestry. The FSD soil data intentionally limits the landscape's lower-lying, wetter areas, such as riparian zones or catchment drainage zones, due to conservation reasons, poor soils, or accessibility.

The *TWI* results for the HCD DEM surface have performed exceptionally well at high resolution. The 5 m and 10 m *TWI* products have the lowest *SD* and MAD values compared to the NNR and MA generalisation methods. This result is not surprising given that we would expect the hydrologically conditioned DEM to

accurately represent topographic drainage patterns such as flow accumulation and flow direction within the landscape. This indicates that while the HCD products intentionally eliminate surface artefacts such as sinks or depressions, they do not necessarily eliminate all zones of water accumulation or flow direction, for that matter. The HCD surfaces are, in fact, the best suited for identifying hydrological flow paths within the landscape. However, since *TWI* is highly dependent on *SlopGrad*, and the HCD slope surfaces did not perform particularly well across the various pixel resolutions, I did not expect the *TWI* to be so well represented. If *SlopGrad* is poorly defined, then it is likely that higher-order derivatives, like *TWI*, will also be poorly described (Shi *et al.*, 2012). However, Wilson (2012) was able to show that DEM errors propagated strongly to *SlopGrad* but only moderately to *TWI* estimates, although the coefficient of variation for *TWI* values varied more spatially than that of *SlopGrad*. Interestingly, his results confirm that *TWI* values are less sensitive than *SlopGrad* to the input DEM but cautioned that the choice of flow routing algorithm, rather than DEM resolution, may significantly influence the upslope calculation of contributing areas.

These results further highlight that at specific DEM resolutions, namely 5, 10, 20, 30 and 60 m, the generalisation approach does not significantly influence *TWI* predictions with minor discrepancies of *SD* between the three generalisation methods. This is also true for the NLD DEMs at 5 and 30 m, respectively. However, note that as the DEM resolution decreases to 90 m, so does the *TWI* accuracy, represented by high *SD*, for each DEM surface, indicating that we can be less confident of the *TWI* estimates at coarser resolutions. There is a definite pattern of decreased *TWI* prediction accuracy as pixel resolution decreased in both LD and NLD products. This is expected given that *TWI* relies on predicting the hydrological flows from the upslope area, which depends on pixel resolution. Increased prediction error in upslope area estimation will result in less accurate *TWI* predictions (Sörensen *et al.*, 2006).

Thomas *et al.* (2017) point out that coarser DEM resolutions are more likely to produce narrower slope distributions and lower *mean SlopGrad*, due to the loss of topographic detail resulting from surface smoothing, with lower *SlopGrad* on steeper slopes and higher gradients on shallower slopes (Chang & Tsai, 1991; Thompson *et al.*, 2001). Furthermore, larger *mean TWI* values are also expected due to less irregular flow paths and larger minimum areal units, with larger values in upper landscape positions and lower values in lower landscape positions (Thompson *et al.*, 2001; Wu *et al.*, 2008). These studies concluded that larger *mean TWI* values were expected from coarser-resolution DEMs, primarily due to the influence of catchment area distributions rather than *SlopGrad* (Sorensen & Seibert, 2007; Gillin *et al.*, 2015). The upslope area calculation depends on the method used to calculate the accumulated area of upstream, routed to downstream cells, and the study's spatial context. A key outcome is that as *mean TWI* increases, the spatial pattern of standard deviation and so too residual error also increased across generalisation methods and compounded by increasing decreasing pixel resolution.

Conventionally, the area estimation from a cell has been transferred in the steepest downslope direction to one of the eight neighbouring cells (D8 flow routing flow direction algorithm) (Sörensen *et al.*, 2006). This neighbourhood approach is computationally similar to both the NNR and MA approaches. This could explain why the *TWI* products from these interpolation approaches performed particularly well despite not being hydrologically corrected, i.e., DEM surface remained unfilled.

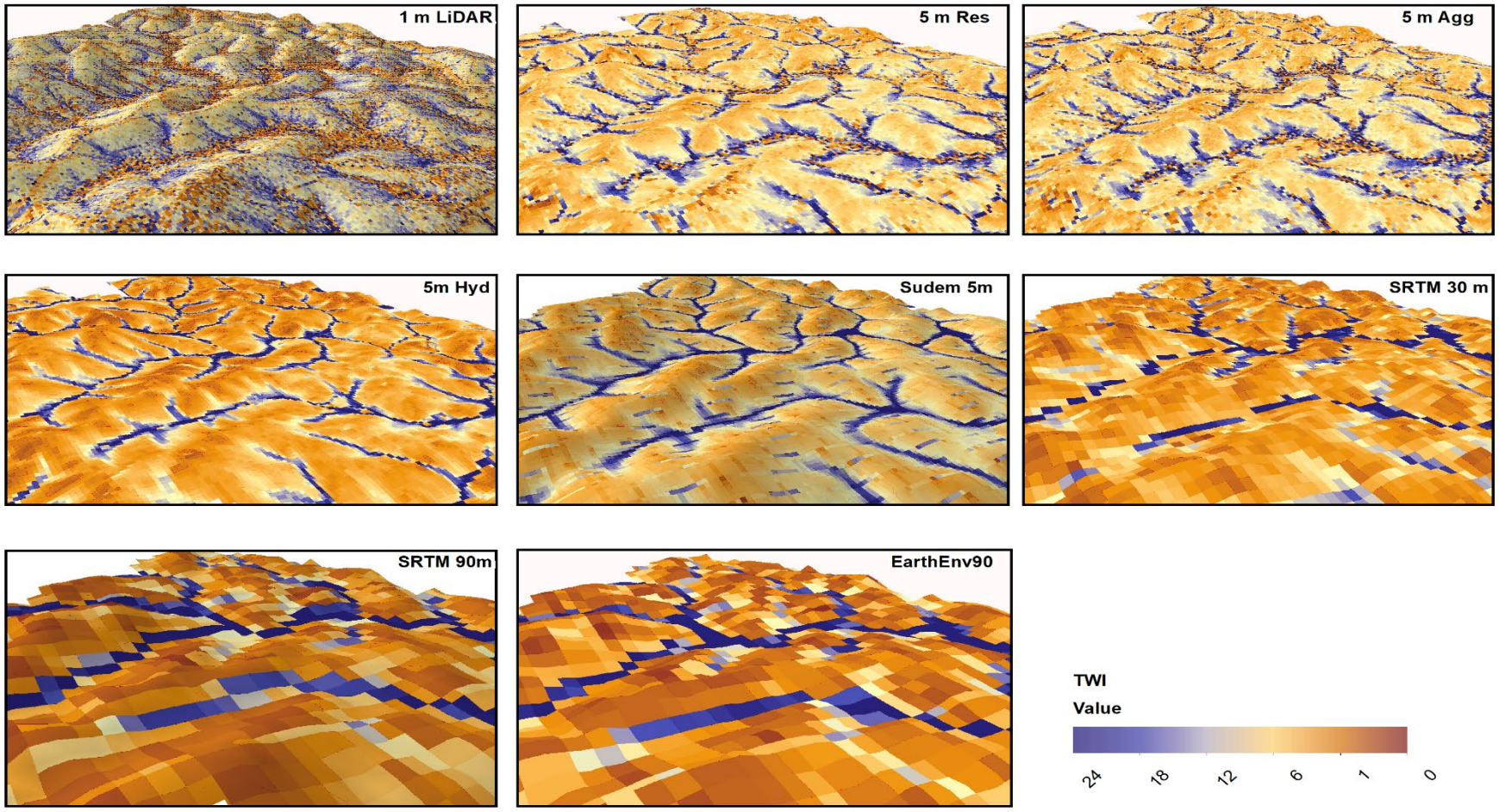


Figure 2.23: Representation of TWI calculated for Braemar. Note the high-detailed 1 m LiDAR and 5 m SUDEM TWI surfaces compared to the 30 and 90 m TWI surfaces' smooth and generalised surface.

Therefore, accurate inspection of *TWI* requires special consideration of sensor type or pixel resolution and a clear understanding of what flow algorithms best suit the study objective. Let us consider the ANOVA and HSD results for *TWI* presented in **Table 2.6** and **Table 2.7** across the various DEM resolutions and generalisation methods. Several key observations can be made: firstly, the LD and NLD DEM surface with resolutions greater than 20 m differed significantly from the 1 m actual surface model at almost all altitude ranges. This suggests a distinctive altitudinal influence on the estimation of *TWI* in the landscape with an intermittent clustering of *TWI* values in the lower and higher altitude regions for selected generalisation methods and pixel thresholds, i.e., with 200 to 550 m elevation range and 5 to 20 m pixel resolution providing the most statistically significant comparisons to the actual surface model. These results are consistent with those of Hancock (2005), who also found *mean TWI* values to be sensitive to pixel resolution changes greater than 20 m with decreased correlations in accuracy in resampled DEMs up to 60 m. This is mainly due to the influence that varying pixel resolutions have on the behaviour of the area–slope relationship and the cumulative area distributions within the study site. The results demonstrate that care must be taken when extracting catchment information from a digital elevation model and that the reliability of the data must be related to grid size.

Simply put, local slope influences hydrological variables. The *TWI's* ability to predict observed patterns of saturated areas is sensitive to the algorithms used to calculate up-slope contributing area and slope gradient. The coarser-resolution products capture micro-topographic features in lower detail. Therefore, they cannot statistically replicate the flow accumulations and diversions characterised by their high-resolution counterparts at lower altitudes. The application of coarse DEMs to capture or adequately represent topographic features may lead to inaccurate topographic indices predictions due to surface generalisations of the terrain (Hancock, 2005; Vaze *et al.*, 2010).

Outside of the altitudinal influences, the *TWI* results based on DEM resolution and generalisation method should not be entirely dismissed as inaccurate since descriptive statistic results in **Table 2.4** at the landscape, or "global" scale, confirm many similarities with the reference surface at selected resolutions. Consequently, the consideration of *TWI* variables at resolutions above 20 m needs to be reserved for applications geared more for landscape or regional assessments of hydrological significance. *TWI's* visual impression in **Figure 2.23** highlights the difference between the 1 m LiDAR surface and the generalised LD and NLD *TWI* surfaces. The 1 m LiDAR *TWI* displays significant terrain and slope detail. However, the wetness index representation is far too "noisy" for any practical synoptic application - *TWI* surface(s) provides too much detail and redundancy in representing surface flow across the landscape. A discussion provided by Kuo *et al.* (1999) solidifies these results by suggesting that very high DEM resolution data may not be appropriate for *TWI* mapping in many topographic settings. The 5 m SUDEM delivers perhaps the best general representation of flow diversion while the 90 m products still adequately represent the general accumulation zones of saturation in the landscape. Likewise, the 5, 30, and even 90 m NLD *TWI* surfaces provide reasonable interpretations of high and low moisture accumulation regions within Braemar from a landscape perspective. Studies that evaluated groundwater flow directions were generally better represented by landscape topography rather than small-scale surface variations. Therefore smoother topographic surfaces, defined by coarser resolutions, are better suited to represent near-surface flow pathways and water table positions (Gillin *et al.*, 2015).

Similarly to Sörensen *et al.* (2006), the results seem to demonstrate further that different methods of calculating the *TWI* indeed produce a high variation in wetness index values between the various DEM surfaces with not one single method performing optimally across the full range of resolutions and altitude ranges. Haas (2010) provides a detailed review of studies that have evaluated the vertical and horizontal influence on *TWI* value

extraction within different landscapes. Therefore, the optimal DEM resolution for deriving *TWI* depends on the hydrological processes being modelled, the scale of the topographic features controlling them (Quinn *et al.*, 1991) and the DEM resolution for representing slope-area variables. Therefore, user considerations should focus on balancing appropriate levels of topographic accuracy, data processing, and storage requirements combined with the need for interpretable outputs (MacMillan *et al.*, 2003; Hengl, 2006; Liu, 2008).

2.3.5.7 Curvature

The final topographic variables assessed were planform (*PlanCurv)* and profile (*ProCurv)* curvatures. Curvature surface results for the LD and NLD DEMs were extracted in ArcMap based on Zevenbergen & Thorne (1987). *ProCurve* can be interpreted as the slope shape in the maximum slope's direction and predominantly governs water and sediment flow across the landscape. At the same time, *PlanCurv* is the slope shape perpendicular to the slope direction and measures convergence or divergence of water in the landscape (Moore *et al.*, 1993). Using both *PlanCurv* and *ProCurv* to classify the shape of the slope allows users to define the terrain into convex (erosional), concave (depositional) or flat (straight or planar features). For *PlanCurv*, a positive value indicates the divergence of surface flow and a negative value of the flow concentration.

For *ProCurv*, positive values indicate the concave profile and negative values the convex profile. Curvature values, therefore, do not show rates of erosions but rather indicate areas of potential erosion or deposition by identifying regions of terrain convexity and concavity (Prasannakumar *et al.*, 2011). For instance, negative values for *Plan* and *ProCurv* values indicate areas with greater potential for erosion. It is expected then that surface runoff tends to accelerate in areas with negative *ProCurve* values and converge in areas with negative *PlanCurv* values (Mueller & Sassenrath, 2015). Curvature values are thus suitable indicators of soil wetness and soil drainage class (Troeh, 1964), known to affect soil development and morphology (McBratney *et al.*, 2003; Lagacherie, 2008). *PlanCurv* is also sometimes used to describe the curvature of contour lines and should yield similar results to tangential (horizontal) curvature provided the contour lines represent the shape of the land surface (Wilson & Gallant, 2000). Readers are referred to Olaya (2009) for further detailed descriptions of several other curvatures and their potential significance in soil-landscape applications.

Unsurprisingly, curvature plays a significant role in holistically and quantitatively evaluating soil-landscape relationships using key terrain attributes. Soil surveys rely on relief properties and criteria necessary for characterising soil mapping units. Accordingly, using DEM derived properties, like curvature, as landscape differentiating criteria can generate satisfactory soil characterisation results, particularly in reconnaissance soil surveys (Dobos & Montanarella, 2006). In particular, the spatial distribution of soil moisture content is strongly related to variables such as slope gradient, slope aspect, curvature, and topographic position (Florinsky, 2016), with both *Pro* and *PlanCurv,* regarded as crucial topographic factors that determine overland and intra-soil water dynamics (Kirkby & Chorley, 1967). Still, while curvature values calculated from DEMs are considered critical parameters for DSM and soil-landscape mapping, the accuracy of estimates are heavily dependent on the spatial resolution of the DEM, neighbourhood size or search radius of the various curvature algorithms, as well as the size of the area under consideration (Smith *et al.*, 2006; Wu *et al.*, 2008). Terrain curvature estimates are considered suitable predictors of soil development. They affect the spatial distribution of soil properties influenced by the flux of water and materials horizontally from upslope areas or vertically from surface soil horizons (Maynard & Johnson, 2014). Given that the mathematical calculation of curvature is derived in units of radians, values are generally small, and several authors have suggested reporting the values as either radians/ 100 m or using a 1/100 m scale to allow for more practical interpretations of the results (Zevenbergen &

Thorne, 1987; Reuter, 2004). For this study, I evaluated curvature values using mm^{-2} units (Thompson *et al.*, 2001).

The curvature results extracted for Braemar were the most difficult to interpret and were perhaps the most sensitive to DEM resolution and neighbourhood size (Roecker, 2013). Since curvatures values can take on any real number value $(\mathbb{R}; +; -; <)$ including integer values, the ANOVA analysis and correlation assessment were not done on the curvature surface products. Mitášová & Hofierka (1993) have suggested that *mean Plan/ProCurv* (Hengl & Rossiter, 2003; Li *et al.*, 2004) results are suitable for describing surface curvature variance. However, as the results show, *mean* curvatures for Braemar may not convey enough information about the positional variables of accumulation or deposition other than indicating that the surface is generally convex or concave at a central point and provides little information about the local variability related to abrupt topographic change. It must be pointed out that the assessment of DEM surface curvature, like the other terrain attributes extracted, has been evaluated by comparing the descriptive statistics for each generalisation, sensor, and resolution combination. Thus, *mean*, *min*, *max* and even MAD measures are therefore purely intended to be interpreted as statistical measures describing the distribution of the FSD soil-point observations (n= 1 168). Specific morphometric variables such as curvature can be further mathematically expressed to define explicit measures of additional curvature derivatives such as *mean* curvature that describes mean-concave and mean-convex terrains (Olaya, 2009). Additionally, minimal curvature, which represents the curvature of a regular section with the lowest value of curvature among all regular areas of the topographic surface at the given point and maximum curvature that represents a normal section with the highest value of curvature among all normal sections of the topographic surface at the given point (Shary *et al.*, 2002; Florinsky, 2016). Thus, these models are often combined, i.e., maximum and minimum curvature, to provide a comprehensive interpretation of surface curvature. Therefore, this study's curvature results should not be interpreted and compared in the context of these types of curvature methods. A complete discussion of these and alternative curvature measures is provided by Shary (1995).

Studies have shown that curvature values may poorly represent the land surface unless calculated over a neighbourhood size and pixel resolution commensurate with the extent of the landforms they are expected to characterise (Roecker, 2013). *ProCurv* values are typically positive for convex landforms (e.g. hills and ridges) and negative for concave features (e.g. depressions and valleys), and linear for zero values. Similarly, negative *PlanCurv* values suggest a convergent landscape with positive values indicating a diverging slope with zero values related to linear or flat curvature surfaces (Jenness, 2013; Florinsky, 2016) (**Figure 2.24)**.

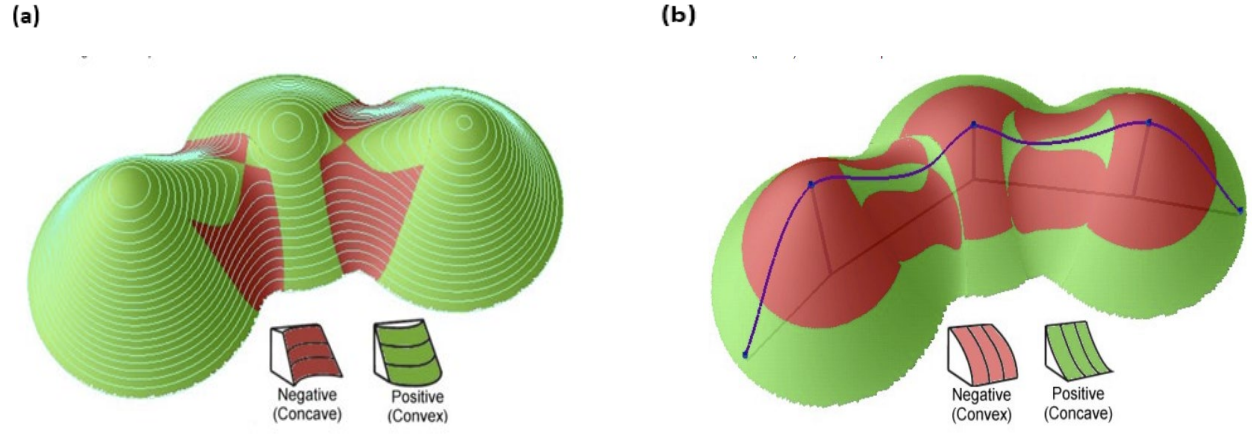


Figure 2.24: (a) *PlanCurv* is the curvature in a horizontal plane. It can be described as the curvature of the hypothetical contour line that passes through a specific cell. *PlanCurv* is positive for cells with laterally convex contours and negative for cells with concave contours. *PlanCurv* can be used to differentiate between ridges and valleys. (b) *ProCurv* is the curvature of the surface in the steepest slope's direction (in the vertical plane of a flow line). It affects the acceleration and deceleration of flow across the surface. *ProCurv* affects the flow velocity of water, draining the surface and influencing erosion and deposition. In locations with negative *ProCurv*, erosion will prevail, and in locations with concave (positive) curvature, the deposition will dominate (Florinsky *et al.*, 2002).

The generally low *mean Plan*/*ProCurv* values observed for this study have been observed in similar studies Mora (2015), suggesting a slightly convex, diverging linear (flat) landscape. These results make sense since the samples used for DEM data extraction were limited to suitable areas for commercial timber afforestation with steep slopes and valley-bottom regions most likely not sampled due to their limitation on timber production and reserved for conservation importance. Moreover, the *Plan/ProCurv mean* values for most DEM surfaces were close to zero, indicating similar concavity and convexity levels across platforms within the study area. Notwithstanding, the 90 m HCD P*lanCurv* product had a *mean* value of -1.43 m m^{-2}, suggesting that the terrain surface predominantly comprises converging land surface features. This result is entirely expected, given that the HCD interpolation approach transforms the terrain surface to account for flow accumulation and direction. This generates a surface that accounts for hydrological surface flow. However, I cannot comment on why this result manifested exclusively at 90 m and not at the other finer resolution HCD products or even coarse NNR and MA generalisation approaches. This result is similar to that observed by Kienzle (2004), who showed that coarse resolution DEMs (> 90 m) significantly affected calculations of overland flow and divergence or convergence and subsequent erosion or deposition simulations for detailed hydrological and environmental terrain analyses. Notably, the curvature results derived from the HCD may also be affected by the pit-removal algorithm used to derive the hydrological flow surface. Recently, Wang *et al.* (2019) were able to show that DEMs resulting from the application of a depression-carving algorithm for removing spurious artefacts in the DEM surface might not reasonably calculate topographic attributes (such as slope gradient, curvatures, specific contributing area, and topographic wetness index) on those intact cells (especially those neighbouring the carved cells) in depressions. Their results suggest that errors in HCD methods are limited to void-filled areas and the neighbouring cells.

Terrain attributes respond differently to resolution change, especially when the resolution is coarsened, i.e., 5→50 m, with plan and profile curvatures being the most sensitive among these attributes (Zhou *et al.*, 2008). This study has confirmed, and corroborated by other authors, that coarse-resolution products produce lower slope gradients on steeper slopes, steeper slope gradients on flatter slopes, narrower ranges in curvatures, and larger specific catchment areas in upper landscape positions and lower curvature catchment areas. Concomitantly, in lower landscape positions, flow-path lengths tend to decrease, and the accuracy of terrain attributes, including curvature, tends to decrease (Chang & Tsai, 1991; Gao, 1997; Goyal *et al.*, 1998; Chaplot *et al.*, 2000; Schoorl *et al.*, 2000; Wilson & Gallant, 2000; Thompson *et al.*, 2001; McMaster, 2002). This is mainly attributed to topography's smoothing, resulting in larger grid sizes, causing terrain detail, such as shorter, steeper slopes, to be lost. This smoothing effect also produces narrower ranges in slope curvatures (Wolock & Price, 1994; Zhang & Montgomery, 1994). The impact of underestimating *PlanCurv* or *ProfCurv* can lead to the misinterpretation of dispersion and convergence areas in the landscape, particularly for erosion, sedimentation, and hydrological analysis (Kienzle, 2004). In an effort to better understand the variability associated with pixel resolution and sensor platform, I have considered alternative descriptive statistics to the *mean Curv* values such as *SD, min, and max Curv values*, which might offer better insight into the variability and representation of slope shape and direction across the site. With that in mind, the *SD* results exhibit a linear

decrease in deviation as pixel resolution decreases from 1 m to 90 m across all sensor platforms for both *Plan and ProfCurv* values.

As DEM resolution decreases, the distribution of *Curv* values is, as expected, more concentrated to intermediate *Curve* values than to extreme values (Sorensen & Seibert, 2007) with lower overall *Curv* variability. A more pragmatic way to interpret the results is that steeper slopes and a broader curvature range are observed with higher resolution DEMs, representing higher detailed surface features (Thompson *et al.*, 2001). Areas of high *SD* of *Curv* values are typically found along steeper rocky slopes and linear features such as roads, stream channels, and slope breaks (Brubaker *et al.*, 2013).

Interestingly, if we compare these results with that of the *SlopGrad* values, we then observe a similar trend of decreasing curvature *SD* values with decreased pixel resolution (60-90 m). Similar results were reported by Evans *et al.* (2009). They observed that while *mean* surface curvature values varied predictably, the *SD* of surface curvature, *ProCurv* particularly, depended on *SlopGrad*. *PlanCurv* values generated from both the SRTM and SUDEM products showed similar trends in *mean* curvature values to the reference surface with low overall *SD* values suggesting that these NLD and low-resolution products can represent the general planform curvature of the site. Sharma *et al.* (2011) investigated the effects of DEM cell size and terrain complexity, and entropy for hydrological modelling. They reported that the *mean* local entropy of DEM at different resolutions follow a similar trend as that of global entropy, i.e., decrease in entropy with increasing cell size for all the terrain complexity classes. However, like this study's result, their findings for the normalised terrain entropy showed a reverse trend, i.e., increase in information content with increasing cell size, indicating that their 90 m terrain products correspond to the original 20 m contour line's information content data used for interpolation.

This study further demonstrates that while the *mean* curvature values may be small, the *max & min* curvatures values exhibit extreme values for higher resolution surface curvature products. These extreme positive and negative values serve as an excellent proxy to define the rugosity of the terrain. Considering the positive and negative *ProCurv* values observed, they may indicate abrupt changes in *PlanCurv* terrain surface, possibly at a localised scale, and the occurrence of depositional and erosional characters (Prasannakumar *et al.*, 2011). The results seem to agree somewhat with those of A-Xing *et al.* (2008), who concluded that both *ProCurv* and contour curvature is more sensitive to neighbourhood size (window size) than *SlopGrad*, and sensitivity is much higher at high-resolution pixel sizes than at low resolutions. It is clear then that these extreme values are limited in occurrence and may represent features such as road cuttings, banks, ridgelines, contour banks, or even depressions that are not representative of the landscape's general "regional" curvature. They have been detected since it is possible to detect abrupt changes in the micro-relief at higher pixel resolutions (Brubaker *et al.*, 2013).

Notwithstanding, these results also need to be considered with the possible presence of typical DEM analysis phenomena, such as GiBB-like-phenomena (Gousie & Franklin, 2005), further amplified through DEM generalisation. The Gibbs' phenomenon is particularly a factor for consideration where elevation discontinuities or sharp elevation changes occur, resulting in artificial depressions or peaks near such locations. These artefacts manifest as over and undershoots around an elevation discontinuity and other artefacts observed in grids interpolated from contour DEMs due to gaps in slope gradient. They are (conditionally) referred to as 'Gibbs-like phenomena'(Zhou *et al.*, 2008).

Therefore the use of interpolation techniques influences the accuracy of a DEM. Florinsky (2016) point out that all DEM surfaces are exposed to GiBB phenomena since all DEMs contain jump discontinuities of the

elevation function due to interpolation. These include sharp elevation changes (or steep gradients), such as terraces, escarpments, abrupt slopes, peaks, and pits. GiBB phenomena also occur near pronounced systematic and random DEM errors (Goodchild & Mark, 1987). For instance, near an abrupt linear step in elevations resulting from processing orthophoto imagery as separate patches and subsequent joining them to assemble a DEM, or at a point with a false elevation value of 100 m within an area with an average altitude of 10 m as may be the case with coarse-resolution DEM data. DEM errors caused by the Gibbs phenomenon can propagate through DEMs' processing and produce new errors in "secondary" or generalised terrain surfaces derived from a DEM (Florinsky, 2016). In general, second-order derivatives, such as curvature values, are susceptible to potential errors in the source DEMs since they are more inclined to highlight local terrain irregularities (Amatulli *et al.*, 2018).

The MAD *Plan* and *ProfCurv* values (**Table 2.4**), also referred to as difference curvature (Florinsky, 2016), exhibit interesting properties for all the derived DEM surfaces. Specifically, we observe that *PlanCurv*, MAD values are centred between 27 and 28 m m^{-2}. *ProCurv* has values centred on 18 and 19 m m^{-2} respectively for all sensor platforms and pixel resolutions. Interestingly all sensor platforms exhibit almost identical MAD *Plan/ProfCurv* values regardless of DEM resolution. Realistically this is unlikely. However, the scale-generalisation dichotomy inherently influences the uniform detection of certain terrain features across the site within each resolution. These results highlight that specific topographic attributes are linear combinations of others. For instance, *mean* curvature values combine horizontal and vertical curvatures (Florinsky, 2016). I was unable to identify any supporting evidence to corroborate these findings but offers the following plausible interpretation: It is expected that DEM generalisation may increase the absolute values of correlation coefficients between related terrain variables or phenomena owing to the removal of minor variations. Here again, the statistical trends for curvature follow those of elevation data (§ *2.3.1*) when compared at a global scale. Based on the descriptive statistics, MAD does not have the desired outcome when averaged and presented for regional areas since it does not adequately detect multi-scale and multi-resolution variability. Instead, the results indicate that MAD across sensors remains consistent for regular terrain features and divergent and convergent slopes regardless of sensor resolution.

Furthermore, the computation of the MAD statistic (**Equation 2.2**) defines the absolute deviation of values from the central *mean* value. For the curvatures, MAD values are reported as a standard set of values for all DEM surfaces. This would imply that mathematically all DEM surfaces exhibit the same magnitude of deviation from the reference surface at similar regions in the study area and for the same terrain features, i.e., prominent hilltops, crests, ridges, or divergent slopes. Since curvature is understood to be the second-order derivative of *SlopGrad,* which itself is a direct derivative of elevation*,* each DEM surface, regardless of resolution, can generally detect significant variations and general trends across the landscape (the high correlation between DEM products confirms this). Thompson *et al.* (2001) reported similar findings when comparing the predictive performance of terrain attributes for soil mapping of two DEMs at resolutions of 10 and 30 m. Their results indicate that while the 30 m terrain attributes were derived from the 10 m DEM source, the predictive performance of each for A-horizon depth was similar, regardless of DEM resolution.

The semantic interpretation of terrain complexity has found success in curvature application as a metric used to depict topographic turbulence and roughness (Shary *et al.*, 2002). Huaxing (2008) elaborates on this complexity concept by stating that since curvature is the derivative of specific points on the mathematical curve used to depict how fast the angle of a tangent moves over a given arc to the length of the arc; we can then consider *SlopAsp* and shape complexity as being equivalent to the *PlanCurv.* Similarly, *SlopGrad* and slope change are then equal to the *ProCurv,* where the only difference between these geometrical indices is direction. It is necessary to understand that issues such as curvature, and terrain complexity for that matter, can be considered in the vertical and horizontal direction, i.e., local relief and local standard deviation are supposed

to define characteristics of vertical complexity, while features such as among other things *SlopAsp*, determine horizontal complexity. Importantly, these complexity characteristics are independent. Total curvature values may be large even if local relief is low, explaining the high observed MAD curvature values irrespective of DEM resolution.

What is promising is that these results agree with similar studies comparing DEM resolutions and curvature trends. Specifically, the results have shown an overall decreasing trend in curvature values as DEM resolution decreases, i.e., as DEM grid spacing increases. This is particularly useful for future studies of the coastal region of KwaZulu-Natal, aiming to compare curvatures values (*max* curvature) extracted for different DEM resolutions. Here *max* curvature defines the maximum curvature at a location within the landscape. Negative values indicate concave shapes, and positive values indicate convex shapes (Shi, 2010), and $Curv_{max} > 0$ typically corresponds to ridges rather than depressions. The mainly decreasing positive curvature values with increasing grid spacing is a concept that has been described in detail by Zevenbergen & Thorne (1987), Chang & Tsai (1991), Shary *et al.* (2002) and Florinsky (2016). These studies reported that an increase in DEM grid spacing transforms small steep zones to broad smoothed-surface areas, ultimately "flattening" the terrain and reducing the curvature estimations. Thompson *et al.* (2001) found that decreasing the horizontal resolution of a DEM from 10 to 30 m resulted in a smoother, less defined landscape with more moderate slope gradients and reduced curvatures. Simiarly, Chang & Tsai (1991) reported that curvature error values tend to increase with DEM generalisation and decreasing spatial resolution, and Carter (1992) found that curvature values tend to correspond more to a "reference" value as grid spacing increases. In this study, however, the coarse-resolution DEMs do not approximate the reference surface. In fact, none of the LD and NLD DEM surfaces can accurately represent the large *max PlanCurv* values, 561.6 m m^{-2}, of the 1 m reference LiDAR. In contrast, the next highest *PlanCurv* values recorded are for the 5 m HCD at 327.86 m m^{-2} and 5 m NNR DEM 137.33 m $m^{-2,}$ respectively (**Table 2.4**). These results suggest that the evaluation of local curvature variables cannot be adequately studied simply by a comparison of calculated and "reference" values using the LiDAR DEM surface since it is unreasonable to suggest that a high-resolution `reference' surface is the correct one since it too is only an approximation to the real-world surface (Florinsky, 1998b).

Furthermore, since altitude classes did not further segment curvature into statistically meaningful estimates, the *mean* curvature results are expected to represent regional rather than local curvature deviation within the study site. This may also explain the unusual curvature results between the various LD and NLD DEM surfaces, especially at finer DEM resolutions. The curvature results suggest that it is difficult to outright regard the high-resolution reference LiDAR curvature values as the most accurate and therefore disregard the utility of the subsequent coarse-resolution DEMs. Furthermore, the accuracy comparisons between DEM products has not been explicitly limited to local or regional extents but instead defined as a generalised comparison of derived DEM terrain attributes with a modelled reference surface. Grunwald (2016) points out that it is vital to have a DEM with sufficient detail to characterise the topographic variability that impacts soil formation in terms of spatial resolution. Therefore, the association between soil and landscape terrain is best addressed at resolutions < 100 m. Fine-resolution DEMs are expected to provide a more accurate representation of the terrain's shape and be justified by mechanistic soil formation models for local terrain attributes such as slope, aspect, and curvature.

In contrast, when using data with resolutions > 100 m, the utility of local terrain attributes may not be relevant, especially in physiographical problematic areas (Grunwald, 2016). For instance, curvature derived at coarse resolutions are associated with broad-scale relief patterns and should not be compared to fine-scale reference products. Topographic features more considerable than the derived DEM resolution but smaller than the neighbourhood size will be suppressed and smoothened during the topographic attribute computation (Smith *et al.*, 2006; Shafique *et al.*, 2011). The variability in the observed curvature statistics depends on the spatial

extent that these variables are calculated, which varies with DEM resolution causing all topographic attributes to change accordingly. In this case, coarse resolution DEM attributes may provide better information related to landscape position. This is useful for defining meaningful soil-landscape attribute relationships. Whether factual or anomalous, the curvature results suggest that higher resolution DEMs may not be necessary for generating applicable soil-landscape models.

Thompson *et al.* (2001) comment that DEM resolution, precision, and source issues will become increasingly important with the further development and implementation of broad-scale or regional soil-landscape models, where access to detailed topographic field surveys is not feasible. However, the values of curvature measures such as MAD and SD for terrain interpretations need further investigation for broader applicability towards the pedometric agenda. Readers are referred to Grohmann *et al.* (2010) and Florinsky (2016) for a continued discussion on the performance of various statistical methods for representing local DEM attribute metrics.

2.4 CONCLUSION

Digital elevation model data are elemental in deriving primary topographic attributes that are input variables to various regional soil-landscape models. DEMs' utility to extract different topographic indices as primary inputs to various DSM applications allows the generalised soil-formative relationship between topography and soil characteristics to be measured quantitatively. Traditional landscape-scale approaches to extracting and analysing soils remain subjective and an expensive last resort for large-scale regional soil distribution and variability prediction. Therefore, the ability to rapidly and objectively represent soil-landscape relationships between soil properties and landscape position using emerging technologies and elevation data in a digital environment and at varying scales is fundamental for using soil-landscape mapping as a regional planning tool. Indeed, much progress has been achieved in outlining that a DEM's quality directly impacts the quality of the derived terrain attributes, thereby resulting in soil maps. Users' ability to select the right DEMs is pivotal to the outcome of any soil-landscape model application. There is, however, still varied consensus on the effect of DEM source and resolution on the application of these topographic attributes to landscape characterisation within South Africa. Open-source elevation data for South Africa are available from several primary sources and resolutions: SRTM, EarthEnv and SUDEM. Limited research has been conducted in a local context comparing the extraction of terrain attributes to high-resolution digital terrain data such as LiDAR that are becoming increasingly available. However, the utility of LiDAR to topographic analyses presents its own challenges in terms of operational-relevant resolution, processing demands and limited spatial coverage.

In most cases, data precision, grid resolution, and orientation errors present no issue for visualising and mapping three-dimensional elevation surfaces. Computationally, however, if these errors are not managed, they could impact derived parameters like slope and aspect, primarily dependent on the DEM surface quality. There is a need to quantify the impact that generalisation approaches have on simplifying detailed DEMs and compare the accuracy and reliability of results between high resolution and coarse resolution data on the extraction of localised topographic variables as a primer for future soil-landscape or digital soil models. This study addressed two main research objectives: first, to thoroughly examine the variability of spatial resolution, data source of DEMs and DEM generalisation effects on the extraction of selected topographic indices from an area located along KwaZulu-Natal's southern coastal region. Secondly, to investigate the statistical significance that elevation partitioning may have on the extraction and statistical significance of selected terrain properties, specifically, *terrain altitude, slope aspect, slope gradient, profile curvature, planform curvature, topographic wetness index, LS factor, surface roughness and lastly terrain ruggedness index*. These described terrain characteristics are among the key inputs to digital soil surveys derived from digital data with geographic information systems.

Unique to this regional case study is the application of three commonly used neighbourhood generalisation interpolators. The nearest neighbour, mean cell aggregation, and hydrologically corrected topo-to-raster methods were used to simplify a high-resolution 1 m LiDAR DTM to lower DEM resolutions and comprehensively evaluate the functionality and utility of these three generalisation methods under the influence of varying DEM resolutions. Whilst similar studies have evaluated the impact and applicability of fine and coarse resolution DEMs on terrain analyses, few studies have been conducted under local conditions by assessing the combination of interpolators at the given DEM resolutions: 5 m; 10 m; 20 m; 30 m; 60 m; and 90 m derived from LiDAR, SUDEM, SRTM and EarthENV platforms. Therefore, this study endeavoured to demonstrate the importance of utilising DEM surfaces that are "fit-for-purpose" for direct application to DSM for describing the scale-dependent topographic relationships relevant to future soil-landscape modelling applications under local terrain conditions. It must be emphasised that the DEM generalisation methods used in this study are not necessarily the most advanced for terrain surface modelling. However, they present the

most readily available methods accessible to non-specialists in many commercial GIS software packages and represent a suite of methods used to produce many of the DEMs in use for DSM and soil-landscape analysis today.

This study has shown that sensor selection and DEM resolution determine the application's utility by default. Careful consideration must be given to the generalisation approaches used in representing the scale-dependent terrain processes to be modelled. Notably, the scale-interpolation dichotomy inherently influences the accurate detection of certain terrain features across the site with optimal results obtained where the neighbourhood size and DEM pixel resolution are commensurate with the rate and extent of topographic variability within the landscape. This study has proved that terrain attributes respond to resolution change in characteristically different ways, especially when manipulating DEM resolutions between generalisation methods. Notably, we would expect higher agreements between derived surfaces and the reference surface at lower altitude ranges and less complex terrain. However, this study highlights a robust statistical association for particular terrain attributes, notably *SlopGrad*, *TWI* and *LS* factor, even at altitudes greater than 650 m a.s.l. While spatial resolution across sensor platforms may influence the level of detail between DEMs, the magnitude of DEM error variation is not always commensurate with the magnitude of error observed in their derived surface parameters. i.e., slight deviations between DEM surfaces may translate into significant errors in derived products, e.g., *SlopGrad*, *LS* and *TWI*, as the results have shown. The study highlights the need to quantify the scale dependency of terrain features within an area and recognise where these features can be accurately represented for deriving terrain attributes in the landscape. For instance, terrain parameters such as *SlopGrad* and *TWI* generally exhibited strong dependencies on elevation and spatial resolution regarding the statistical similarity to the actual surface, with lower accuracy observed at higher altitudes and coarser-resolution surfaces.

High spatial resolution data such as LiDAR provide new opportunities for a wide range of geomorphologic and pedological applications through visual and computational analysis. While LiDAR data is increasingly being used at an operational level and may offer far more accurate DEMs and higher spatial resolutions than other DEM sources, they often have limited spatial coverage with higher acquisition and processing costs. They may contain unnecessary artefacts or features that are not important for modelling soil-landscape relationships. End-users still question the necessity of purchasing such data, favouring readily available medium to coarse resolution surface data sets. This suggests that higher-resolution DEMs, derived from LiDAR or satellite platforms, may find optimal operational and financial return with small-scale, local applications and may not be necessary for optimising large-scale soil-landscape models. Indeed, for a coastal region of KwaZulu-Natal, this study has confirmed much of what has been established internationally that DEM resolutions and terrain variable extraction for soil-landscape modelling should align with the landscape-scale processes being modelled. Consequently, DEM resolution, precision, and source issues will become increasingly mainstream. Further developments and implementation of soil-landscape and geomorphological models using GIS and remote sensing platforms must be explored for DSM, mainly where detailed topographic field surveys are not feasible.

A significant outcome of this research highlights the variability of terrain attribute evaluation using descriptive measures of central tendency regionally, i.e., across the entire study site using all sample points and a local scale defined by segmented altitudinal class breaks. The study further demonstrates the importance of sensor selection, generalisation, and landscape position on the extraction and statistical similarity between terrain attributes derived at different spatial resolutions. Based on the overall findings of this study, the recommendation is that users' selection of DEM surface models should be research-driven, and landscape focused. DEM selection should prioritise performance either at a regional scale or a localised neighbourhood scale.

The findings on the influence of DEM horizontal displacement based on DEM vertical error needs attention. Users should be cognitive regarding the potential *in situ* horizontal errors on data precision, sampling area and upslope catchment area calculations associated with DEM vertical error. Moreover, end-users should expect lower prediction accuracies in derived terrain and surface products concerning DSM applications if these irregularities are not acknowledged and addressed in soil-landscape applications.

Though the study set out to explain the influence of DEM source and generalisation on the extraction of terrain variables within the southern coastal region of KwaZulu-Natal, there is no reason to suggest that these results are limited to this region. Users should therefore consider the results as being representative of the entire moist coastal region of KwaZulu-Natal. The overarching output of this research is a comprehensive quantitative assessment for analysing and interpreting the relationship between DEM resolution and terrain attribute extraction that can be summarised as follows:

- It is a common belief that a more detailed depiction of surface topography yields more accurate modelling results. In this study, the highest resolution DEM products generated at 5 and 10 m were not ideal DEM products for extracting terrain properties such as *LS* or *TWI*.
- The results further show that optimal correlations between terrain properties and moderate resolution DEMs (30-90 m) are suitable for describing specific local terrain variations. This means that freely available satellite sensors that operate at these resolutions may still offer a viable option of describing landscape form and process.
- However, DEM coarsening and generalisation beyond 30 m using the topo-to-raster generalisation method consistently resulted in a significant error in extracting specific terrain attributes. This approach should therefore be limited to hydrological or watershed management applications.
- For index-based terrain features such as *TWI* and *TRI*, the 1 m LiDAR provides too much terrain detail, perhaps more than is necessary for most pedometric applications with possible limitations related to data redundancy, processing time and scale-related issues of modelling soil-landscape relationships.
- The research further highlights that as the landscape exhibits more uniformity in morphology, the need for high-resolution data becomes negligible. Freely available data such as SUDEM, SRTM and EarthENV can account for the necessary variation in terrain properties such as *elevation*, *SlopGrad* and even *TWI*, particularly at a global-landscape scale.
- A positive outcome regarding selecting optimal generalisation methods is that the results confirm that both the NNR and *MA* methods are the most accurate and precise (repeatable, stable estimates) in their surface generalisations and retain the terrain properties across most tested resolutions.
- This study has also highlighted some significant limitations concerning the retrieval of terrain properties such as *elevation* and *SlopGrad, SlopAsp* from hydrologically corrected DEM surfaces. In particular, HCD derived surface parameters were not reliable for the actual surface estimates across a range of DEM pixel resolutions. This cautionary note urges against using hydrologically corrected DEM surfaces to extract and analyse topographic parameters unrelated to hydrology.

To conclude, the generalisation approaches and data source combination for the optimal description of surface properties has been explicitly described. Finding the right combination of when to upscale surface data, what DEMs to use and at what spatial scale to operate at to ensure that surface integrity is most optimal is mainly still context-specific. Many disciplines rely on the accurate representation and extraction of topographic parameters for various studies. The research presented herein demonstrates a robust framework to interpret optimal sensor choice and spatial scale for understanding the geomorphological processes in the landscape relevant to present and future DSM applications for the southern coastal area of KwaZulu-Natal.

CHAPTER 3

3 MULTI-RESOLUTION SOIL-LANDSCAPE CHARACTERISATION: ANALYSING THE UTILITY OF GEOMORPHONS TO CLASSIFY LOCAL SOILSCAPES FOR IMPROVED DIGITAL SOIL MODELLING[1]

"Intuitively, I understand that a system is complex if it has many constituent objects and the interactions among these objects are also 'many'. Yet, diversity, heterogeneity and connectivity are also key characteristics of a complex system, as are non-linear interactions and openness to environmental fluctuations."[2]

[2] Papadimitriou, F. (2009). Modelling spatial landscape complexity using the Levenshtein algorithm. *Ecological Informatics, 4*(1), 48-55.

[1]This chapter was published in Geoderma Regional: Atkinson, J.T, de Clercq, W.P., Rozanov, A.B., 2020. Multi-resolution soil-landscape characterization in KwaZulu-Natal: Using geomorphons to classify local soilscapes for improved digital geomorphological modelling. Geoderma Regional 22: 1-18

[1]The work was also presented at the International Union of Soil Sciences Digital Soil Mapping & GlobalSoilMap Workshop, 12-16 March 2019, Santiago, Chile.

3.1 INTRODUCTION

Digital terrain representation techniques have been developed since the middle of the 20th century, relying heavily on developments in geo-computational technology, modern mathematics, and graphical computer operations. However, the last decade has seen renewed interest in pushing the boundaries of topographic quantification and geomorphic regionalisation of land surfaces within a GIS environment (Bishop *et al.*, 2012; Florinsky, 2016). Over this period, land-surface analysis and classification have seen rapid improvements in the rate and quality of geomorphometric computational approaches, a key factor for the utility of digital geomorphic mapping (DGM) (Bishop *et al.*, 2012; Rigol-Sanchez *et al.*, 2015). DGM has provided users with a new set of tools with which to explore the conceptual issues of simulating single or even multi-landscape level functional and structural processes and hierarchical organisation of heterogeneous surface composition or character (Walsh *et al.*, 1997; Tate & Atkinson, 2001; Bishop *et al.*, 2003). Continual advances in quantitative modelling of surface processes combined with new Spatio-temporal geo-computational algorithms have revolutionised the auto-classification and mapping of landform components through the automated analysis of high-quality DEMs (Pike, 2000). Importantly, DGM has provided users with an outlet to transcend traditional humanistic and deterministic approaches to spatial organisation and landscape phenomena visualisation.

Conventional methods offer a high degree of regional or site-specific accuracy driven by expert input. However, human error often inhibited replication, analytical reasoning and standardisation of procedures, subjectivity and biases typically characterised by qualitative geomorphological and physiographic terrain analysis (Baker, 1986). Incidentally, with advanced DGM approaches, users are now able to quantify better and represent landscape morphology (Pike, 2000; Reuter *et al.*, 2009), evaluate surface-biophysical associations (Florinsky, 1998b; Liang, 2007; Smith & Pain, 2009; Tarolli *et al.*, 2009; Gregory & Goudie, 2011) and implicitly explore spatial landscape complexity (Papadimitriou, 2009) through a broad suite of diversity and contagion indices of landscape heterogeneity (Li & Reynolds, 1993). Current international literature on landform classification and soil-landscape analysis using DGM is extensive (Libohova *et al.*, 2016). Previous works have focused on deriving landform elements based on geomorphometric variable extraction (Evans, 1972; Pike, 1988; MacMillan *et al.*, 2004; Olaya, 2009) which can either be cell-based or object-based (Drăguţ & Blaschke, 2006; van Asselen & Seijmonsbergen, 2006; Drăguţ & Blaschke, 2008; Ghosh *et al.*, 2009). Further studies have explored the capability of manually designed classifiers integrating expert evidence (Pennock *et al.*, 1987; Skidmore *et al.*, 1991; MacMillan *et al.*, 2000; Gallant *et al.*, 2005; Iwahashi & Pike, 2007; Minar & Evans, 2008). While others have found varying levels of success with machine learning classification approaches (Brown *et al.*, 2001; Hengl & Rossiter, 2003; Prima *et al.*, 2006; Ghosh *et al.*, 2009) and methods focused on delineating the terrain into target units of classification or physiographic terrain units (Adediran *et al.*, 2004; Stepinski & Bagaria, 2009; Drăguţ & Eisank, 2012).

The specific application of geomorphons in DGM is yet to be fully explored under Southern Hemisphere conditions where a variety of landscape pattern-process functions inevitably influence a unique set of soil-landscape and pedo-hydro-geomorphic processes (Partridge *et al.*, 2010; Grab & Knight, 2015; Holmes *et al.*, 2016). There is an immense opportunity in a regional context to further examine the effectiveness of geospatial tools in understanding the scale-dependency of landscape pattern-process assemblages and calibrating model predictions to local geographical settings (Van Zijl & Le Roux, 2014). Regarding the pioneering of geomorphological studies that now form the baseline for many DGM principles, I acknowledge that much work has already been done regarding exposing the geomorphic principles that resonate in today's diverse South African landscape and landforms. At a national scale, the work by Lester King (King, 1942; King, 1967) in delineating 18 geomorphic provinces for Southern Africa based on geomorphic history, geological structure,

climate location and altitude provides a departure point in assessing how far the discipline of geomorphological mapping has progressed in South Africa. Recently, Partridge *et al.* (2010) refined the concept by identifying 34 geomorphic provinces and 12 sub-provinces relating to drainage structure and slope by combining macro-reach descriptors, statistical analyses, and digital terrain derived data (Holmes *et al.*, 2016). These studies have contributed much towards outlining the contemporary challenges of landscape classification and landform discretisation studies in South Africa. However, these terrain categories remain limited to macro-scale geomorphic classification and offer a minimal contribution to representing distinct, local-scale geomorphic sub-regions (Schumann *et al.*, 2011). Further notable studies explicitly relevant to this study's basis include works by authors such as Cooper (1991) to define a genetic geomorphological classification for Natal river-mouths (Rowntree et al., 2000) and provide a geomorphological classification system for the zonation of South African rivers. These studies have contributed much towards outlining the contemporary challenges of landscape classification and landform discretisation studies in South Africa.

Arguably the most prevalent representation of the landscape pattern-process for South Africa must be the National Land Type Survey Database (LTS) (ARC, 2003). A Land Type Unit represents a delineated area at a map scale of 1:250 000, displaying marked uniformity regarding terrain form, soil pattern and climate (Schoeman *et al.*, 2013), producing a homogenous distribution of soil properties across the landscape. Soil attributes are represented by a probabilistic organisation and symmetry of soil property associations and topographic processes known as a topo-sequence (Bushnell, 1943), or terrain morphological unit (TMU) linked to the following five terrain units: Crest (TMU 1), Scarp (TMU 2), Mid Slope (TMU 3), Foot Slope (TMU 4) and Valley Bottom (TMU 5). In the absence of readily accessible conventional or semi-detailed soil survey data for a vast majority of South Africa, local context studies still rely on the LTS database for describing a variety of process-pattern interactions within the landscape. Perhaps the most contemporary research to leverage off the LTS underpinnings and offer a refined regional model of landform classification is the work done by Van den Bergh *et al.* (2009). Their study successfully applied spatial modelling approaches to refine terrain morphological units and soil associations assemblages to resemble medium-scale, 1:50 000, soil-landscape relationships for the entire KwaZulu-Natal Province. However, a significant critique of the study is that despite improving on the terrain representation through the disaggregation of the coarse land type classes the study does not fully explore the potential of refining (expanding) the existing land type classes to cater for more geomorphic landforms (or hillslope processes for the matter). Incidentally, the pursuit of a universally acceptable domain ontology for defining a framework in landform characterisation, particularly in South Africa, remains ever elusive given the fugacious nature, regionality, scale-dependency and preference of geospatial technologies in addressing geomorphological problems (Cavazzi *et al.*, 2013).

Oddly, DGM technologies remain a source of *felix culpa* to the end-user. A major caveat of using DGM technologies for landscape analysis is the abundance of empirical black-box toolsets currently available to the general end-user. DGM technologies provide an objective and repeatable definition for geomorphological mapping of elementary terrain forms. However, many users are not adequately prepared for the technical rigour required to define and translate the scientific underpinnings of geomorphological concepts into a GIS environment. This presents limitations to parameterise or regionalise these terrain models, leaving the end-user often undecided or unconvinced of the resulting modelled landform representations. Concomitantly, users are now called upon to have a semi-detailed understanding of key discipline-specific concepts' practical and empirical underpinnings. Many considerations now exist before a terrain map can even be derived. So, while current DGM software/tool technologies expose a younger, more technocratic generation to the appreciation for quantitative evolutionary assessment of terrain features, so too does it prejudice more traditionalist-geoscientists reluctant, intimidated and disinterested in exploring these new avenues of research (Bishop *et al.*, 2012). DGM approaches that simplify and translate “human knowledge” into automatic terrain classification offer great potential for improved landscape analysis.

A design that has attracted attention in DGM internationally is geomorphon mapping proposed by Jasiewicz & Stepinski (2013). Geomorphons represent the fundamental micro-structures within a landscape and simultaneously define terrain attributes and landform types (Jasiewicz & Stepinski, 2013). Geomorphons are analogous to textons (Julesz, 1981) of a landscape. Their extraction from a DEM comes at a small computational cost considering that they simultaneously represent quantitative and stratified terrain attributes and landform types (Jasiewicz & Stepinski, 2013). Carré *et al.* (2007) successfully investigated a similar concept termed a "terron" for soil mapping that quantifies the relationship between terrain attributes and soil properties using a fuzzy k-mean classifier. They concluded that the typology of a terron would be a suitable surrogate to investigate potential soil-landscape relationships as the terron would represent terrain attribute information without the need for physical site excavation.

The geomorphon approach's product is the stratification of the landscape into ten unique but recognisable landform elements: *Peak*, *Ridge*, *Shoulder*, *Spur*, and *Slope*, *Hollow*, *Footslope*, *Valley*, *Pit* and *Flat* (Jasiewicz & Stepinski, 2013). This unconventional approach to landform classification uses a computer vision approach to search for changes in elevations along a continuous line-of-sight in eight cardinal directions away from a central zenith point, at nadir, to recognise the variation in topographic patterns in the landscape (Libohova *et al.*, 2016). Two essential concepts truly distinguish the geomorphon approach apart from its contemporaries. First, landform discretisation is based on geomorphometric variables and the complete topographic pattern corresponding to specific landform elements. The geomorphon approach leverages computer vision utilities rather than merely differential geometry tools to automate and replicate the same classification level performed by human analysts (Silva *et al.*, 2016a). Secondly, is the ability of geomorphons to determine a local pattern from a DEM using a neighbourhood approach with size and shape that self-adapts to the local topography using a line-of-sight principle (Lee, 1991; Nagy, 1991; Yokoyama *et al.*, 2002), enabling better matching of landform elements to the appropriate spatial scale. This aspect in itself is a revolutionary advantage to end-users not sufficiently skilled in the concept of DGM scale-dependency (how patterns change with scale), which is pivotal in feature selection and terrain generalisation (Cavazzi *et al.*, 2013). No *consensus omnium* yet exists for defining the optimal grain size for environmental factors to use in terrain analysis, and scale-dependency in soil-landscape studies is still mostly unresolved with minimal empirical guidelines available (Bishop *et al.*, 2012). The main reason for this is that determining an optimal grid size for classifying terrain heterogeneity depends on terrain provenance, terrain complexity, observed terrain phenomena, neighbourhood size and DEM generalisation approach (Wu, 2004; Hengl, 2006). This limits the direct applicability of most DGM approaches and models of soil-landscape studies to other regions (Quinn *et al.*, 1991; Vaze *et al.*, 2010). The main challenge with selecting optimal DEM pixel resolution is that firstly at finer resolutions, terrain variables contain too much unnecessary detail or "noise" that may lead to lower modelled accuracy. Secondly, at coarse resolutions, terrain variables may show too much generalisation or "smoothing" and not adequately represent terrain attributes or the land surface, ultimately limiting soil-landscape models' predictive capacity and accuracy (Cavazzi *et al.*, 2013; Atkinson *et al.*, 2017).

Minar & Evans (2008) point out that very few investigations have attempted to formalise landform taxonomy and methodology. Except for Ruhe & Walker (1968) widely accepted five hillslope profile positions for terrain classification (of which the LTS are loosely based upon), consensus both locally and internationally on terrain characterisation is still varied (Beven, 2001; McDonnell & Woods, 2004; Weiler & McDonnell, 2004). The utility of superior DGM tools, such as the geomorphon approach, simplifies the spatial definition and classification of terrain morphology within complex soils-landscape systems that necessitate further review in a local context.

This is the background against which I argue, supported by similar soil-landscape research (Smith & Hudson, 2002; Van Niekerk, 2010; Zerizghy *et al.*, 2013; Mashimbye *et al.*, 2014; Malan, 2016; Miller & Schaetzl,

2016), that geomorphic classification and delineation need to be better represented in discrete landform mapping endeavours in South Africa. Central to this directive is the shift away from an *ad-hoc* to a more formal and systematic approach to landscape character assessment (Wascher, 2005) across a range of DEM pixel resolutions, using modern geographic technologies and incorporating better base maps of topography into the landform mapping process (Miller & Schaetzl, 2014). Here landscape character is defined as the distinct, recognizable and consistent pattern of landscape elements that distinguish one landscape from another (Swanwick, 2002). Therefore, this definition underlines the explicit recognition of individual landscape elements that constitute the landscape, thus allowing for a systematic comparison of areas based on their landscape character (Galatowitsch *et al.*, 2009). In view of these factors, I believe the current study approach can enhance the understanding of soil and landform properties at a local to semi-regional scale in South Africa through improved landform classification and terrain-unit separation.

This study addresses the need to explicitly investigate how DEM spatial variations and geomorphon parameterisation influence the geomorphic terrain characterisation for soil-landscape feature extraction. The present study examines whether the geomorphon approach can accurately describe terrain features in three derived DEM models: the SUDEM, SRTM, and ASTER GDEM2 datasets generated two operationally relevant pixel resolutions, 30 m and 90 m. To test the self-adapting ability of the geomorphon approach under regional conditions, a total of 4 750 digital gridded terrain samples were used to quantitatively assess how DEM model and DEM resolution influence the extraction, generalisation and representation of digitally-derived terrain attributes. These include *slope gradient*, *elevation* and *terrain surface classification*. This study demonstrates that variation in the resulting terrain unit representation is constrained by discontinuities of selected elementary landscape patterns such as geomorphon frequency and soil-associated landscape properties such as *soil complex, soil texture, and soil depth*. A *Similarity Index* is introduced to compare the results between geomorphon surfaces to gauge the degree of recall and precision between the derived geomorphic landscape features for each DEM surface. Finally, the regional geomorphon-soil relationships' findings are presented in a readily interpretable and qualitative manner, providing a "quasi-landscape signature" for potential localised geomorphons. Applying the study findings may benefit practitioners looking to align or refine modelled terrain classification approaches with expert perception and formalises heuristic approaches.

This study outlines two key objectives: 1) to qualitatively assess the similarity in terrain representation and selected soil-landscape covariate extraction between four geomorphon surfaces derived at different spatial resolutions using a single set of model optimisation criteria in a mountainous region in Central KwaZulu-Natal. 2) then, to evaluate how well aligned these derived geomorphon surfaces and their (scale-dependent) topographic properties are to the heuristic underpinnings of accepted, operative soil-landscape relationships for the study region.

3.2 MATERIALS AND METHODS

3.2.1 Site description

This study was carried out in a 16 km x 11 km area (11 200 ha) located near the town of Bergville in the Central Drakensberg Region, KwaZulu-Natal (**Figure 3.1)**. The study area, which extends from *28º 38' 13. 17'' S* to *28º 38' 56. 99''S* and from *29º 0' 57. 01''E* and *29º 17' 11. 28''E* was selected as it provided a suitable range of terrain types from moderate to steep sloping hills to open and incised valleys showing evidence of anastomosing channels formed by old stream meltwaters. Moreover, regional-scale DEMs depict these rolling hills as distinct, recognisable features of the landscape. Changes in elevation from the lower-lying troughs to hillslope highs are approximately 45 m, and the distances between these features are about 150 m.

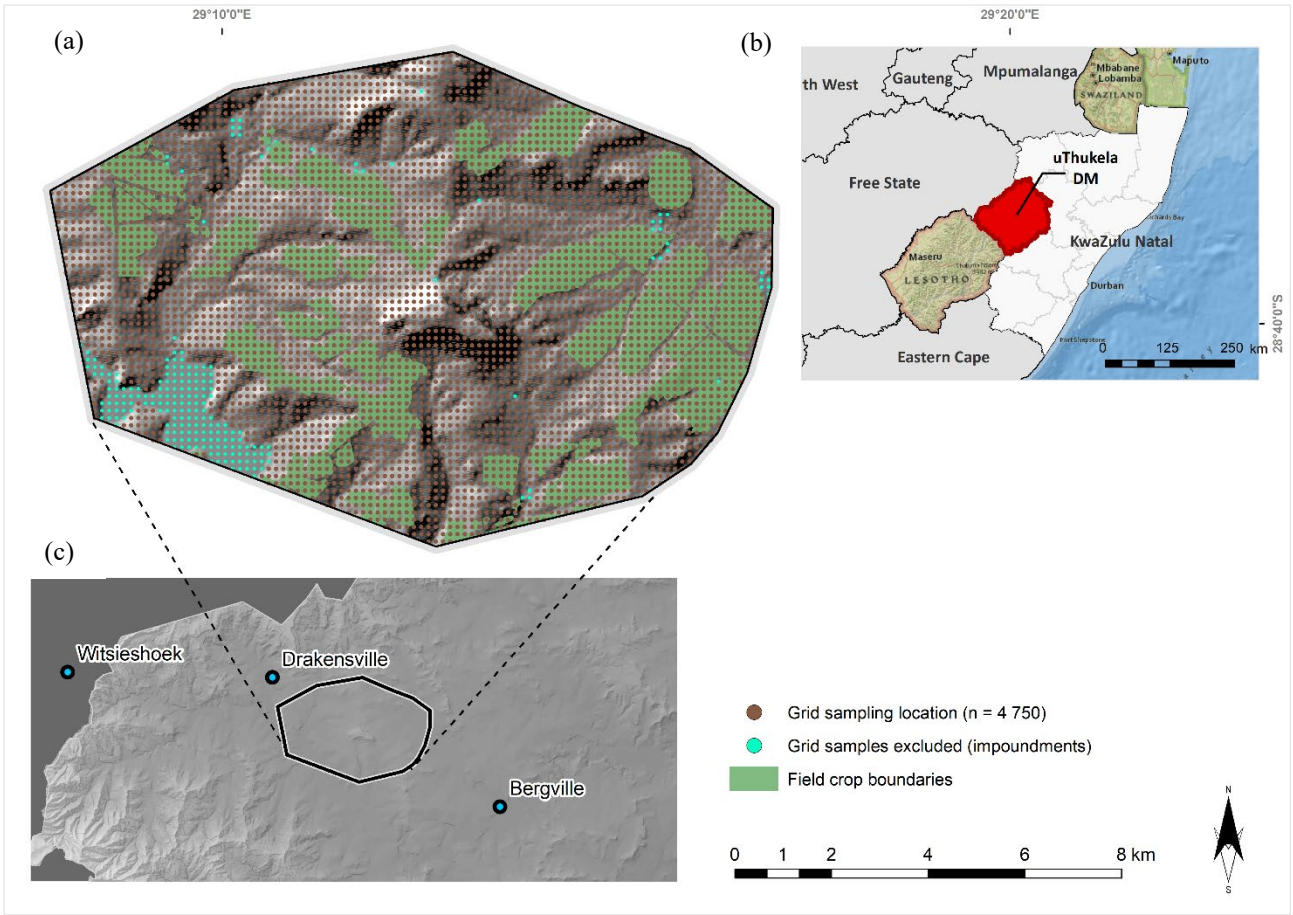


Figure 3.1: Bergville area - (a) Location of study site showing stratified sample observations (b) geographical overview of the study site (c) DEM delineated catchment area

Morphologically, features in the area represent the endemic erosional cycles associated with the process of pediplanation, characterised by scarp retreat with constant declivity across the landscape due to fluvial and

gravitational influences, exposing much of the old relic granites (King, 1967). The primary geological material across the site is of the Karoo Systems' Beaufort Series. The lithology consists mainly of blue, green, red, or even purple mudstones and shales alternating with yellow, fine-to medium-grained feldspathic sandstones (Van der Eyk *et al.*, 1969). The most distinctive landscape element in the study area is the noticeable presence of a table-topped hill (mesa) feature with well-defined scarp slopes extending to lower ranging hillslopes and, finally, an alluvial toe slope at the base of Woodstock Dam. The less dominant terrain features are characterised by moderate to steep undulating sandstone deposits through the site. The elevation ranges from 1 164 m above sea level in the east to 1 472 m in the center of the study area, with an average altitude of 1 318 m. MAT is in the region of 18.4°C, with temperatures dropping as low as 1.2°C between June and July and reaching 31.1°C during the summer months of December to January. The median annual rainfall for the region is 820 mm, with a higher concentration from December to February (Camp, 1995). The dominant land use in the area is natural grassland and veld (6 387 ha) with other prominent land uses including extensive commercial arable and livestock agriculture (3 961 ha), woodland and open bushland (153 ha) and water bodies (453 ha) with the prominent Woodstock Dam occupying 373 ha of the total catchment impoundment footprint. Most of the urban villages and settlements (32 ha) are limited to the lower-lying regions. In comparison, surface erosional features (55 ha) such as channeled gullies and "dongas" prevail in the moderate to lower surface gradient regions (GTI, 2015).

3.2.2 Data acquisition and analysis

3.2.2.1 Sample data points

To characterise landscape homogeneity, first, we must appropriately describe its spatial heterogeneity. Using the definition by Kolasa & Rollo (1991), a surface property is heterogeneous if its measurements vary in space. A more thorough description is that spatial heterogeneity in categorical landscape maps is defined as the complexity for both the composition (diversity) and configuration (spatial arrangement) of a particular land feature (Lausch, 2015). Landscape composition accounts for the types of categories present, including how many of these groups are present while ignoring the specific spatial arrangement of these classes on the landscape. In this study, spatial configuration refers to the particular spatial arrangement of the different geomorphon types in a landscape (Wang *et al.*, 2017). Characterising the spatial heterogeneity of geomorphons derived under a set of controlled parameters may help design the spatial resolution for future earth observing missions *vis-à-vis,* under local settings (Morisette *et al.*, 2002). Concerning the field soil survey campaigns, landscape spatial heterogeneity characterisation is essential for defining appropriate sampling and validation schemes that capture the soil and terrain surface property's spatial scale and optimises the field collection resources (Stein & Ettema, 2003; Morisette *et al.*, 2006). In this study, spatial heterogeneity has been defined through two components, namely: the *spatial variability* (DEM derivative variability) of the surface property over the observed study area; and the *spatial structure* (geomorphon variability) or landscape objects or patches that represent themselves independently and repeatedly within the study site at a characteristic length scale (spatial scale) (Garrigues *et al.*, 2006).

The study has adopted a purposive grid sampling approach (Aguilar *et al.*, 2005; Nanni *et al.*, 2011; George *et al.*, 2018) using a 150 m x 150 m sampling grid frequency resulting in a total of 4 750 derived gridded sample-point locations for the entire 11 200 ha site. This sampling design's choice is based on two fundamental research-specific considerations: 1) foremost, the design is intended to capture as much surface heterogeneity from the DEM-resulting geomorphon surface derivatives as possible by exhaustively surveying the entire study

region using all 4 750 point samples. This systematic gridded sample approach proved most optimal in comprehensively accounting for the terrain feature variation within the study site, limiting minimum biased feature extractions and resulting in higher reliability of final results in comparison with other possible sampling techniques, i.e., stratified cluster sampling or conditioned Latin-hypercube sampling (George *et al.*, 2018). Furthermore, the applied sampling approach will limit the possibility of drawing an unrepresentative sample of terrain features and enable the second-order geographical variance, i.e., global variance and spatial structure, to be better represented (Wang *et al.*, 2010). From a baseline perspective, characterising the landscape spatial heterogeneity in this manner may help define the spatial resolution for future predictive or statistical-based landscape analyses (Garrigues *et al.*, 2006). 2) Secondly, the choice of the sample-grid resolution is selected by design to corroborate any derived results with in-field applications to soil and landscape assessments typically done at 150 m x 150 m resolution being the *de facto* grid resolution for field-based surveys in the Province of KwaZulu-Natal (Smith, 2006). Incidentally, a review of the model proposed by Hengl (2006) to determine maximum location accuracy (MLA) or average size area (ASA) confirms that the 150 x 150 m gridded approach applied in this study is an appropriate ground scale to describe landscape variability with 30 m and 90 m DEM surfaces.

3.2.2.2 Digital elevation data

DEM choice generally contains tradeoffs between cost, accuracy, spatial coverage and grid size and how they are prepared and corrected (Robinson *et al.*, 2014). However, the quality of the DEM ultimately determines the accuracy and reliability of the spatial geomorphometric analysis. For this study, four DEM surfaces were analysed: a locally derived 30 m and 90 m (generalised from 5 m DEM surface using a nearest neighbor approach) Stellenbosch University Digital Elevation Model (SUDEM) (Van Niekerk, 2014; Van Niekerk, 2016) as well as the readily available C-Band, 90 m Shuttle Radar Topography Mission (SRTM) and 30 m Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model V2 (GDEM2) surface models. The SRTM and GDEM2 remain the most widely applied satellite-derived, near-global and medium/ high-resolution DEM data sets and were selected due to their reputable operational application and general ease of open-access to the end-user (Gesch *et al.*, 2012). The SRTM is reported to have an absolute vertical (orthometric) accuracy of 16 m or less, a relative vertical accuracy of 10 m or better and an absolute horizontal accuracy of better than 20 m (Rexer & Hirt, 2014; Sharma & Tiwari, 2014; Jarihani *et al.*, 2015). While the GDEM2, released in October 2011, has an improved vertical accuracy (orthometric) of 17 m and absolute horizontal accuracy of ± 30 m (Meyer, 2011; Gesch *et al.*, 2012; Jarihani *et al.*, 2015). The SRTM data were acquired using interferometry with C- and X-band SAR to derive a digital surface model (DsM), meaning that elevation heights represent the top of reflective features such as vegetation, anthropogenic features or bare soil, and are described above the WGS 84 geoidal datum (Prasannakumar *et al.*, 2011). Owing to its widely used applications, there has been continual improvement efforts and rigorous quality control of the SRTM products resulting in a highly trusted and widely used data product with current versions of the SRTM now offered as void filled products, including the 90 m elevation dataset used in this study (Jarvis *et al.*, 2008).

The GDEM2 is a product of a joint venture between NASA and Japan's METI. The GDEM2 surface product(s) have noticeably improved quality than the ASTER GDEM1, attributed to reduced artefact incidence, improved spatial resolution, and water masking accuracy (Rexer & Hirt, 2014). Despite the overall improved image quality, the GDEM2 dataset remains beset with region/ scene-specific anomalies and artefacts that require processing and removal before use, particularity at higher latitudes (> 60N) and lower latitudes (< 60S) (Robinson *et al.*, 2014).

Furthermore, improvement in both vertical and horizontal accuracy has come at the cost of increased noise in the datasets as well as reflectance-bias to land cover and reflective surfaces, with accuracy measurement being lower for forested areas, buildings etc. compared to bare areas (Meyer, 2011; Gesch *et al.*, 2012). The high-resolution SUDEM terrain products are gaining popularity in a variety of regional terrain analysis, digital-soil mapping, and hydrological modelling applications. The SUDEM is perhaps the most accurate and readily available localised high-resolution DEM for South Africa. SUDEMs are offered as a series of products developed by CGA situated at Stellenbosch University, South Africa. By processing large scale contours and spot heights with local interpolation approaches, DEM products are designed and offered at four different processed levels (Van Niekerk, 2012). Selected regional studies have shown that the SUDEM products, even when generalised to lower resolutions, can yield more accurate terrain feature representations than higher-resolution products such as LiDAR generalised to the same resolutions (Atkinson *et al.*, 2017). The SUDEM is an appealing option in terrain surface analysis in South Africa as it is both cost-effective and retains high feature quality. While there are spurious issues involving contour and spot height error estimations in the data, many of these errors (gaps) do not feature in the final interpolated product(s). The baseline level 2b product used in this study offered far superior data accuracy and image quality than other readily available primary surface models obtained from public-sector data custodians such as the South African Chief Directorate National Geospatial Information (NGI, 2018).

This study utilised the Earth Science Research Institutes ArcGIS® (ArcMap™, Version 10.5) and GRASS GIS (v 7.4.1) (Neteler & Mitasova, 2007) as the primary GIS platforms for all geospatial and modelling applications and interpretations for characterising the selected DEMs for geomorphon feature detection. All DEM products were re-projected to meters from a Geographic Coordinate System to the Universal Transverse Mercator (UTM) projection (Zone 36S) before analysis. Minimal post-processing procedures were applied to the DEM models used in this study, as the products were already suitably post-processed by the respective data vendors. These image enhancements included eliminating voids, spike and pit removal, and water body levelling, resulting in improved surface accuracy and height validation (Tachikawa *et al.*, 2011; Yang *et al.*, 2011). Further post-processing with majority-filtering and mean-filtering was applied to each DEM surface. This assisted in smoothing erroneous pixel values before the final reclassification of the geomorphon rasters into the ten discrete categories. It is important to note that while pre-processing is vital for model accuracy, DEM preparation can impact other model performance components (Callow *et al.*, 2007).

Consequently, none of the DEM models were further modified to generate hydrologically-conditioned DEM surfaces. This was a necessary and deliberate consideration since the study required identifying certain discontinuous terrain features such as pits (geomorphon Unit 7) and hollows (geomorphon Unit 10) within the landscape. Correcting for surface flow accumulation will eliminate these features from the geomorphon evaluation. Furthermore, Atkinson *et al.* (2017) reported that using hydrologically corrected DEMs to derive selected surface parameters is not reliable for extracting selected terrain variable estimates across a range of pixel resolutions.

A similar methodology was applied by Siart *et al.* (2009) using a hydrologic surface analysis approach in ArcGIS to detect hollows and sinks in the elevation. In this regard, a sink is defined as an area surrounded by higher elevation values and corresponds with a depression. Such terrain features are characterised by internal drainage and may be related to natural processes, but they may equally be attributed to imperfections in the DEM (Mark, 1988).

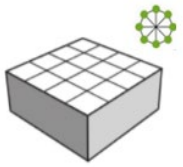
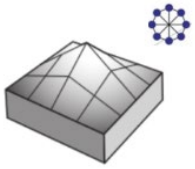
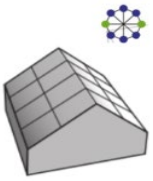
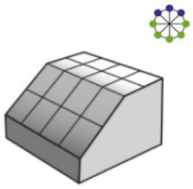
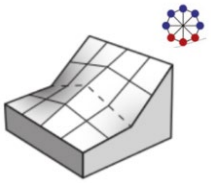
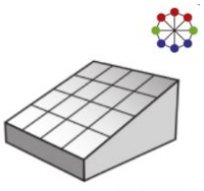
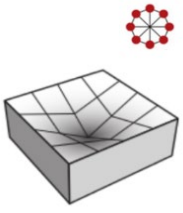
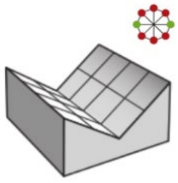
To ensure DEM comparability, all DEMs were processed to equal spatial resolutions of analysis of 30 m and 90 m, respectively (Bubenzer & Bolten, 2008). While both the GDEM2 and SRTM DEMs were used in their

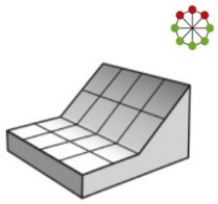
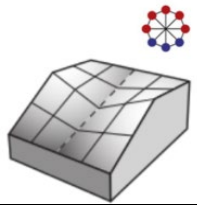
native resolutions of 30 m and 90 m, respectively. The 30 m and 90 m SUDEM DEM models were derived from a high-resolution 5 m SUDEM using the nearest neighbour generalisation approach outlined by Atkinson *et al.* (2017). Having been derived directly from the 5 m DEM, and therefore assuming the highest elevation accuracy, the 30 m SUDEM was used as the reference DEM to evaluate the geomorphon surfaces' accuracy and associated terrain variables from the other DEM sources. A vital part of this research involved identifying the relationship between geomorphon surfaces and terrain character and the similarities in terrain products across these scale-specific geomorphon surfaces. It is, however, beyond the scope of this research to explicitly address all the issues of DEM error assessment in a single chapter and readers are referred to the work of Florinsky (1998a), Thompson *et al.* (2001), Sorensen & Seibert (2007), Prasannakumar *et al.* (2011), Shafique *et al.* (2011) and Atkinson *et al.* (2017).

3.2.3 Geomorphon pattern characterisation

For creating the 2D geomorphon models, the DEMs were input to GRASS GIS (v 7.4.1) with the workflow using the *r.geomorphon* extension. The raster-based *r.geomorphon* application delineates the ten most universally accepted landform units by applying a pattern recognition algorithm based on a 3 x 3 local neighbourhood search radius from a central focal point (Jasiewicz & Stepinski, 2013). Each geomorphon coincides with the most common slope positions used to describe landscape features (**Table 3.1**). A vital function of the *r.geomorphon* extension is the ability to process extensive DEM datasets through optimal memory management (Jasiewicz & Stepinski, 2013). It does this by limiting geomorphon calculations to a local, user-defined neighbourhood extent (line of sight), ensuring that only a relevant, small portion of the entire DEM is read to the computer memory during processing. A central feature in generating geomorphic surfaces is *a posteriori* selection of window size and length scale (Wood, 2002). The *r.geomorphon* algorithm will then analyse the extent and shape of the featured neighbourhood to classify the landform elements by automatically self-adjusting/adapting to the local terrain's geometry derived from the DEM surface. Notably, model optimisation depends on three core parameters: maximum search radius (lookup distance), flatness threshold (t-degrees) and Skip radius (cells). A further consideration is the search radius (L-cells), representing the maximum distance for each pixel's line-of-sight (LOS) calculations. In order to evaluate how well the *r.geomorphon* tool can adapt to feature scale recognition and consistently detect similar local terrain features, landform feature resolution must be relative to DEM spatial (pixel) resolution. Therefore, users must systematically and iteratively, albeit on a trial-and-error basis, calibrate the geomorphon model inputs with parameters such as search and skip radii values to determine the appropriate scale of analysis for the study. Users can effectively skip small terrain variations to capture more massive landforms and uniformly optimise the final geomorphon surface outputs by adjusting these parameters. For this study, the optimisation approach enabled the us to determine whether geomorphic features detected at 30 m are diligently represented at 90 m and v*ice versa* across different DEM surfaces.

Table 3.1: Geomorphon (ternary patterns) corresponding to the ten most common landform elements evaluated in the study. The table, adapted from (Spijker, 2013), includes the symbolic 3D landform representations, including the terrain representations from neighbourhood pixel perspective (Red – Higher, Blue – Lower, Green – Same value) (Jasiewicz & Stepinski, 2013)

Geomorphon Number	landform	Description	Image
1	Flat	Includes areas with no elevation difference. In this case, this means a slope > 5º.	
2	Peak	An area in which the surrounding area is lower than a central cell.	
3	Ridge	A ridge is a continuous range of mountains or hills that form an elevated crest for some distance.	
4	Shoulder	A shoulder is a fairly steep, convex, transitional, and erosional zone from the summit to the back slope.	
5	Spur	A spur is a lateral ridge or tongue of land descending from a hill or mountain.	
6	Slope	Features include areas with a slope between 15º and 45º.	
7	Pit	An area in which the surrounding area is higher than a central cell.	
8	Valley	A depression feature with a predominant extent in one direction.	

9	Footslope	A footslope is a linear or concave landform downhill from an inflection point where the slope angle decreases and upslope eroding sediment collects.	
10	Hollow	A lateral, lower area of land descending from a hill, mountain or prominent ridge crest.	

This study adopted a similar approach to Luo & Liu (2018) and Gruber et al. (2015) by iteratively selecting model parameters that offered a reasonable balance between capturing landform elements' accuracy and model computational cost. After subjecting all the DEMs to several training iterations of the *r.geomorphon* model, with each model iteration characterised by a set of specific parameter combinations. It was confirmed that applying a constant outer search radius of 30 cells across all DEM surfaces allowed the classification outputs by solely observing surface variation resulting from pixel resolution and not the line of sight variations (Libohova *et al.*, 2016). The following optimisation parameters best represented the geomorphon surfaces across all DEM models (**Table 3.2**).

Table 3.2: Distance parameters (search, skip, flat distances) used in *r.geomorphon* for computational optimisation of multi-resolution DEMs.

DEM Source	Resolution	Outer Search Radius		Inner Search Radius		Flatness Threshold	Flatness Distance
		Cells	Distance	Cells	Distance		
SUDEM	30 m	30	900 m	6	180 m	1.2°	0
SUDEM	90 m	30	2 700 m	6	540 m	1.2°	0
ASTER GDEM	30 m	30	900 m	6	180 m	1.2°	0
SRTM	90 m	30	2 700 m	6	540 m	1.2°	0

Table 3.2 demonstrates that a constant outer search radius of 30 cells was used to map the g qeomorphon features between the DEM surfaces. The 30-cell search radius equals search distances between 900 to 2 700 m for the 30 and 90 m DEM surfaces, respectively. The *r.geomorphon* tool allows the user to model both inner and outer search radii using either cell number or ground distance (m) depending on user preference (Jasiewicz & Stepinski, 2013). Finally, the flatness threshold of 1.2° used in this study corresponded well with similar studies by Luo & Liu (2018) and Trentin & Robaina (2016). They settled on a maximum search radius of 2 500 m and 1 800 m and a flatness threshold of 1° and 2°, respectively. The SUDEM 30 m and 90 m and the SRTM DEM identified nine geomorphic features, while the Aster GDEM2 identified all ten geomorphic features, including the “Pit” terrain feature.

3.2.4 Extracting the primary topographic covariates

All geomorphon features were assessed against three primary compound terrain attributes digital elevation height (*Elev)*, slope gradient (*SlopGrad*) and terrain surface classification (*TSC*). These indices were selected since they are readily extractable in a GIS environment, are simple to interpret and compare across datasets, can be analysed using descriptive or non-parametric statistical approaches and are regularly used as input variable parameters in similar landscape assessment studies (Kumar, 2013). Many studies still rely on local elevation and slope as easily interpretable parameters to identify and compare landform types (Saadat *et al.*, 2008). In fact, selected studies have shown elevation range alone to be sufficient to identify hill and mountain areas. The slope gradient is a fundamental input parameter in almost all landscape classification models. It denotes the change in steepness or inclination, either as a percentage or in degrees, of a given surface over its horizontal plane (Chabala *et al.*, 2013). Slope gradients, represented as percentage values for this study, were extracted using the *Slope* tool in ArcMap's *Spatial Analyst* suite of tools. Two second-order terrain attributes included and particularly prominent in quantitatively evaluating soil-landscape relationships (Speight, 1977; USDA, 2017) are profile curvature and terrain roughness (Florinsky *et al.*, 2002; Iwahashi & Pike, 2007; Prasannakumar *et al.*, 2011). However, the mathematical interpretation and application of these attributes are not straightforward. While the simplest definition of surface roughness is that it represents the contrast of fine and coarse topographic spacing of features encompassing both relief frequency and spacing within the horizontal plane; in truth, there are still no standard methods for quantitatively characterising the anisotropic nature of surface roughness (Pike, 2000).

Moreover, while curvature values are useful indicators of areas of potential erosion or deposition (Olaya, 2009), depending on either the convexity and concavity of the landscape, the mathematical calculation of curvature is derived in units of radians. Values are generally small and impractical for analysis and interpretation of the results Zevenbergen & Thorne (1987). Consequently, it was decided to evaluate the *Terrain Surface Classification* attribute, as suggested by Iwahashi & Pike (2007), as a dual-proxy indicator based on slope gradient curvature (convexity) and surface roughness for land-use classification. The calculation of *TSC* requires input data for DEM curvature and roughness. These attributes were calculated using SAGA (Version 2.14) for the DEM products using the default 1 000 m *LOS* distance parameter and executed in the *Terrain Analysis* module. The result is a new thematic surface classified into eight landscape categories (Watkins, 2015). To avoid data redundancy in derived model representations and to align the *TSC* estimations to the relevant assessment scale of the study, the 8 *TSC* classes were reduced to four categories, namely: *Gentle Slope*, *High Convexity (GSHC*); *Gentle Slope, Low Convexity (GSLC); Steep Slope, High Convexity (SLHC) and Steep Slope, Low Convexity (SLLC).*

3.2.5 Extracting legacy soil data: Tugela Basin database

This study obtained the primary source of synoptic soil information that provided reasonable area coverage from the semi-detailed survey of the Tugela Basin's Soils (TBS). It was Van der Eyk *et al.* (1969) who first produced a compilation of hardcopy soil maps at a scale of 1:50 000 for the Three Rivers Area (Umvoti, Umgeni and Illovo) and 1:18 000 topographic coverage for the Tugela Basin, in which the study area resides. The survey was a mammoth undertaking, even by modern standards. More than 22 000 soil profiles were examined in detail across the 31 500 km^2 coverage with locational positions hand-recorded on hard-copy photo-mosaics and topographical maps during the study. Over 250 modal profiles and 819 soil samples were collected for each soil pit, which was analysed further to determine particles size distribution, soil reaction

characteristics (pH, KCL), and soil organic carbon levels. However, the survey's purpose was not to provide a blueprint for detailed land-use planning. Therefore, the data do not show the distribution of individual soil series or phases of series for the region. Instead, the data reveals the broad zonal arrangements in the soil series distribution, with each mapping unit potentially comprising as many as several associated series (Van der Eyk *et al.*, 1969).

Nevertheless, nearly five decades on, the survey remains the basis on which many subsequent classification systems such as the Land Type Survey and KwaZulu-Natal land Capability Classification (Camp, 1995) have been developed. For this study, a digital version of the TBS survey data was used. Initially, the data was acquired in a vector format but later rasterised for automated overlay analysis with the other covariate terrain datasets. In the absence of readily interpretable taxonomic criteria, the study applied categorical properties (as defined in the soil dataset's original attribute space) associated with soil properties, such as dominant soil complex (mixed category unit), dominant clay content, and dominant soil depth class. An important element of this study is that soil complexes are understood to be soil mapping units consisting of two or more dissimilar components that occur in a regularly repeating pattern. The major components cannot be distinguished from one another at the scale of mapping (Schoeneberger *et al.*, 2012). I acknowledge that these indicators do not provide an entirely definitive definition for soil taxonomic classification by modern standards or approaches. However, these descriptions do provide a suitable aggregated mapping unit to denote the distribution of soils in the absence of more specific soil taxa (Van den Bergh *et al.*, 2009; Van Zijl *et al.*, 2014). In fact, these workarounds must be encouraged, rather than avoided, particularly when digitising historical or legacy soil datasets for updating current land and soil resource information (KZNDARD, 2015). Notwithstanding, the legacy soil data remain relevant to providing a first approximation and departure point to assess the soil-landscape component between derived geomorphon surfaces. This is especially useful where soils of different types are mixed geographically so that the scale of the map makes it undesirable or impractical to show each one separately.

3.2.6 Data analysis and measures used for map comparisons

The 4 750 sample data points were used to extract the raster values from each *geomorphon*, *Elev*, *SlopGrad*, *TSC*, *soil complex, clay percentage,* and *soil depth* data set respectively in ArcGIS using the *Extract to Multipoint Feature*. The data were then imported to *STATISTICA* 13.2 for statistical and exploratory data analysis. Data analysis for similarity assessment of geomorphon features between DEMs was primarily limited to simple quantitative assessment metrics of central tendency and variance around the mean value. Each geomorphon surface feature was classified according to *count, trimmed mean, min, max, standard deviation, coefficient of variation and standard error* values for elevation and slope.

Also, data for *Elev, SlopGrad*, *TSC, soil complex, clay content and soil depth* were categorically classified and represented by sample-point frequency to compare dominant soil-landscape features across each geomorphon surface. Finally, to independently evaluate the predicted geomorphon surface products' accuracy and similarity, the resulting maps (90 m SUDEM, 90 m SRTM and 30 m GDEM2) were compared to the designated reference surface (30 m SUDEM). A cell-by-cell comparison of raster values for each geomorphon surface layer and the derived datasets was used to analyse feature similarity using recall and precisions measures. Specifically, the novel application of a composite measure (BK) (Leifman *et al.*, 2003) to evaluate geomorphon classification similarity (Smirnoff *et al.*, 2008) was used in this study (**Figure 3.2**). This BK metric (normalised between -1 and 1 but typically reported as $n \times 100\%$) has been used in other studies to evaluate the accuracy of different classification methods and has the advantage of being readily produced from

a confusion matrix (Sharifzadeh *et al.*, 2005) of predicted vs observed observations. Geomorphon feature similarity was then assessed by the following: *Recall = Cells of Unit X placed as X on resulting map / All cells of Unit X on the original map* then (1- *Recall).* This is considered to be an indication of the classification of false negatives. Similarly, the *Precision = Cells of Unit X placed as X on resulting map / All cells of Unit X on the resulting map* represents the portion of cells rightfully placed in the units of the resulting map with (1- *Precision)* representing the so-called false positives (Smirnoff *et al.*, 2008).

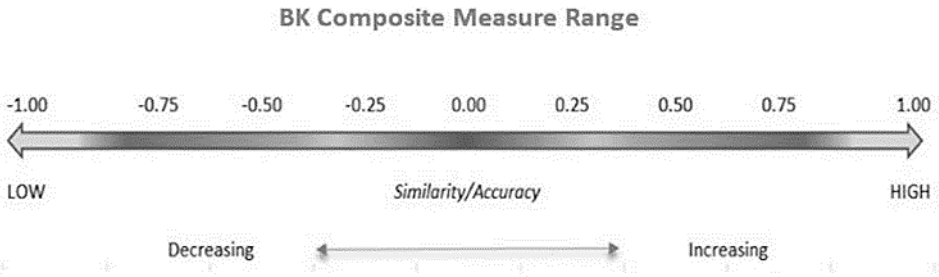


Figure 3.2: BK composite measure range used to evaluate the similarity between the results of two computed Geomorphon surfaces. Values above zero indicate close correspondence to the original map, while values below zero imply a high degree of dissimilarity with reference Geomorphon surface.

The ability to quantify the overall similarity between different landscapes by merely evaluating the associated and derived surfaces' descriptive data is a critical element of the outlined methodology. A significant advantage of using the composite BK approach is that it allows two or more generalised maps to be compared, regardless of spatial resolution differences, with either serving as a reference map. The observed BK values then represent the degree of similarity/ dissimilarity between two or more maps. The high-resolution SUDEM 30 m, including all associated derived products, was used as a reference (observed) map. In contrast, the other DEM surfaces and their related products served as the predicted (produced) surface(s). By combining the recall and precision, the BK = *Recall + Precision – 1* score can be calculated with the range of the composite values represented between -1 and 1, with higher similarity tending towards more positive values reaching 1.

3.3 RESULTS AND DISCUSSION

3.3.1 Geomorphon similarity, cross-tabulation, and BK composite measure

The overall similarity results of the unsupervised surface classification for the BK composite comparison are presented both graphically and through a confusion matrix (**Table 3.3 to 3.5**). The results reflect the geomorphon feature comparison of the SUDEM 90 m (**Figure 3.3b**), SRTM 90 m (**Figure 3.3c**) and ASTER GDEM2 (**Figure 3.3d**) to the reference SUDEM 30 m surface (**Figure 3.3a**). Further visual inspection of the geomorphon surfaces reveals that the reference 30 m SUDEM exhibits the most surface detail. In contrast, the 90 m geomorphon surfaces depict a relatively smooth, generalised landscape with geomorphon features still distinctive and most features well discretised compared to the 30 m SUDEM. In contrast, the 30 m GDEM2 surface displays noisy and incoherent geomorphon consistency with minimal feature characterisation due to the "fractured" nature of the DEM surface.

Considering these first approximation visual comparisons, we can conclude that while the reference 30 m SUDEM surface may contain detailed geomorphon detail, particularly with the representation of discrete features, such as *ridges*, *slopes* and *valleys*, the 90 m geomorphon surface products can graphically represent similar morphon features. Perhaps not in extent, but certainly in terrain position. This outcome was not entirely unexpected, given the correlated *mean Elev* values across all DEM surfaces (**Table 3.6 to 3.9**). Nonetheless, the result remains positive in the context of this study, given that the consensus is that as pixel resolution of the DEM surface decreases, slope values are underestimated, typically leading to a misrepresentation of terrain features at coarser resolutions (Warren *et al.*, 2004; Atkinson *et al.*, 2017). However, deferring to the quantitative result comparisons for the overall similarity performance, i.e., the BK composite index, the results between the predicted and reference surfaces are not particularly promising. The performance described here refers to the quality of the classifier (predicted surface) and its actual usefulness for automatic mapping of the "true surface" or some permutation thereof (Jasiewicz *et al.*, 2015). The low BK values displayed in **Table 3.3 to 3.5** indicate that most predicted geomorphon units are statistically dissimilar to the reference geomorphon surface. In particular, the predicted geomorphon features that showed the lowest similarity in both 90 m DEM products were the *flat*, *shoulder*, *spur* and *foot slope* terrain features. Coincidently, the same geomorphon features also exhibited higher slope gradient standard deviations to the SUDEM 30 m surface (**Table 3.6 to 3.9**).

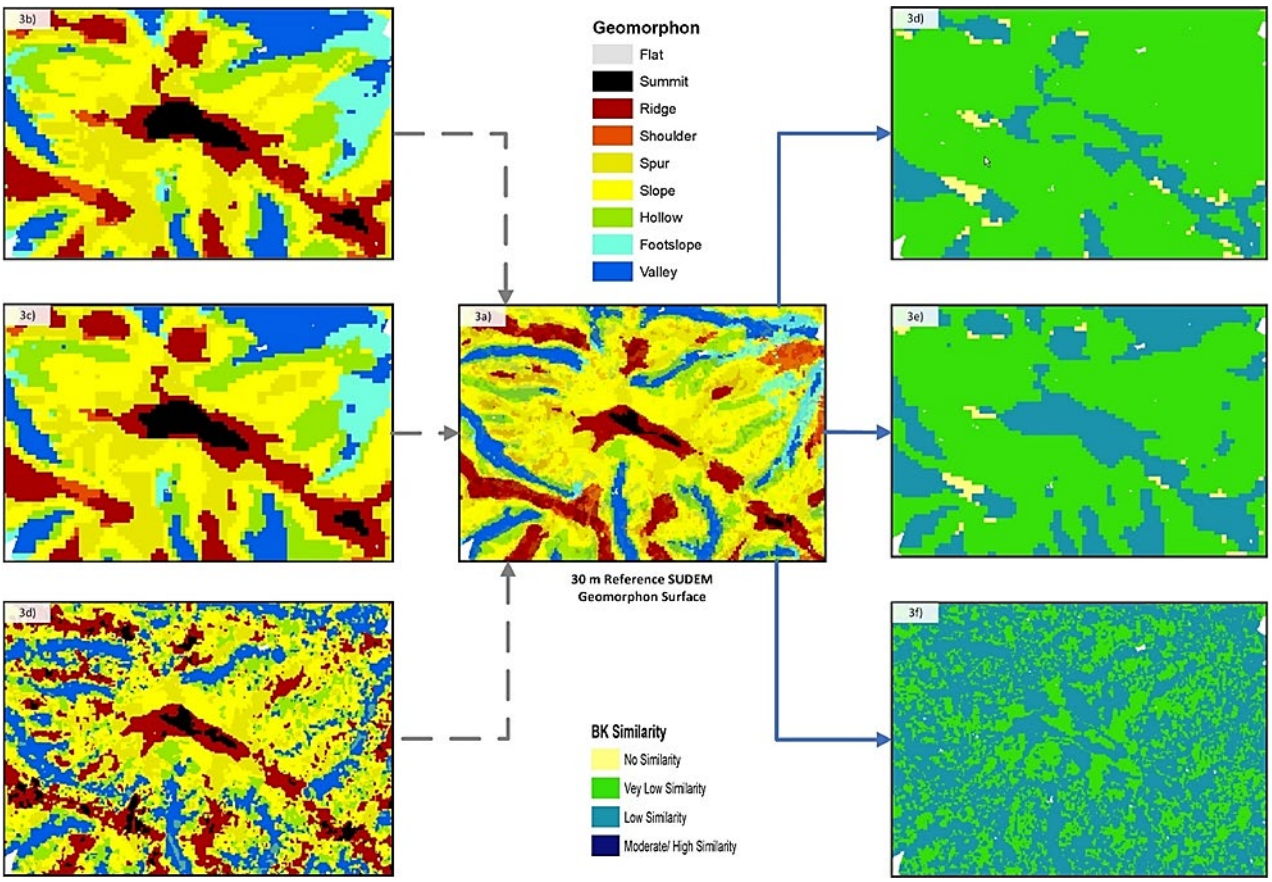


Figure 3.3: Overview of a centrally located Sub-AOI showing a 2-D comparison of (a) reference SUDEM 30 m geomorphon surface with (b) SUDEM 90 m surface (c) SRTM 90 m surface and (d) GDEM2 30 m geomorphon surface. Figures (e-f) show the BK similarity for the SUDEM 90 m, SRTM 90 m and 30 m GDEM2.

Interestingly, a study conducted by Gruber *et al.* (2017) similarly found that geomorphons poorly mapped features such as *shoulders*, *foot slopes*, and localised *flat* regions despite applying best parameter settings for each topographic position with only minor gains in accuracy related to scale variation. In contrast, the 30 m GDEM2 displayed a lower similarity to the reference 30 m SUDEM surface than the 90 m products with almost all morphon features. GDEM2 displayed lower BK values with features such as *flat*, *shoulder* and *foot slope* with BK scores even approaching -1.00, suggesting complete dissimilarity with the reference surface. Conventionally, these results would not be expected given that the GDEM shares the same resolution with the reference 30 m SUDEM. Admittedly, the GDEM2 coverage for the study site is inherently poorly represented, with high intra-class fragmentation and heterogeneity of surface elevation. In this instance, the overall poor performance of the GDEM2 geomorphon classification is predominantly a manifestation of the sub-optimal quality of the DEM surface – for reasons highlighted previously. While the results of the GDEM2 geomorphon comparisons show just how compromised the quality of the surface representations are for this region, it is difficult to completely disregard the GDEM2 surface for soil-landscape studies given that the GDEM2 global elevation remains positively correlated to the reference 30 m SUDEM elevation surface (**Figure 3.6**). For a more robust evaluation of the extraneous predicted surface (dis)similarity to the reference surface, two additional measures of model effectiveness were evaluated. Model recall (*producer accuracy*) is defined as a measure of how well the DEM process correctly identified a specific geomorphon type; and model precision (*user accuracy*) representing a measure of how well the DEM process captured all occurrences of any specific geomorphon type (Leifman *et al.*, 2003).

Table 3.3: BK similarity comparison of geomorphometric features between the reference SUDEM 30 m and SUDEM 90 m DEM surface.

		PREDICTED (90m SUDEM)											SIMILARITY		
		FLAT	SUMMIT	RIDGE	SHOULDER	SPUR	SLOPE	HOLLOW	FOOTSLOPE	VALLEY	DEPRESSION	TOTAL	RECALL	PRECISION	BK
ACTUAL (30m SUDEM)	FLAT	16	0	0	2	0	0	0	45	2	0	65	0.25	0.16	-0.6
	SUMMIT	0	19	11	0	0	0	0	0	0	0	30	0.63	0.28	-0.1
	RIDGE	5	42	429	41	89	33	1	4	1	0	645	0.67	0.55	0.21
	SHOULDER	5	0	27	36	14	35	15	30	6	0	168	0.21	0.15	-0.6
	SPUR	6	5	192	50	224	194	27	36	18	0	752	0.3	0.36	-0.3
	SLOPE	35	2	111	88	257	607	168	207	123	0	1 598	0.38	0.53	-0.1
	HOLLOW	2	1	13	11	30	205	155	73	119	0	609	0.25	0.34	-0.4
	FOOTSLOPE	28	0	0	5	5	18	7	150	135	0	348	0.43	0.24	-0.3
	VALLEY	1	0	4	1	4	64	80	83	297	0	534	0.56	0.42	-0
		98	**69**	**787**	**234**	**623**	**1 156**	**453**	**628**	**701**	**0**	**4 749**			

Table 3.4: BK similarity comparison of geomorphometric features between the reference SUDEM 30 m and SRTM 90 m DEM surface.

		PREDICTED (90 m SRTM)											SIMILARITY		
		FLAT	SUMMIT	RIDGE	SHOULDER	SPUR	SLOPE	HOLLOW	FOOTSLOPE	VALLEY	DEPRESSION	TOTAL	RECALL	PRECISION	BK
ACTUAL (30 m SUDEM)	FLAT	6	0	0	0	2	1	2	44	10	0	65	0.09	0.29	-0.6
	SUMMIT	0	22	8	0	0	0	0	0	0	0	30	0.73	0.28	0.01
	RIDGE	1	48	433	18	102	30	3	8	2	0	645	0.67	0.52	0.19
	SHOULDER	1	0	37	22	21	32	14	34	7	0	168	0.13	0.2	-0.7
	SPUR	1	8	207	28	221	213	30	16	28	0	752	0.29	0.34	-0.4
	SLOPE	6	1	125	35	257	651	219	151	153	0	1 598	0.41	0.54	-0.1
	HOLLOW	0	0	13	7	36	205	157	41	150	0	609	0.26	0.3	-0.4
	FOOTSLOPE	6	0	0	2	8	25	9	149	149	0	348	0.43	0.31	-0.3
	VALLEY	0	0	5	0	6	52	87	42	342	0	534	0.64	0.41	0.05
		21	**79**	**828**	**112**	**653**	**1 209**	**521**	**485**	**841**	**0**	**4 749**			

Table 3.5: BK similarity comparison of geomorphometric features between the reference SUDEM 30 m and ASTER GDEM2 30 m DEM surface.

		PREDICTED (30 m ASTER GDEM)											SIMILARITY		
		FLAT	**SUMMIT**	**RIDGE**	**SHOULDER**	**SPUR**	**SLOPE**	**HOLLOW**	**FOOTSLOPE**	**VALLEY**	**DEPRESSION**	**TOTAL**	**RECALL**	**PRECISION**	**BK**
ACTUAL (30 m SUDEM)	**FLAT**	0	1	4	5	7	16	4	5	15	8	65	0	0	-1
	SUMMIT	0	17	11	2	0	0	0	0	0	0	30	0.57	0.13	-0.3
	RIDGE	0	79	364	17	111	53	11	1	7	2	645	0.56	0.44	0
	SHOULDER	0	6	61	4	31	35	16	0	14	1	168	0.02	0.06	-0.9
	SPUR	0	10	199	8	269	178	50	15	22	1	752	0.36	0.35	-0.3
	SLOPE	0	11	150	22	295	702	237	19	154	8	1 598	0.44	0.55	-0
	HOLLOW	0	0	13	5	32	147	210	18	178	6	609	0.34	0.32	-0.3
	FOOTSLOPE	1	2	24	3	26	86	54	8	122	22	348	0.02	0.08	-0.9
	VALLEY	1	0	5	2	7	54	82	36	318	29	534	0.6	0.38	-0
		2	**126**	**831**	**68**	**778**	**1271**	**664**	**102**	**830**	**77**	**4 749**			

The *summits*, *ridges*, *slopes*, and *valleys* show relatively high recall or precision accuracy values concerning the SUDEM surface in all three predicted surfaces. For instance, while the 90 m SRTM surface yielded low overall BK values, the recall accuracy – that a mapped geomorphon is correctly predicted – is high with 73%, 67% and 64% for the *summit*, *ridges* and *valleys,* respectively. Interestingly, even the GDEM2, despite the low similarity in the BK metrics, can still map 57% and 56% of the *summit* and *ridges* in the reference surface 30 m SUDEM surface. These results indicate that some replication between the predicted and reference surfaces, albeit for limited geomorphon features and not complete landscape level, is possible even with coarse DEM resolutions of 90 m despite inherent artifactual inconsistencies in selected DEM surfaces like the GDEM2.

These results present both advantages and drawbacks for future terrain analysis using geomorphons. Firstly, the application of coarse resolution DEMs to geomorphon feature discretisation holds promise for soilscape studies in South Africa, particularly in regions where open-source global DEM datasets such as SRTM are still a primary source of terrain data. However, since the GDEM2 can comparatively represent approximately 60% of the reference DEM surface and geomorphon features despite its mediocre DEM properties begs the question: how much of the model precision in the output predicted geomorphon surfaces is a result of statistical or spatial chance or even endogenous error? Furthermore, the BK approach used in this study is adapted from the method developed by Smirnoff *et al.* (2008), who applied the metric to compare the results obtained from surface generalisation using a cellular automata algorithm on a cell-by-cell basis. While novel when applied to this study, the application of discrete point feature comparisons may have led to a far more restrictive sample-space analysis of results than a complete pixel-based grid comparison of datasets. This may have contributed to the overall poor BK results but moderately acceptable precision and recall results between the predicted and actual geomorphon surfaces. Further investigation into selecting a robust and suitable "benchmark" for similarity assessment between geomorphon surfaces may need further definition for future applications in terrain analyses.

Reflecting on the collective findings of the multi-resolution surface similarity analysis for this study, I am compelled to raise a crucial question regarding the future utility of the geomorphon approach in similar regional environments: *is the geomorphon approach truly as scale-independent as initially anticipated?* The preliminary results of this study so far, and corroborated by results of similar terrain evaluation studies by Gruber *et al.* (2015) and Luo & Liu (2018), suggest that the answer is perhaps more polarised and less anecdotal than initially expected. Indeed, geomorphons may offer a novel approach for identifying landforms by leveraging the recognition of geomorphological phenotypes using machine vision principles, unconstrained by scale issues (Jasiewicz & Stepinski, 2013). Nevertheless, the application of "machine vision" or "visual perception" to identify terrain-feature patterns, whilst immensely contemporary and efficient, should signal to users that DEM parametrisation inputs and landform relevant outputs can be infinitely more complicated to define and interpret by end-users with extreme subjectivity and variance in surface outputs both within and between DEM scale. For geomorphons to be functionally relevant, it is necessary to reconcile the optimal scale for simple static terrain visualisation versus the representation of spatial features and associated environmental phenomena or processes across the landscape, particularly for soil-landscape studies.

Incidentally, the monologue by Levin (1992) states that *"scale represents the window of perception, the filter or the measuring tool through which a landscape may be viewed or perceived"*. Clearly, the primary consideration for users of geomorphons (and the success they have thereafter) is that the geomorphometric application and interpretation must be governed by primary geomorphological factors such as the characterisation of local or regional landforms in the landscape (Pike, 2000). This, therefore, leads to an important distinction in applying the geomorphon approach to either specific or general landscape geomorphometry applications and, more importantly, the extraction of morphologically visible "texton" features in the landscape (Evans, 1972; Drăguţ *et al.*, 2013).

Therefore, it is incumbent upon users to invest adequate time in pre-planning and troubleshooting the optimal synergies between DEM quality and geomorphon parametrisation before their strategic and operational uptake in terrain or soil-landscape based studies.

Indeed, the intelligible multi-scale format of the geomorphon approach may appear to wholly endogenize the otherwise complicated inputs of other pixel-based terrain models. However, MacMillan & Shary (2009) concluded that it would be nearly impossible and impractical for any terrain model using any fixed dimension of search radius to perfectly capture the variation of landform features of interest in any given area. With the geomorphon approach, this expectation is not unwarranted. By users specifying an infinitely large search radius (*L*) and low flatness threshold (*t*) when trying to define geomorphons in the landscape, the vision-based model should, in theory, identify all possible landform elements within a particular landscape (Jasiewicz & Stepinski, 2013). However, shorter (*L)* values will attempt to register surface microfeatures, while applying larger (*L)* values will result in smaller features lumped into other more common landforms with the apparent implication of false-positive morphon classifications.

Given the low frequency of "narrow" and "finer" resolution features such as *flats*, *shoulders* and *summits* (**Figure 3.4 and 3.5)**, it would not be surprising if many of these more minor features were grouped with the more extensive and more frequently occurring elements such as *slopes* and *valleys* and *ridges.* This may be particularly possible in the 90 m DEM surfaces during the generalisation or "smoothing" process. In fact, Luo & Liu (2018) found that geomorphon maps created using different search radii converge and yield no significant comparisons as the maximum search radius increased to a threshold of zero variance. A similar trend was observed in this study between the 30 m SUDEM surface and 90 m products. Despite the difference in pixel resolution, the threshold values set in **Table 3.2** resulted in similar geomorphon features detected across the landscape at different extents, abundance and detail. Inset maps for each of these results are presented in **Figure 3.3a to 3.3d,** while **Figure 3.3e to 3.3g** represent the degree of similarity between the geomorphon features for each DEM surface. Similarly, selected findings from the present study on this particular aspect are corroborated by Libohova *et al.* (2016), who reported little to no significant difference between geomorphon products with a different line of sight within the same pixel resolution. Consequently, the line of sight threshold(s) evaluated in this study would be considered consistent and applicable for mapping most terrain features across that landscape across the observed DEM resolutions.

The findings of this research gain further momentum, given the detailed considerations and comparison of automated landform detection methods that Gruber et al. (2015) outlined. Their study concluded that there is no significant difference between terrain features characterised at meso (50-100 m) and macro (250-400 m) extents when uniform *(L)* values were tested with the only important consideration that at a mesoscale level, some terrain classes were not reliably detectable. While the approach may not provide new perspectives to geomorphon application at a global scale, i.e., > 5 000 m, or generated much new evidence to the broader international dialogue, what is encouraging is that the present results are in line with reviews of a similar localised scale and data specifications, precisely that of (Gruber *et al.*, 2015; Libohova *et al.*, 2016; Trentin & Robaina, 2016). This outcome is encouraging for future studies aiming to optimise geomorphon outputs for regional landscape characterisation using high and medium resolution, open-source DEMs.

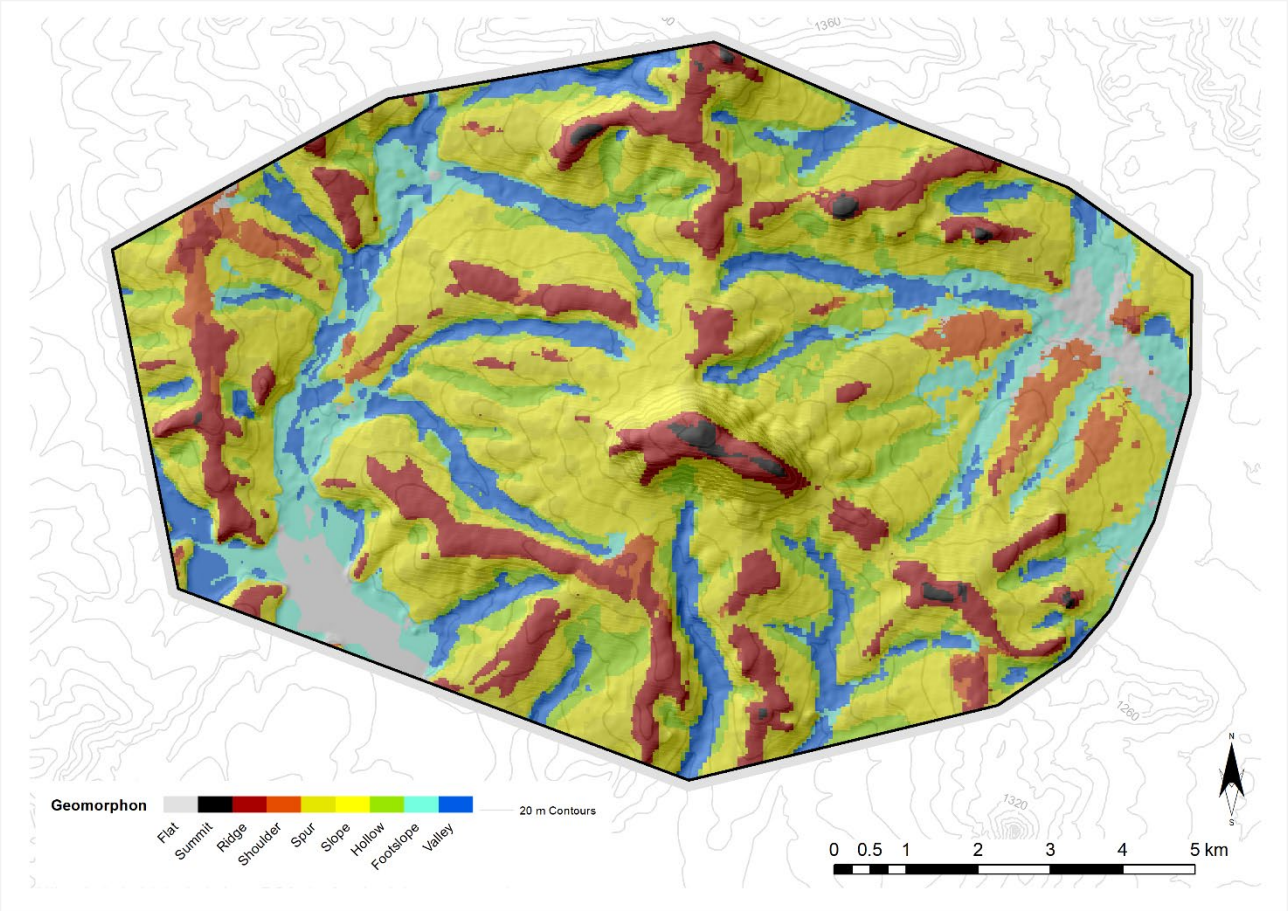


Figure 3.4: Reference map of geomorphons for the Bergville study region generated from the 30 m SUDEM.

What is unequalled in this study is that it has shown that while geomorphon features are detectable across most DEM surfaces of varying resolution, the accuracy of the geomorphon characterisation at a regional scale is not linearly correlated with pixel resolution, i.e., finer resolution DEMs do not necessarily produce "improved" geomorphon surfaces. In fact, improved geomorphon representation may not translate into enhanced soil-landscape applications. Geomorphon surfaces should be applied and interpreted with the area's conditions under observation and the analyses' objectives in question. Contextually, the use of GIS to derive reliable estimates of terrain features and the influence of DEM resolution on soil-landscape measurement and representation accuracy is well established. Numerous studies have addressed, *inter alia*, issues such as data precision (Theobald, 1989), grid resolution or grid size (Chang & Tsai, 1991; Garbrecht & Martz, 1994; Hodgson, 1995; Florinsky & Kuryakova, 2000; Tang, 2000) as well as grid orientation (Jones, 1998), influences of slope algorithms (Dunn & Hickey, 1998) and more recently computational approaches (Warren *et al.*, 2004) and DEM uncertainty and error propagation (Goulden *et al.*, 2016). The consensus is that as pixel resolution of the DEM surface decreases, DEM terrain attribute metrics are underestimated, typically leading to a misrepresentation of terrain features at coarser resolutions. However, in this study, while the 90 m DEM surface products may not be entirely statistically similar to the 30 m reference surface, they are also not entirely inaccurate in their representation of specific morphological features present in the study site. Stated differently, the 30 m (except for the GDEM2) and 90 m geomorphon products represent the region depending on the landscape analysis context. To represent generalised terrain features, the use of coarser DEMs helps reduce the small-scale variation in terrain properties and better define the broad landscape pattern (Stein *et al.*, 2001).

In contrast, if the objective is to locate specific terrain features (hot spots), DEM aggregation should be avoided favouring higher resolution surface models such as the SUDEM 30 m product. Adjunct to this discussion on optimal pixel resolution, Liang *et al.* (2004) observed impacts of different spatial resolutions on modelling surface runoff. They concluded that spatial resolution could be improved only to a critical level, after which no substantial improvements on model performance will be observed. So, while grid resolution plays a vital role in the efficiency of the mapping geomorphons, its contribution can only be optimised to a certain level to satisfy both processing capabilities and representation of spatial variability. Likewise, Silva *et al.* (2016a) highlighted that pixel size alone and not threshold *LOS* values had a more considerable influence on the chi-squared values of geomorphon surfaces and soil type phenomena. Hengl (2006) corroborated this pixel-specific concept with similar findings suggesting that the choice of optimal pixel size may need to differ according to different target terrain variables investigated.

In fact, Gruber *et al.* (2015) explicitly outline the need to adapt model thresholds or membership functions according to the topographic detail of an area for better terrain differentiation. A key finding from their study encapsulates the dichotomy and complexity of terrain detail *vs* membership function *vs* pixel resolution considerations. Their study found that, in general, macro-scale terrain classification maps could replicate a larger number of topographic position classes (*quantitative*). In contrast, the mesoscale models contained fewer types but were more similar (*qualitative*) to a reference surface. For this study, if one considers the frequency distribution of the geomorphons across the site (**Figure 3.4**), we see a similar pattern that in general, the 90 m products have a higher frequency of morphon features but a generally lower similarity (accuracy) to the reference 30 m SUDEM surface (**Table 3.3 to 3.5**). Given the spatial modelling efficiency of the geomorphon approach, It is not surprising then to find that the study by Gruber *et al.* (2015) recommends that geomorphons are better suited to macro-scale applications. To implicitly evaluate the utility and scale-sensitivity of geomorphons for soil-landscape assessments, I aimed to assess how the predicted geomorphon terrain features were measured against specific qualitative descriptions for well-accepted soil-landscape relationships and terrain properties for the study region.

3.3.2 Geomorphon area frequency distribution by DEM surface

The ability to quantify the overall similarity between different landscapes by merely evaluating the descriptive data of the associated and derived surfaces is a critical element of the present methodology. **Table 3.6** to **Table 3.9** outline the summary statistic results for the respective resolution-defined geomorphon surface products, including the descriptive statistics for *Elev*, *SlopGrad*, and geomorphon feature area defined in hectares (Ha). By comparing the point-sample count for each geomorphon surface, interpreted here as a quasi-indicator for in-field sample survey frequency, the dominant (count) and largest (area) terrain geomorphon identified in all DEM surfaces were *slope* features. A key observation is that for selected geomorphon features, the areas derived from the coarser-resolution 90 m SUDEM and 90 m SRTM surfaces are consistently more extensive than for the same geomorphon features derived from the reference 30 m SUDEM surface. Furthermore, the results depict a higher intra-class similarity of geomorphon properties within rather than between DEM surface resolution. Specifically, geomorphon features derived from the 90 m SUDEM and 90 m SRTM DEM surfaces are like those derived from the reference surface (30 m SUDEM). Put differently, the DEM generalisation from 30 to 90 m provides better geomorphon representation than the 30 m GDEM2 surface. The utility of the 90 m geomorphon products to represent the reference surface better than the GDEM 30 m surface is also evident in the lower-lying terrain features such as *foot slopes* and *valley bottoms* and certain higher altitude features such as *summits, ridges and shoulders*. These results are not surprising, considering that lower

resolution DEMs have been shown to typically not be suited for detecting finer resolution terrain features (Cavazzi *et al.*, 2013).

Interestingly, similar results were observed in separate studies by Libohova *et al.* (2016) and Silva *et al.* (2016b). Jasiewicz & Netzel (2014) provide the necessary context for the observed trends in this study, relating the observations to the computational efficiency (or limitations) of the geomorphon approach to ternary pattern characterisation. While the final geomorphon model-outputs represent the ten most common and recognisable terrain features, there are many more ternary elements and associated patterns, many of which are also considered abnormalities, calculated within a given feature space. It is also evident that there is an uneven distribution of distinct geomorphons; the 30 most common geomorphons constitute 85% of all cells. Select studies have reported that as many as 6 561 theoretical ternary patterns can be identified within a region (Jasiewicz & Stepinski, 2013). Many of these patterns are, however, rotational or reflection variations of other patterns. The geomorphon model does optimise this pattern-redundancy to the final sequence of approximately 500 theoretical patterns for computational efficiency.

Nevertheless, for geomorphometric mapping, it is necessary to reduce the number of landform elements. Essentially, this is accomplished by dividing geomorphons into classes (many-to-one) that correspond to the most commonly apparent landform features. Multiple terrain features may systematically be forced into one of the archetypical ten unique geomorphon classes, and therefore the misclassification of cells due to uncertainty can be expected. The landscape terrain consists of a series of landform elements, i.e., *valley*, *slope, ridge, foot slope* etc., and the nature of the land surfaces dictate the quality of the landscape patterns. The dominant elementary forms are often *flat* areas and *slopes* in higher relief areas (Jasiewicz *et al.*, 2014). The remaining terrain elements such as *peaks, hollows, spurs* and *shoulder*s are less abundant and therefore often do not dominate the frequency distribution of the geomorphon features in the landscape. Moreover, it is possible to have multiple ternary elements and different ternary patterns representing other instances of the same morphon class, i.e., *slope, valley* etc., within and between locations (Jasiewicz & Stepinski, 2013). The less constrained the rule-criteria are for allocating the ternary patterns to the morphon instances (combinations). The more uniform the landscape is, the higher the frequency of features assigned to a typical class and the higher the likelihood of Type I and II classification errors.

In contrast, the accurate detection of specific (and less common when compared to *slopes* and *valleys*) landform features such as *pits/ depressions and summits/ peaks* require a single exact geomorphon to be allocated to the final 10 class classified map. Typically, there are fewer instances or interpretations of these less-common ternary patterns before final classification, generally resulting in a far lower computational frequency of these features within the final 10-class geomorphon surfaces. Jasiewicz et al. (2014) recommend using a histogram as an option for rapidly assessing the terrain-pattern "primitive feature" signature because of its rotational invariance, i.e., two patterns rotated in respect to each other will be identical. Primitive features are simple measures designed to provide small, local pieces of information about the landscape pattern. Typically, a geomorphon element at a given cell could be a primitive feature, and a histogram of all geomorphon elements over the landscape could be the landscape signature. **Figure 3.5** depicts the histogram of geomorphon terrain features vs feature area for each observed DEM surface within the Bergville study site.

Table 3.6: Descriptive statistics for DEM elevation, slope and feature area for 30 m SUDEM reference Geomorphon surface.

	SUDEM 30	Height (m)						Slope (%)						Area
	Count	**Min**	**Max**	Mean*	SD**	CV***	SE****	**Min**	**Max**	**Mean**	**SD**	**CV**	**SE**	**Ha**
Flat	65	1 169	1 255	1 181	16.2	1.4	2.0	0	6	2	1.6	65.3	0.2	145
Summit	30	1 244	1 472	1 379	80.7	5.9	14.7	0	19	6	4.2	68.4	0.8	70
Ridge	645	1 182	1 462	1 278	56.6	4.4	2.2	0	55	7	7.3	101.4	0.3	1 450
Shoulder	168	1 176	1 292	1 225	31.8	2.6	2.5	0	15	3	1.9	53.4	0.1	392
Spur	752	1 178	1 445	1 259	44.1	3.5	1.6	0	60	9	7.5	88.3	0.3	1 670
Slope	1 598	1 172	1 417	1 247	40.1	3.2	1.0	0	55	8	5.5	73.1	0.1	3 600
Hollow	609	1 178	1 375	1 246	31.1	2.5	1.3	1	52	7	4.2	58.3	0.2	1 359
Foot	348	1 168	1 292	1 201	26.9	2.2	1.4	0	11	3	2.2	62.0	0.1	796
Valley	534	1 169	1 335	1 230	32.9	2.7	1.4	0	20	6	3.5	62.3	0.2	1 233
Pit	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3.7: Descriptive statistics for DEM elevation, slope, and feature area for the 30 m ASTER GDEM2 predicated Geomorphon surface.

	GDEM2	Height (m)						Slope (%)						Area
	Count	**Min**	**Max**	**Mean**	**SD**	**CV**	**SE**	**Min**	**Max**	**Mean**	**SD**	**CV**	**SE**	**Ha**
Flat	2	1 170	1 176	1 173	4.2	0.4	3.0	9	29	19	14.0	73.6	9.9	4
Summit	126	1 173	1 470	1 287	67.1	5.2	6.0	1	47	14	9.2	64.1	0.8	270
Ridge	831	1 168	1 469	1 266	56.4	4.5	2.0	0	86	13	10.2	77.3	0.4	1 885
Shoulder	68	1 164	1 386	1 242	51.4	4.1	6.2	1	47	14	7.5	54.5	0.9	144
Spur	778	1 164	1 432	1 256	47.5	3.8	1.7	0	87	14	9.6	69.8	0.3	1 770
Slope	1 271	1 159	1 436	1 248	44.1	3.5	1.2	0	66	13	8.0	63.8	0.2	2 839
Hollow	664	1 163	1 367	1 240	36.4	2.9	1.4	1	61	12	7.6	62.1	0.3	1 480
Foot	102	1 161	1 313	1 222	31.4	2.6	3.1	1	33	12	6.8	56.9	0.7	240
Valley	830	1 150	1 328	1 225	33.6	2.7	1.2	0	50	12	6.8	58.5	0.2	1 928
Pit	77	1 154	1 267	1 197	28.9	2.4	3.3	2	33	13	7.7	58.4	0.9	155

Table 3.8: Descriptive statistics for DEM elevation, slope and feature area for the 90 m SRTM predicated Geomorphon surface

	SRTM 90	Height (m)						Slope (%)						Area
	Count	Min	Max	Mean	SD	CV	SE	Min	Max	Mean	SD	CV	SE	Ha
Flat	21	1 172	1 213	1 183	11.4	1.0	2.5	0	10	4	2.5	63.9	0.5	44
Summit	79	1 295	1 476	1 416	56.7	4.0	6.4	0	45	16	13.9	87.7	1.6	172
Ridge	828	1 194	1 419	1 277	42.3	3.3	1.5	0	58	8	7.2	96.3	0.3	1 887
Shoulder	112	1 188	1 302	1 243	32.0	2.6	3.0	1	16	5	2.5	50.3	0.2	242
Spur	653	1 197	1 420	1 268	36.9	2.9	1.4	0	47	8	6.4	77.4	0.3	1 489
Slope	1 209	1 184	1 398	1 250	33.8	2.7	1.0	1	51	7	4.0	58.9	0.1	2 705
Hollow	521	1 190	1 317	1 246	25.5	2.0	1.1	1	20	6	2.9	46.8	0.1	1 156
Foot	485	1 166	1 288	1 199	23.8	2.0	1.1	0	25	5	2.9	63.8	0.1	1 082
Valley	841	1 168	1 323	1 223	30.2	2.5	1.0	0	18	5	3.1	60.3	0.1	1 938
Pit	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3.9: Descriptive statistics for DEM elevation, slope, and feature area for 90 m SUDEM predicted Geomorphon surface.

	SUDEM 90	Height (m)						Slope (%)						Area
	Count	Min	Max	Mean	SD	CV	SE	Min	Max	Mean	SD	CV	SE	Ha
Flat	98	1169	1250	1185	14.6	1.2	1.5	1	14	4	2.4	57.5	0.2	228
Summit	69	1296	1472	1418	59.3	4.2	7.1	0	45	16	13.7	83.7	1.7	155
Ridge	787	1193	1426	1278	46.0	3.6	1.6	0	52	8	8.0	99.1	0.3	1776
Shoulder	234	1187	1346	1241	30.8	2.5	2.0	0	21	6	3.5	60.3	0.2	557
Spur	623	1193	1387	1270	34.4	2.7	1.4	0	41	8	5.7	71.9	0.2	1362
Slope	1156	1185	1357	1249	32.3	2.6	1.0	1	36	7	3.6	53.5	0.1	2614
Hollow	453	1191	1324	1248	25.6	2.1	1.2	2	17	6	2.5	41.0	0.1	1020
Foot	628	1168	1288	1204	24.7	2.1	1.0	0	24	5	2.8	60.6	0.1	1417
Valley	701	1170	1300	1221	27.0	2.2	1.0	0	21	5	2.6	56.3	0.1	1586
Pit	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**Mean: Trimmed Mean; **SD: Standard Deviation; ***CV: Coefficient of Variation; ****SE: Standard Error of the Mean*

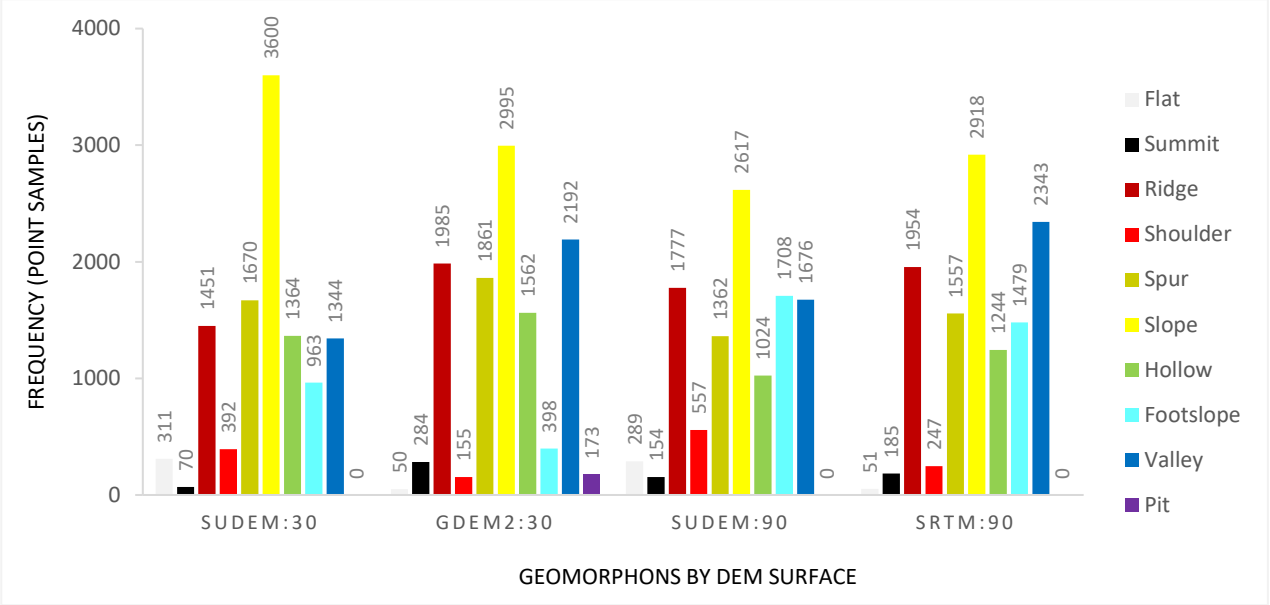


Figure 3.5: Histogram of geomorphon frequency for the SUDEM 30, GDEM2 30, SUDEM 90 and SRTM 90 DEM. With a designed sampling strategy of 1 point per ha, the values above each geomorphon feature are taken to represent the surface area (ha) of each feature.

By analysing the histogram patterns (landscape signature) for each DEM, we observe that the reference SUDEM 30 is the only geomorphon surface dominated by a *slope-spur-ridge-hollow* pattern. The remaining DEM surfaces all display similar patterns with minor variations. The SRTM 90 m: *slope-valley-ridge-spur*, SUDEM 90 m: *slope-ridge-valley-footslope* and the GDEM2 30 m: *slope-valley-ridge-spur*. Despite the variability between DEM geomorphon feature frequency and extent, the histogram indicates that all DEM surfaces can define the same general landscape pattern, i.e., higher altitudes, i.e., *ridges, spurs*, and lower-lying features like *valleys* and *foot slopes.* Although derived from the same 5 m DsM, the two SUDEM surfaces do not show as much geomorphon feature similarity as expected. This is evident because almost 3 500 ha of *slope* geomorphon area is classified in the 30 m SUDEM surface in contrast to the 2 500 ha in the 90 m SUDEM product. This reaffirms that the influence of pixel resolution and pixel generalisation on the detection and classification of certain features in the landscape must still be acknowledged in geomorphon surface analysis (Thompson *et al.*, 2001; Atkinson *et al.*, 2017). Astonishingly, the GDEM2 geomorphon results are more comparable to the SRTM 90 surface than the SUDEM 30 surface despite sharing the same pixel resolution as the former. However, the GDEM2 histogram also reveals certain geomorphon surface trends that suggest inherent spurious anomalies in the DEM surface. For instance, the GDEM2 is the only surface that has identified approximately 200 ha of *pit* geomorphon features while only classifying 50 ha as *flat* areas and further classifying approximately 2 200 ha of *valley* features in the study site. As a baseline visual approximation of the geomorphon surfaces, the histogram plot provides the basic synoptic overview to confirm that the quality of the GDEM2 for this region may inherently be compromised for terrain-based assessment and feature extraction, including soil-landscape based assessments.

To further evaluate the extent of the surface similarity and dissimilarity between geomorphon features, it is necessary to compare the results for *Elev* and *SlopGrad* derived for each of the four DEM surfaces. The following section provides a valuable synopsis of relevant findings.

3.3.3 Geomorphon feature description based on selected DEM properties

3.3.3.1 DEM elevation

The summary statistic results for the *Elev* and *SlopGrad* values for each of the four DEM surfaces are presented in **Table 3.6** to **3.9** and segmented by the 10-class geomorphon types. Combined with the box and whisker plot (**Figure 3.6**), these results confirm that most geomorphon features positively correlate to the reference SUDEM 30. While the distributions confirm the presence of variable outliers for each geomorphon component, all DEM surfaces demonstrate a high correlation of the *trimmed mean* based on sample point count per geomorphon feature. Given these results, the choice of DEM surface and spatial resolution may appear to have very little influence on the respective DEM geomorphon features' elevation differentiation. However, while altitude is the "pre-cursor" terrain-quality measure for most DEM analyses, users are cautioned in merely using the *Elev* values as the single proxy for terrain similarity, despite the apparent visual correlations. Gonga-Saholiariliva *et al.* (2011) point out that simple global metrics, such as *Elev,* may conceal substantial local variability between data sets.

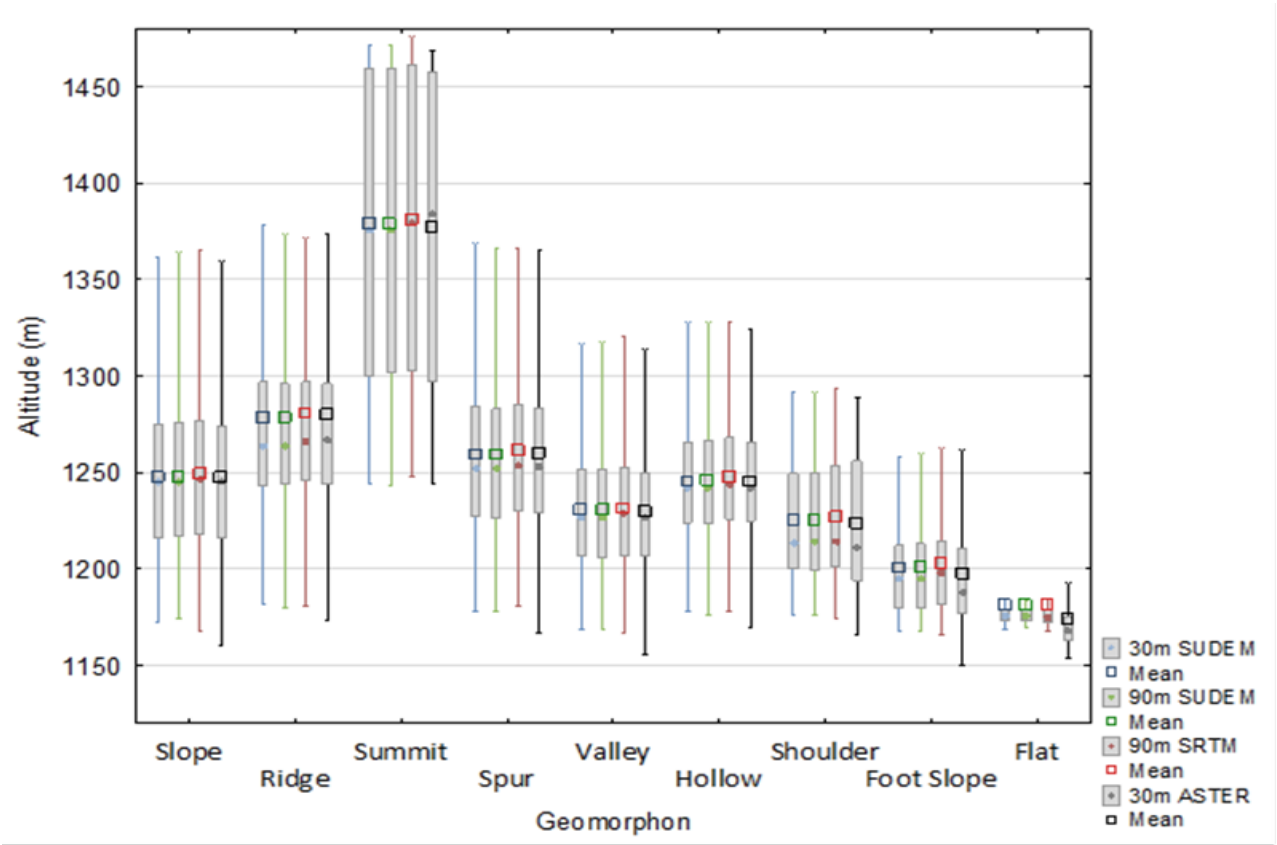


Figure 3.6: Box and Whisker plot showing the elevation distribution for each Geomorphon feature derived from the four DEM for the study site. Note the high inter-correlation between DEM surfaces for each of the geomorphon features (x-axis) for the respective altitude range (y-axis).

Alternative methods may be necessary to provide further insight into the variability between surface products. Importantly, their study offers an additional explanation for these trends by suggesting that even though the altitudinal error is present in the various DEMs, including outliers, neither data scatter in linear regression models nor variance in the vertical accuracy constitute sufficient criteria to achieve a full definition of DEM error because the error is expressed regionally, i.e., across the entire DEM surface. Atkinson *et al.* (2017) were also able to explore global measures of elevation similarity at varying DEM resolutions and generalisation approaches. Their study similarly cautioned on determining DEM product similarity based solely on central estimates of tendency and better quantified the variability in terrain feature extraction by introducing altitudinal breaks into the terrain extraction.

Notably, the findings on the above-observed phenomena are not unique and compare well with Barber & Shortridge (2005) and Luo *et al.* (2012). The high overall standard deviation (*SD*) values for each DEM further reiterate that while the inter-feature *Elev* variability may not appear to be dissimilar across DEM surfaces and between morphon features, the intra-variation of *Elev* within each class is relatively high, indicating high variability from the *mean trimmed* value. When taken in conjunction with the standard error (*SE*) values, we see that certain higher relief features, such as the *summit* and *ridge* features, show the highest variation (error) in elevation and considerably lower reliability and repeatability in elevation estimates across all DEMs than the other geomorphon features. As expected, all the 30 m DEMs' *SD* values were generally higher than the 90 m DEMs. The higher resolution surfaces are expected to capture more terrain detail and, therefore, more variability than their coarser, more generalised counterparts. Moreover, this is most evident in the SUDEM 30 DEM, possibly because the SUDEM is supplementary-derived from high-resolution national 5 m (Van Niekerk, 2012) contour data. The GDEM2 surface exhibited the most extensive range (max-min) in *Elev* values across most morphon features with notable mentions of range-values between 290 and 300 m for higher elevation features such as *summits* and *ridges* respectively and 150 to 200 m for flatter, lower-lying relief features such as *foot slopes* and *hollows*. Elevation range could be considered a distal factor for relative relief and terrain feature morphology with lower altitudes characterised by less terrain variation. In comparison, more broken terrain features could represent higher altitudes. However, further interpretation of the landscape, in terms of *SlopGrad*, is necessary to understand if this variability is authentic or due to DEM surface generalisation error (Atkinson *et al.*, 2017).

3.3.3.2 Slope gradient

It is well known that DEM accuracy has a strong correlation with local morphometric variables. Fisher (1998) found a significant correlation between DEM errors and slope angle, while Gao (1997) reported that DEM errors are typically lower in less complex terrain. Therefore, the *SlopGrad* is more sensitive than *Elev* to local and global terrain variation (Wilson, 2000). From the results presented in **Table 3.6** to **3.9,** we see that *SlopGrad* values show more variability than *Elev* between the four DEM surfaces. Even so, the derived *SlopGrad* values for the four DEM products were well aligned to widely accepted localised-conceptualisations of *SlopGrad* values for original landscape features for the region (Smith, 2006). These findings are promising for soil-landscape applications since they advocate that the categorisation of the modelled geomorphic features by *SlopGrad* remains consistent despite the differences in pixel resolution (grain size) of the DEM sensor platform. The exception appears to manifest with higher altitude, complex terrain features such as *ridges*, *spurs* and *summits,* indicating higher *SlopeGrad SDs* for the SUDEM 30 m, SUDEM 90 m and SRTM 90 DEMs, respectively. This confirms the findings by Guth (1995), who reported that concerning modelling low or high relief terrain, the most considerable differences in calculated gradients occur in lower-lying terrain features such as *pits*, *peaks*, *valleys* and *ridges.*

Notwithstanding, to further endorse the GDEM2 elevation data's spurious findings, the higher than expected *mean SlopeGrad* and *SD* values observed with selected GDEM2 geomorphons confirm the likelihood of contemporaneous DEM error in the surface model. Notably, all GDEM2-identified *flat* features were reported to have a *mean SlopGrad* of 19% and a maximum slope of 29% at low altitudes (1 173 m). In contrast, the SUDEM and SRTM surfaces typically all displayed relatively small *mean SlopGrad* values of approximately 13% at higher altitudes (1 266 m) for steep terrain features such as *spurs*, *slopes* and *ridges.*

Soils do not exist in isolation but are organised within the landscape. In practice, the primary functional relationships of factors, processes, and soil properties enable the recognition and delineation of unique combinations as soil-landscapes (Grunwald, 2016). *SlopeGrad* and soil properties are strongly autocorrelated (Gerrard, 1981), and landscape configuration features can be readily defined by contemporary geomorphological and pedogenic processes. This has led to the realisation that slope forms are associated with specific soil sequences, giving rise to the catena concept (Milne, 1936) as a unit of "mapping convenience". Therefore, practitioners adopt conceptual soil topo-sequence models to enhance the utility of technical soil information (Shahid *et al.*, 2013). A soil topo-sequence shows a catena (or cross-section thereof) of soils and their endogenous relationship in the landscape where soils may differ from one region to another. There is a succession of soil physical and chemical properties associated with varying slope characterisations (Milne, 1935). Accordingly, topo-sequence properties represent a composite of terrain characteristics, i.e., profile curvature, slope gradient, and relative elevation, that allow experts to develop a mental model, albeit subjectively, of soil patterns and properties in the landscape to synthesize and simplify otherwise complex soilscape relationships in-field "on-the-fly" (Milne, 1936; Bushnell, 1943).

So, it is essential that when calibrating the geomorphic/ topo-sequence heterogeneity of derived models, these generalised soil-landscape relationships correspond with well-defined and widely accepted interpretive catenary delineations. These defined relationships should be aligned with knowledge-driven expressions of qualitative mental models of pedogenesis, whether explicit (soil properties) or implicit (soil classes), particularly for South African conditions (Van den Bergh *et al.*, 2009). The summary results for the GDEM2 geomorphons presented in **Table 3.7** are not aligned to accepted generalised soil-landscape patterns for the region, which may present a challenge in defining the utility of the GDEM2 product for other geomorphon soilscape applications in KwaZulu-Natal. Typically, *upper slope* positions (15 to 35% gradient) are regarded as zones of divergence with active denuding and transportation of materials resulting in shallower, well-drained soils. In comparison, near-level (1 to 3%) and lower *slope* positions (4 to 10% gradient) are defined as zones of convergence characterised by the deposition and accumulation of materials with soil-landscape properties associated with solid evidence of soil hydromorphism (Skidmore *et al.*, 1991). Understanding the contemporary erosional and hydrological succession of soil properties within various slope classes is necessary for guiding the description and segmentation of the geomorphon soil-scape "fingerprint" for each DEM surface. Constraints on accurate surface representation using the GDEM2 naturally pose challenges to the credibility of any derived geomorphon to predict soil-landscape properties across the study site.

To implicitly evaluate the utility and scale-sensitivity of geomorphons for soil-landscape assessments, the I aimed to assess how the predicted geomorphon terrain features were measured against specific qualitative descriptions for well-accepted soil-landscape relationships and terrain properties for the study region. The influence of DEM quality on *SlopGrad* was visually emphasised by calculating a terrain profile for each of the DEM surfaces using Global Mapper v20.1.2 (www.BlueMarble.com, 2019) and evaluating the surface properties along a given transect to determine the level of similarity in selected terrain properties.

Interpreting the cross-sectional ground profile results in **Figure 3.7 to 3.10,** we see that the SUDEM 30 m, SUDEM 90 m and SRTM 90 m DEM visually exhibit similar DEM surfaces, the notable difference being that the SUDEM 30 m can provide greater DEM detail for certain terrain features. In contrast, both 90 m DEMs provide subtle generalisations of the terrain and *SlopGrad.* While simplistic, the cross-sectional terrain profiles provide a practical, often overlooked, visual context of the extent of inherent surface variation in specific DEM surface properties. A detailed comparison of DEM terrain attributes are presented in **Table 3.10** and include descriptions of *3D terrain length, vertical terrain difference; min elevation; max elevation* and *max slope* within a north/south transect through the study site. By comparing the terrain profiles for each of the DEM characters, it is clear that the GDEM2 surface is riddled with surface anomalies that manifest as excessive spikes and pits. Undoubtedly then, this would explain the cause for the sub-optimal geomorphon feature characterisation. These findings corroborate well with Rexer & Hirt (2014), who found similar results when conducting accuracy assessment studies of the GDEM2 and SRTM products in Australia. Their findings reported that the major drawback of GDEM2 for landscape analysis is that being an optical sensor, data collection may be impaired by cloud cover over certain areas leading to data anomalies such as voids (holes) or artefacts in the GDEM2.

The irregularities in the GDEM2 can be attributed to the fractured and spurious surface and how these anomalies manifest as excessive spikes and pits. Furthermore, the GDEM2 detects “top of surface” elevation for features on the earth, including buildings and tree canopies, and do not reflect the bare ground or actual terrain elevation (Meyer, 2011; Robinson *et al.*, 2014). Compared to other similar-derived DEM surfaces, Rexer & Hirt (2014) generally reported higher GDEM2 surface offsets exceeding even SRTM elevations in certain areas with negative offsets observed over low- or non-vegetated areas.

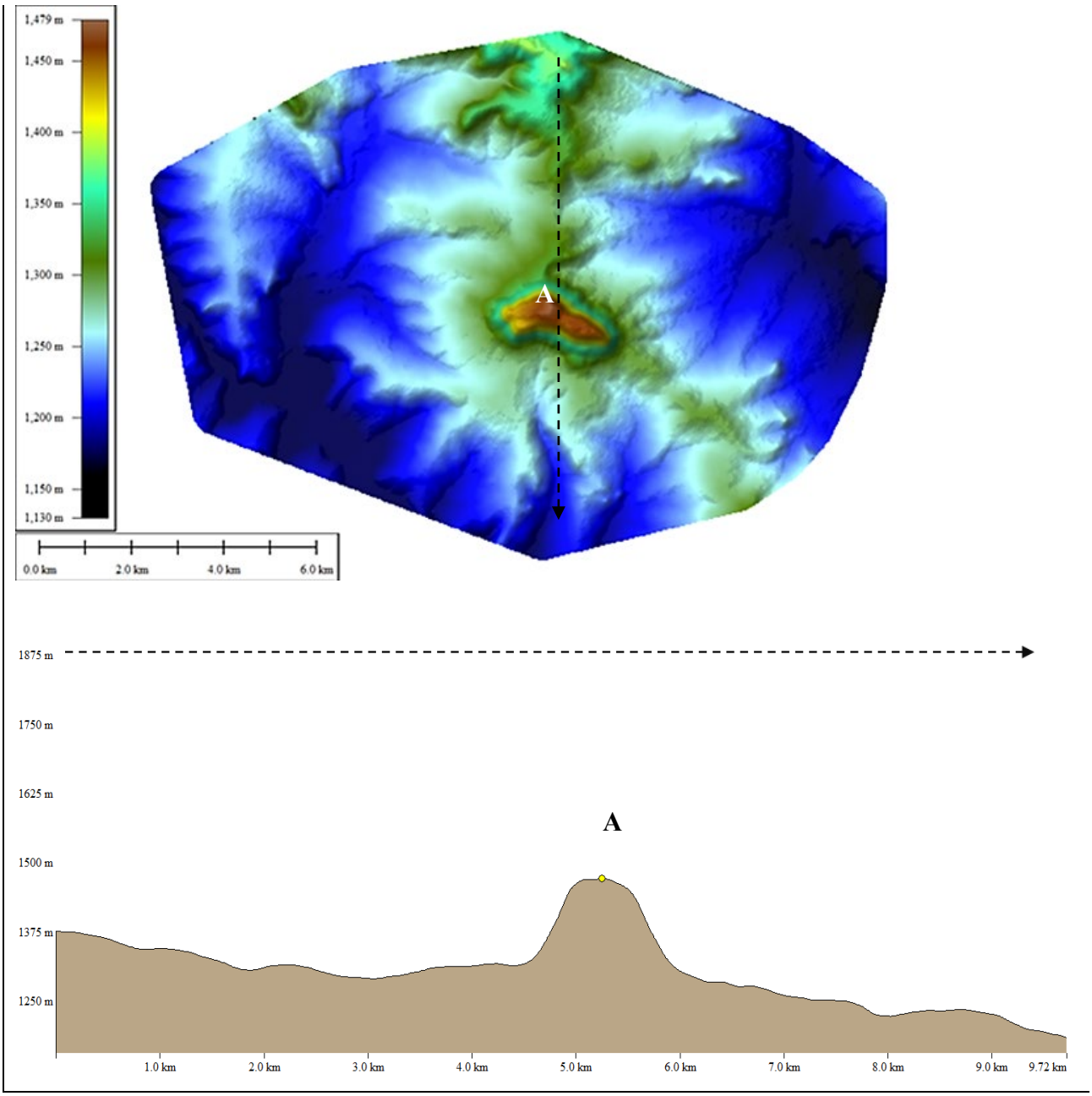


Figure 3.7: North-South 10 km cross-sectional transect of Bergville study site. The DEM image represents the elevation for the 30 m SUDEM, with the transect represented by the black dashed line. Point A represents a prominent landform feature (mesa) in the study site. The 2D profile outlines the 9.7 km transect terrain profile showing variation in elevation and slope gradient.

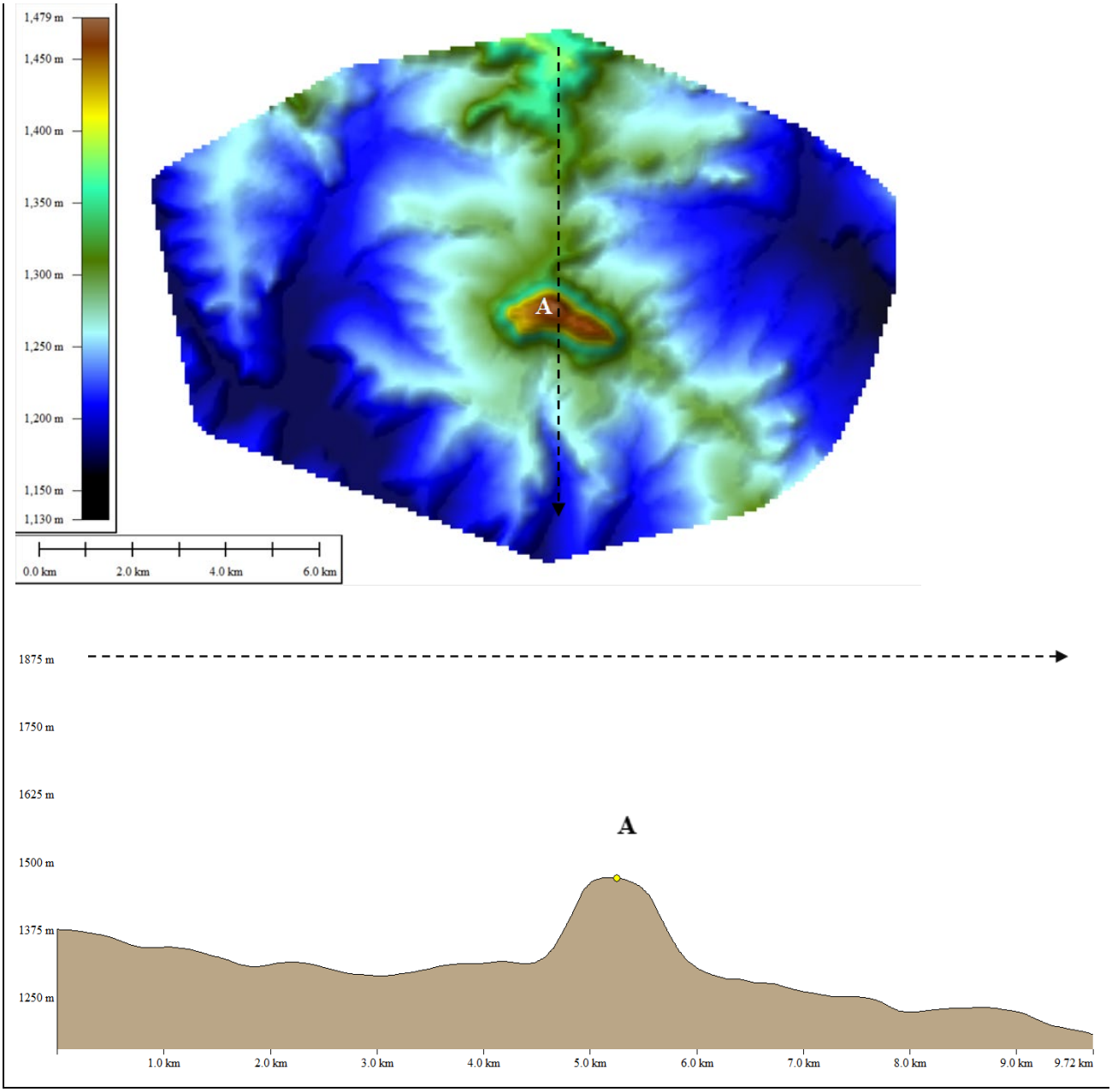


Figure 3.8: North-South 10 km cross-sectional transect of Bergville study site. The DEM image represents the elevation for the 90 m SUDEM, with the transect represented by the black dashed line. Point A represents a prominent landform feature (mesa) in the study site. The 2D profile outlines the 9.7 km transect terrain profile showing variation in elevation and slope gradient.

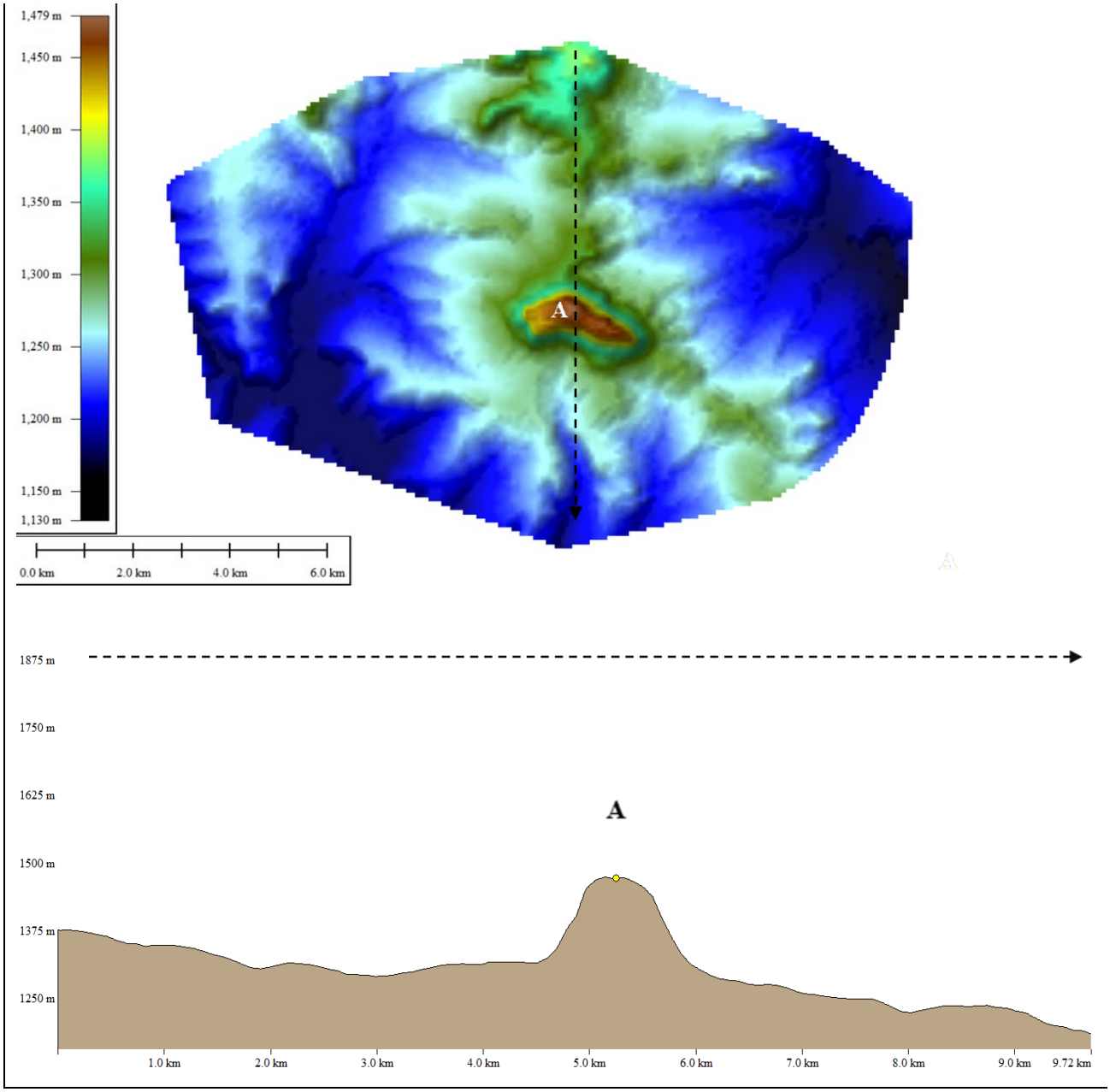


Figure 3.9: North-South 10 km cross-sectional transect of Bergville study site. The DEM image represents the elevation for the 90 m SRTM, with the transect represented by the black dashed line. Point A represents a prominent landform feature (mesa) in the study site. The 2D profile outlines the 9.7 km transect terrain profile showing variation in elevation and slope gradient.

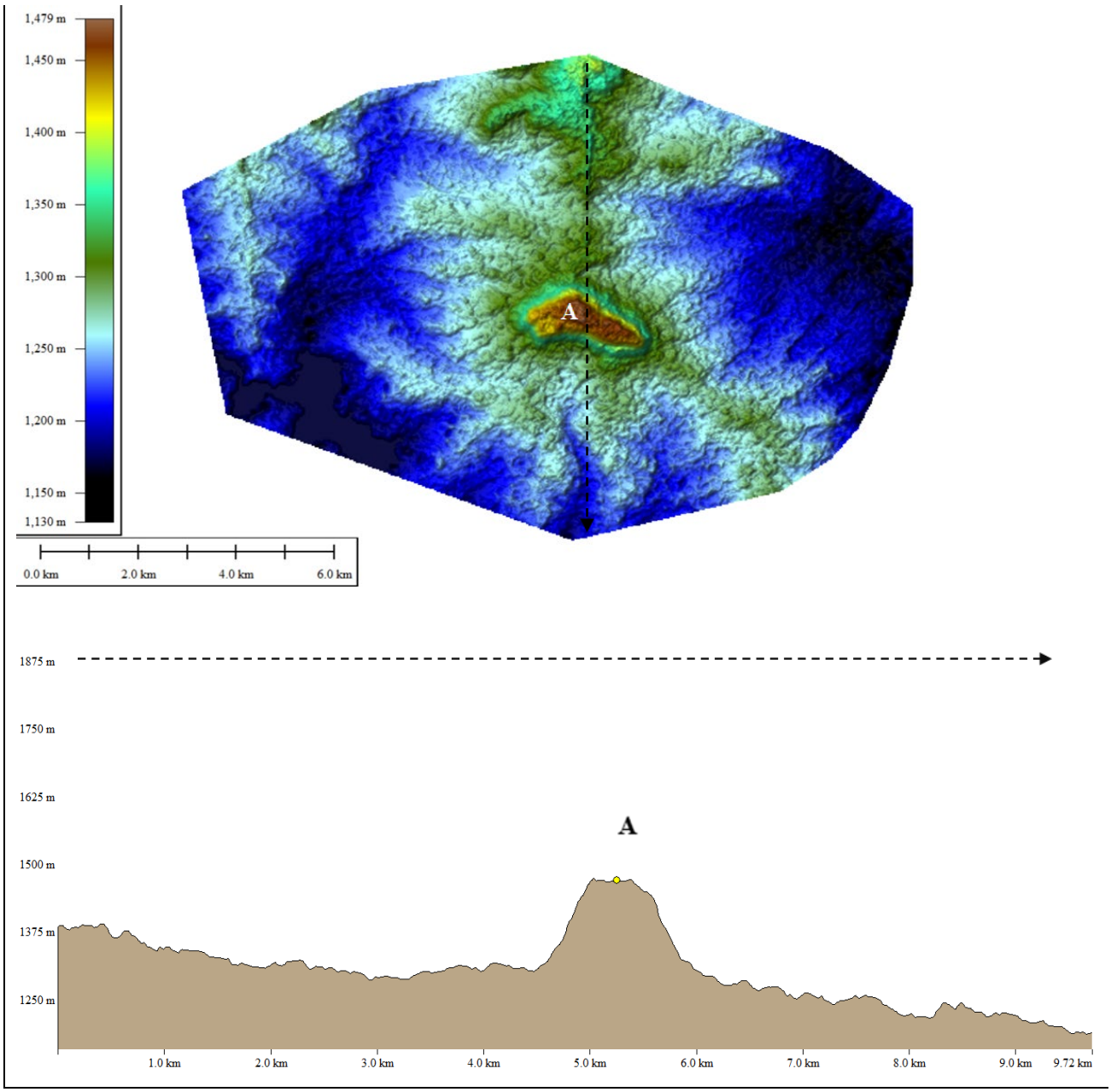


Figure 3.10: North-South 10 km cross-sectional transect of Bergville study site. The DEM image represents the elevation for the 30 m ASTER GDEM2 with the transect represented by the black dashed line. Point A represents a prominent landform feature (mesa) in the study site. The 2D profile outlines the 9.7 km transect terrain profile showing variation in elevation and slope gradient.

The DEM surface attributes presented in **Table 3.10** introduce several interesting findings of note. Firstly, the 3D surface length for the SUDEM 90 and SRTM 90 differs between 80 and 130 m, respectively, from the 3D surface length of the "reference" SUDEM 30. The marginal variation of the coarse resolution DEM surface(s) to the reference surface is remarkable, given that the SRTM DEM is also known to contain spurious terrain anomalies. Yet, these have not manifested within the digitised transect. By contrast, the GDEM2 differs by a distance of 450 m to the reference SUDEM 30 m. It is necessary to point out that all DEM surfaces share the same transect straight-line distance of 9.7 km. So variability in the 3D length is attributed to irregularities in DEM surface properties. The extent of this surface variation between the GDEM2 surface and the SUDEM 30 surface is not inconsequential.

Soil-landscape mapping and DSM analysis require accurate measures of terrain surface attributes given that ground distance(s) estimates translate into the estimation of terrain-feature length, perimeter and surface area. The influence of DEM vertical accuracy on horizontal displacement (HD) has been described earlier in detail *(§ 2.3.4)*. Both Wechsler & Kroll (2006) and Gillin & McGuire (2013) have highlighted the impact of DEM error on distance calculations primarily related to the modelling of selected soil properties and the influence on upslope catchment area boundary analysis.

Table 3.10: 2D and 3D surface attributes for the selected DEM surfaces for the given north/south 9.7 km transect.

	SUDEM 30	SUDEM 90	SRTM 90	ASTER GDEM2
Start Height (m)	1376	1375	1375	1384
End Height (m)	1183	1182	1184	1189
Straight-Line Distance (km)	9.72	9.72	9.72	9.72
3D Distance on Surface (km)	9.85	9.93	9.98	10.30
Vertical Difference (m)	-192.5	-193	-191.5	-195.1
Minimum Elevation on Path (m)	1183	1182	1184	1186
Maximum Elevation on Path (m)	1470	1470	1473	1474
Max Path Slope (Deg.)	25.82° [4.864 km along path]	23.59° [4.864 km along path]	29.11° [4.883 km along path]	39.39° [5.625 km along path]

It is expected then that the considerable variation in ground distance estimation between the GDEM2 and SUDEM 30 m will undoubtedly influence the discretisation of geomorphon feature boundaries and associated representation of soilscape relationships in the study site. The DEM errors observed in the GDEM2 surfaces are likely due to a combined landcover and random/non-random error phenomenon (Tachikawa *et al.*, 2011). Oksanen & Sarjakoski (2005) provide a detailed explanation of how DEM errors influence terrain analyses. The computation of DEM derivatives can be divided into primary topographic variables, which can either be calculated directly from elevation or obtained from its derivatives, such as *SlopeGrad, SlopeAsp, curvature*, and *catchment area*. Secondary variables such as *TWI* and *stream power* index are calculated by combining two or more primary variables. These secondary attributes are essential to DSM applications since they provide an opportunity to describe terrain patterns as a function of process and directly influence soil characteristics, abundance and distribution in the landscape (Wilson & Gallant, 2000).

Systematic and non-systematic DEM errors may also be amplified when second-order derivations, such as *SlopGrad*, are calculated. The results outline that slope values and 3D surface attributes exhibit significant error, while the elevation comparisons between the GDEM2 and reference SUDEM surface are less pronounced. Since secondary attributes are more sensitive to DEM error, Oksanen & Sarjakoski (2005) recommend that error-propagation analysis be based on classification methods that focus on the spatial properties rather than only the source of surface derivative calculation.

Given the performance of the ASTER GDEM2 in regards to *SlopGrad*, *Elev* and geomorphon feature characterisation, users would need to determine whether to proceed with the application of a sub-optimal DEM or attempt to pre-process the DEM further in an attempt to reduce the occurrence of artefacts in the DEM or completely disregard the DEM surface altogether.

The major caveat that many practitioners or "purists" would need to reconcile is that while the surface quality of the GDEM2 for this region is compromised, which limits the objective integrity of any statistical comparisons with the other DEM products, it is equally unjust to out-right disregard the DEM as entirely irrelevant for future DSM and DGM applications in the region. When considered holistically as a global DEM product, the GDEM2 elevation results show a high correlation to the reference 30 m SUDEM surface (**Figure 3.5**). Equally valid is that many of these endogenous errors are masked when using GDEM2 for cartographic terrain visualisation. Moreover, much of the geodetic error is based on the impact of land cover, cloud cover and cloud persistence on GDEM2 elevation, which is images-scene (stereo pair) dependent and varies based on the region under observation (Gesch *et al.*, 2012). Tachikawa *et al.* (2011) report that *mean* elevation errors are higher for fewer scenes, with an increase in the number of scenes able to reduce error significantly between 1 and 10 scenes with marginal improvement beyond 15 scenes. A further proviso to the utility of the GDEM2 is that with limited readily deployable synoptic coverages available to South Africa, future studies may need to prioritise the comparison of SRTM products with other open-access DEM products such as ALOS-2 (also called Daichi-2) world 3D (JAXA, 2019) or EarthEnv90 (Robinson *et al.*, 2014) to broaden the application of sensor platform for better ontological calibration of geomorphons across multi-platform DEM surfaces.

3.3.4 Geomorphon character assessment: generalised soil-landscape properties

The relationship between soil properties, landscape feature type, and feature position is pivotal to discerning and predicting regional soil distribution and variability (Libohova *et al.*, 2016). Landscape topography is consistently acknowledged as a critical pedogenic element in the development and functioning of soils at a local or landscape level through reference to the conceptual soil-landscape paradigm personified in the notion of pedons and polypedons (Simonson, 1968; McBratney *et al.*, 2003). Different locations within the landscape can be partitioned into areas that theoretically represent unique soil modal properties or characteristics indicative of specific soil formation processes. These conceptual soil-landscape models can then be further quantified through spatial analysis of soil terrain attributes at particular points within landscape assemblages such as geomorphons (Libohova *et al.*, 2016).

Consequently, the soil-landscape model's quality directly informs the quality of the mapping results. The assessment of qualitative relationships between topography and soil variation is paramount to digital soil assessment (DSA) (Carré *et al.*, 2007). These soil-landscape relationships can then be further used to explore the continuous spatial distribution of soil properties using pedo transfer functions (Bouma, 1989), providing valuable information regarding soil attributes, quality, and function (Hartemink & Minasny, 2014). However, the interaction between soils and topography and, more importantly, soil-landscape pattern-process relationships can be treated on several levels. For this study, the approach to evaluating the relationship between geomorphon and the soil-landscape pattern-process interaction has been limited to a "black-box" assessment approach whereby the whole system is regarded as a collection of sub-units with no consideration of the dynamics of the internal structure. No attempt has been made to account for the flux of materials and energy through the pedo-transfer system or consider the balance between inputs or outputs directly relating to soil genesis in the landscape (Gerrard, 1981). The comparative assessment of soil pedo-sequence characteristics using geomorphon features is necessary given the complexity of accurately predicting soil properties at each landscape point due to the high spatial variability of soil properties (Burrough, 1993). Moreover, broad-based assessment studies, such as present research, are necessary to qualitatively inform the generalised representation of soil characterisation within topographic features, in this case, geomorphic features, and not necessarily the value of a soil property at each point in the landscape (Florinsky *et al.*, 2002).

The results presented hereafter are solely based on the legacy TBS survey data with generalised soil properties assessed for each scale-specific geomorphon unit. They are merely descriptive and meant as a form of categorical classification of soil-geomorphon association. The study does not attempt to account for the possible modes of origin for soil-landscape pattern-process or topo-sequence formation (Small, 1978). Instead, following on from work by Strand (2011), the results are intended to highlight how soil pattern-processes within geomorphons may contribute to the concept of landscape character assessment (Wascher, 2005), i.e., the distinct, recognizable and consistent pattern of elements in the landscape that differentiates one landscape from another (Swanwick, 2002). This approach attempts to emphasise the recognition of individual components that constitute the geomorphic landscape for the Bergville region. The soil properties for each geomorphon unit may then be used as a basis to establish a systematic approach to formal and parametrised soil-landscape mapping (Mücher *et al.*, 2010), providing a rapid comparison of regions in terms of their landscape character (Galatowitsch *et al.*, 2009). Finally, the meaning of "pattern" and "process" vary substantially between (sub) disciplines (Schröder & Seppelt, 2006). Therefore, I adopt the definition outlined by Grimm *et al.* (2005), which defines *patterns* as observations with a spatial and temporal structure distinct from a random process realisation.

In contrast, *processes* represent different objects' interactions in an environment (Schröder & Seppelt, 2006). These processes can, therefore, also be described through proxy associations of soil topo-sequence. The prevalence of hydromorphic soils indicates the degree of soil saturation or the description of terrain surface classification as a measure of concavity/ convexity, which could act as a proxy for zones of dispersion and accumulation of surface material. Depending on the system boundaries and the soil property of interest, processes may either be endogenic or exogenic. Synergistically, these processes depend on patterns in the landscape, but they also generate their own unique patterns in the landscape, forming complex adaptive landscape systems (Levin, 1998). The generalised classification of soil properties into the geomorphon classes across the different DEM surface resolutions yielded exciting results for the selected soil covariates. These indices were selected based on their comprehensive representation of the TBS survey and their ability to represent the region's essential soil-landscape characteristics. Concomitantly, soil description guidelines for soil classification schemes require the characterisation of landform and topography of soil profile sites (Gruber *et al.*, 2017). Therefore, these secondary soil covariates may provide a reference system for the spatial distribution of broad soil properties for designing a landscape geomorphon signature for the region, which may provide the basis for further soil-landscape disaggregation using these observed soil-landscape characteristics.

3.3.4.1 Geomorphons and soil complex distribution

An approach that is pragmatic and non-parametric (descriptive) was adopted to characterize the soil-landscape geomorphic character of the landscape. Each geomorphon surface was sampled using the terrain attributes intended to capture the region's maximum variation in generalised soil-landscape properties. Histogram modifications were performed on the attribute maps to highlight primary features in each resulting geomorphon surface (Jasiewicz & Stepinski, 2012). The observed soil complex properties characterised in the geomorphon classes are presented in **Figure 3.11 (a to d)**. The results show the relationship between each DEM specific geomorphon feature and soil complex presented as percentages (cumulative) based on the frequency of extracted gridded soils points per landscape feature within the geomorphon surface(s). **Table 3.11** provides an overview of the TBS complex descriptions concerning the major World Reference Base (WRB) Classification system (Tóth *et al.*, 2008). In soil maps, even at large scales, single category units are rarely found. In other words, a complex map might consist of a large number of different taxonomically distinct soil types in the same area (Zhu *et al.*, 1996). A map of the TBS complex soils for the study site is presented in **Figure 3.12.**

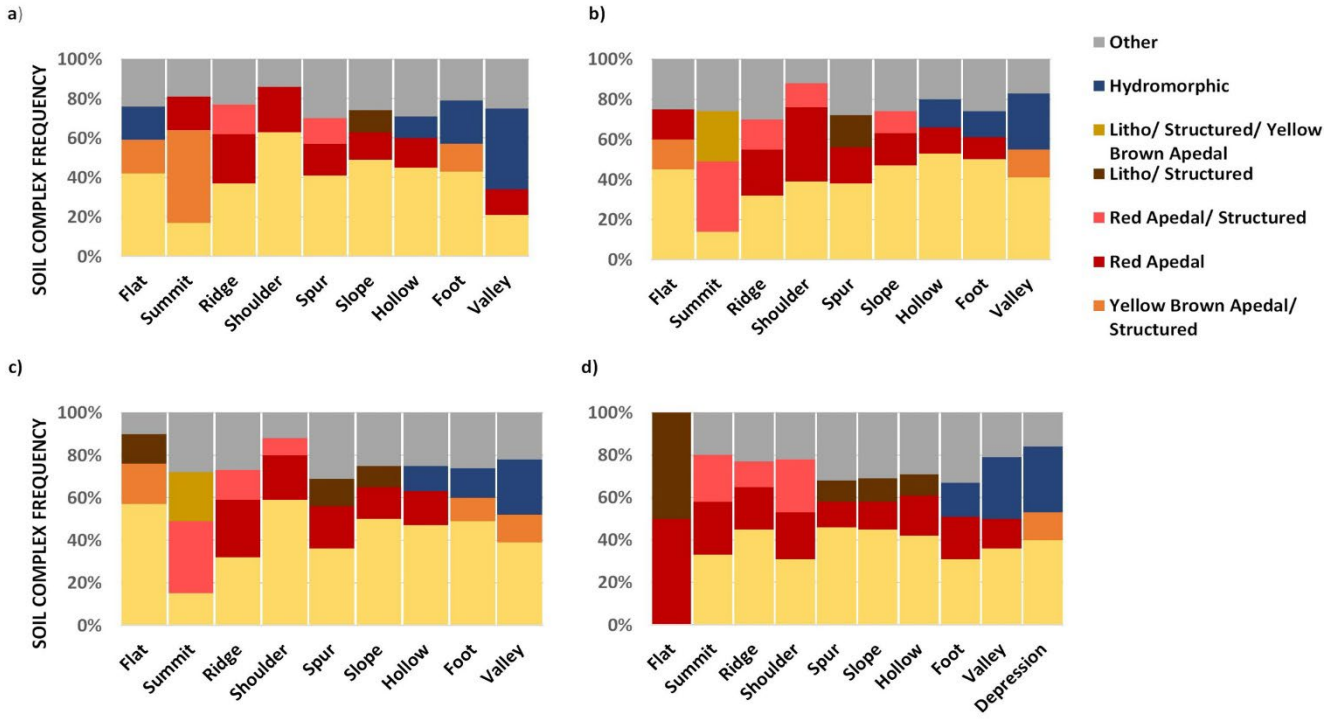


Figure 3.11: Histogram showing percent soil complex cover for a) 30 m SUDEM b) 90 m SUDEM c) 90 SRTM d) and 30 m ASTER GDEM2.

A high degree of variability was observed between the four DEM surfaces and the legacy soil complex distribution in the percentage composition of soil complex and the extent of occurrence for each geomorphon feature. The only two DEM datasets that showed overall generalised soil-geomorphon similarity were the two 90 m DEM surfaces. However, this result was expected given the prior findings of the analysis, which showed that these two datasets were highly correlated on pixel resolution alone. Interestingly, the *hollow*, *ridge* and *shoulder* geomorphons for all modelled surfaces, including the GDEM2, were highly similar both in composition and extent of the specific soil complexes, *vis a vie*, hydromorphic and well-drained apedal soils, respectively. For *hollow* features, all DEM surfaces represented the dominant soil complexes as 45 to 55% yellow-brown apedal soils, 10 to 12% red apedal soils and 10-15% hydromorphic soils. In comparison, the *ridge* features were consistently associated with 30 to 40% yellow-brown apedal, 20 to 25% red apedal and 10 to 15% red structured soils. Notable findings from the analysis include the results for soil complex distribution per geomorphons for the 30 and 90 m SUDEM surfaces. The similarity in soil complex distribution patterns between the 30 and 90 m resolution DEMs suggests that the soil-landscape characterisation and the scale of actual and predicted topographic features present in the study site were well represented. Equally in the scale of topographic features present in the study site and the *Line of sight* optimisation parameters used to model the geomorphons between the 30 or 90 m SUDEM. However, the 30 m SUDEM was able to better characterise the *valley bottom* features as being dominated by hydromorphic soils, as expected by the topo-sequence for the region, with yellow-brown apedal soils similarly typical of lower, permanently and seasonally wet areas in the landscape. In contrast, while hydromorphic soils were at least represented by *valley bottom* regions in the 90 m SUDEM and SRTM geomorphon surfaces at 26 and 28%, respectively, they were not identified as the dominant soil complex. Given the study site's definitive and known hydrography, these valley bottom features should certainly be dominated by hydromorphic soils (Van der Eyk *et al.*, 1969; Elliot & Escott, 2015).

A similar result was observed in several other geomorphon features. The SUDEM 30 m surface presented a more "analogous" soil-landscape pattern for the region, characterising the summit features as being dominated by yellow-brown and structured soil complexes. At the same time, the flat areas of the study site were indicated to contain approximately 17% hydromorphic soils. Which, given the natural physiographic composition of the study site and the degree of seasonal saturation, suggests that these soil-landscape associations are most probable. This confirms that the geomorphons features derived from the SUDEM 30 m positively represent the expected topographic catenary characterisation for the study site.

Table 3.11: Soil complex descriptions used in this study correlated with WRB soil groups.

Soil complex	Soil set	WRB Reference
Hydromorphic	Groundwater affected soils	Gleysols;
	Alternating wet-dry conditions, rich in swelling clays	Vertisols
	Accumulation of Fe under hydromorphic conditions	Plinthosols
Litho/ structured/ yellow-brown apedal	Soils with clay enriched-subsoil	Alisols; Acrisols; Lixisols
	Relatively young soils or soils with little or no profile development	Arenosols; Cambisols
Litho/ structured	Soils with limited rooting	Leptosols
	Soils with clay enriched-subsoil	Acrisols
	Relatively young soils or soils with little or no profile development	Cambisols
	Abrupt textural discontinuity	Planosols
	Structural textural discontinuity	Stagnosols
Red apedal/ structured	Soils with clay enriched-subsoil	Luvisols; Lixisols
	Relatively young soils or soils with little or no profile development	Ferralsols, Arenosols; Cambisols
Red apedal	Relatively young soils or soils with little or no profile development	Ferralsols, Arenosols; Cambisols
Yellow Brown apedal/ structured	Relatively young soils or soils with little or no profile development	Ferralsols, Arenosols; Cambisols
	Soils with clay enriched-subsoil	Alisols; Acrisols; Lixisols

A chi-square contingency test was used to further analyse the degree of soil-geomorphon similarity and explore significant relationships between geomorphon features and the soil complex covariates under the different DEM surface models. The chi-square test is an aspatial goodness of fit test used to assess whether or not the predicted distribution differs from the actual observed distribution H_o*: was defined as no association between Geomorphon features and soil complex and* H_1*: defined as the alternate hypothesis as significantly different at the 5% significance level.* The given model is represented by **Equation 3.1,** where $\underline{o}$ represents the observed association of each combination between geomorphon feature and soil complex. At the same time, $\underline{e}$ denotes the expected association of each combination. The higher the calculated value (X^2) above the critical value (X^2 critical) as determined by a chi-square table, the closer the relationship between soil complexes characterised by geomorphons at different DEM resolutions (Silva *et al.*, 2016b).

$$X^2 = \frac{\sum(o-e)^2}{e} \qquad (3.1)$$

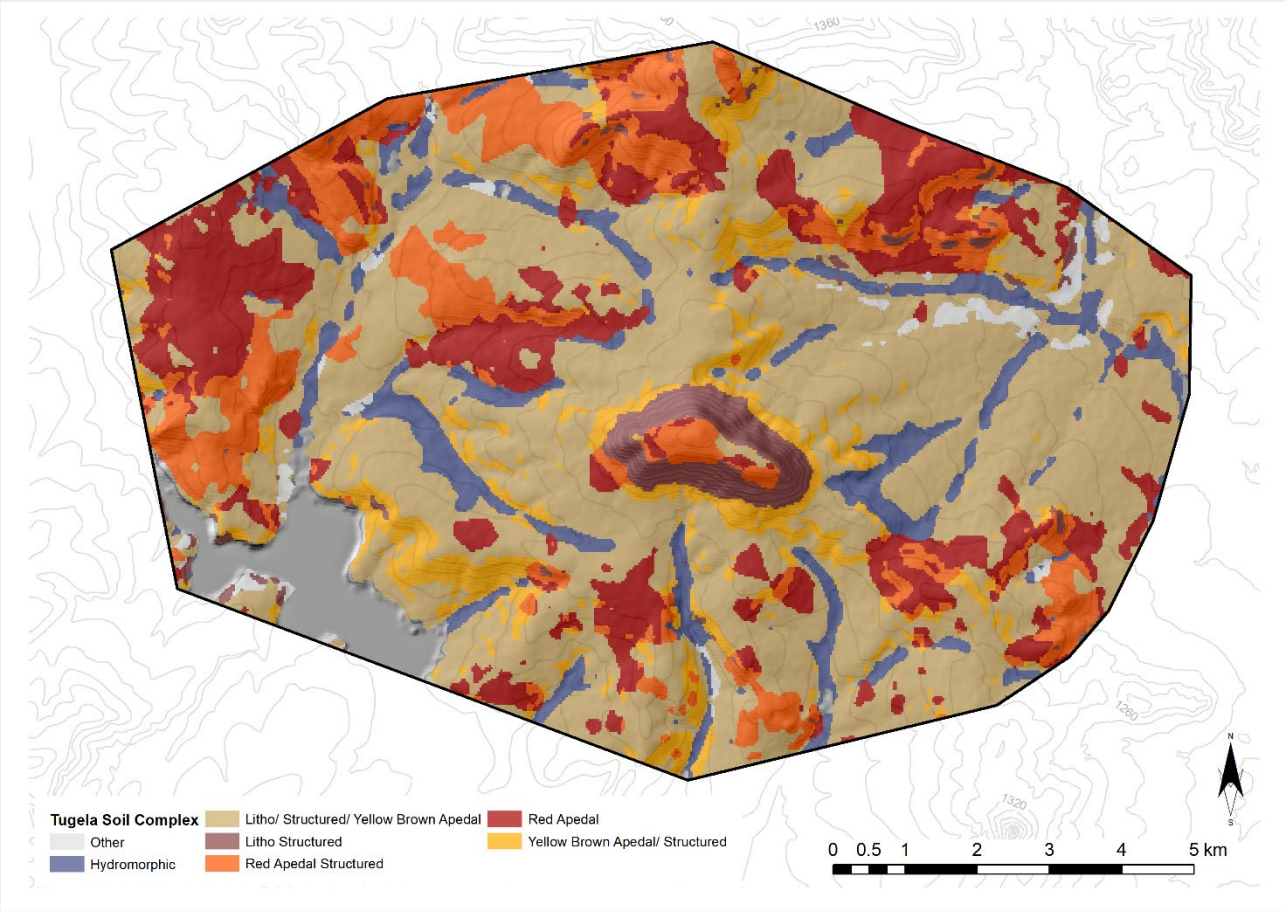


Figure 3.12: Bergville study site with 1:50 000 mapped soil complex units. After Van den Bergh *et al.* (2009)

The chi-square test for association is reported at a 95% confidence level, where $\alpha = 0.05$ and df = n-1. Where the predicted surface is not statistically different from the observed surface (30 m SUDEM), then the DEM datasets are assumed to represent the geomorphon/ soil covariate pattern process. The chi-square test compares the geomorphon and soil complex patterns between the SUDEM 30 m and SRTM 90 m, then the SUDEM 30 m and GDEM2 and finally the SUDEM 90 m and the SRTM 90 m. The results for the 2-way cross-tabulation are presented in (**Table 3.12**) showing the chi-square values (X^2) as well as the two-sided asymptotic significance (*p-value)* for soil complex. While an informative measure of statistical significance, it should be noted that chi-square tests do present certain computational limitations which users must note. Namely, the tests are not designed to represent spatial significance or spatial interaction effects; moreover, the chi-squared approach assumes a normal distribution of the residuals, which is not always the case in modelled spatial analysis (Elhorst *et al.*, 2010).

The X^2 results for soil complex and geomorphon feature between the predicted and observed datasets show better than expected association considering the marginal performance results from the BK geomorphon similarity analysis presented in **Table 3.4 to 3.5**. Most distributions for the soil geomorphon and soil complex relationships failed to model the observed surfaces, with the results found to be statistically significant in favour of H_1. However, selected geomorphon features showed positive soil complex associations between the different DEM datasets. Features such as *summits*, *ridges*, *spurs*, *foot slopes*, and *valley bottoms* showed some degree of similarity between the predicted results and observed surfaces.

Specifically, the soil complex/ geomorphon association was most notable for the *ridge* and *hollow* features showing similarity across all DEM resolutions. These results hold promise for soil-landscape studies exploring the interoperability and utility of the SUDEM and open-source DEM datasets for multi-resolution soil-landscape applications. In contrast, many of the remaining geomorphons failed to model the observed surfaces based on the *p*-values for the predicted and observed surfaces greater than the significant level, a =.05. These included *ridges* (all DEMs), *spurs* (90 m DEMs), *summits* (90 m DEMs), *hollow* (all DEMs), *foot slope* (30 m DEMs) and *valley bottom* (90 m DEMs). Incidentally. In these instances, I failed to accept H_o. It is important to note that while these results are significant, the degrees of freedom for these associations are relatively high, suggesting that the relationship between geomorphon features and soil complex while positive is also highly variable.

Interestingly, the chi-square results, including the 90 m datasets, show a definite pattern of similarity in soil complex and high altitude, narrow and higher slope gradient geomorphon features. Indeed, while higher resolution DEM surfaces are perhaps more suited to detecting subtle terrain discontinuities, the association presented here also relates to the unique arrangement of soil complex to a specific geomorphon. The more unique/ specific the soil complex within the landscape and the smaller the area of the geomorphon then the higher the association of the geomorphic soil complex patterns. This is supported by the fact that *flat* features, with higher variability of well-drained apedal soil complex combinations, showed no association between the DEM surfaces. In contrast, *hollows* characteristically dominated by hydromorphic soils showed a better association between the DEM surfaces. Equally surprising is the predominance of geomorphon dissimilarity between the 90 m SUDEM and 90 m SRTM DEM datasets. These two datasets have the same spatial resolution and genesis and would be expected to have a better agreement of soil patterns. In conjunction with the positive X^2 results between the 30 and 90 m, SUDEM selected geomorphons; these findings suggest that the quality of the DEM, and not only the resolution, is a major contributing factor to geomorphon applications in soil-landscape analyses. These results, along with similar studies Van den Bergh *et al.* (2009) and Dobos *et al.* (2000), further highlight the importance of defining the results of such analysis within the limitations of the geomorphon modelled surface. This is necessary since users can then decide how best to represent the spatial landscape heterogeneity of specific landscape indicators, especially when the rate of change or resolution of the terrain (landscape) discontinuities is unknown.

Given these findings, there may be merit in further exploring the influence that other operational spatial resolution combinations, i.e., 20, 100, 250 m and quantitatively-based statistical DSM approaches (McBratney *et al.*, 2011), could have on geomorphon based soil-landscape associations (Zerizghy *et al.*, 2013). This should include detailed soil taxonomic detail such as soil form (SCWG, 1991) and even modal soil type data (ARC, 2003) instead of broad, generalised soil complex representations. According to the results of analyzing the geomorphon-soil complex pattern, regardless of DEM scale and parameterisation, except for the GDEM2, all DEM surfaces represented the synthesised interpretation of composite terrain characteristics for the study site most accurately as well as possible.

The question of geomorphon feature relevance for defining landscape structure and terrain spatial heterogeneity should be framed in the context of spatial arrangement (geomorphon configuration, i.e., *pattern*) and/or constituent diversity (composition, i.e., *process*) (Lausch, 2015) and ultimately based on user preference. For example, while one geomorphon surface may accurately depict the most dominant soil feature on a topographic series (i.e., 90 m geomorphons), another geomorphon surface with a different resolution could illustrate the expected sequence of soils along with the series (i.e., 30 m geomorphons). These results show that regardless of DEM scale and even geomorphon model parameterisation, geomorphons can repeatedly depict the probable and expected soil associations, thereby topo-sequence, within the landscape that best represents the synthesised interpretation of composite terrain characteristics for the study site.

Table 3.12: Chi-squared (X^2) goodness of fit analysis of Geomorphon association with soil complex distribution by comparing the 30 m SUDEM and 90 m SRTM, the 30 m SUDEM and 30 m GDEM2 and, the 90m SUDEM and 90 m SRTM surface models.

Geomorphon	**Flat**			**Summit**			**Ridge**		
	X²	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**
30m SUDEM vs 90m SRTM	38.934	4	0.001	122.848	5	0.001	0.794[a]	3	0.851[a]
30m SUDEM vs 30m GDEM2	200	4	0.001	75.669	4	0.001	1.669[a]	3	0.644[a]
90m SUDEM vs 90m SRTM	37.311	4	0.001	0,206[a]	3	0.977[a]	0.512[a]	3	0.916[a]
	Shoulder			**Spur**			**Slope**		
	X²	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**
30m SUDEM vs 90m SRTM	8.257	3	0.041	26.134	3	0.001	24.377	4	0.001
30m SUDEM vs 30m GDEM2	38.226	3	0.001	22.253	3	0.001	24.513	4	0.001
90m SUDEM vs 90m SRTM	9.295	3	0.260	0.622[a]	3	0.891[a]	21.145	4	0.001
	Hollow			**Foot slope**			**Valley Bottom**		
	X²	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**	***X²***	**df**	**Asymp. Sig.**
30m SUDEM vs 90m SRTM	0.416[a]	3	0.937[a]	3.061	3	0.382	34.95	**4**	0.001
30m SUDEM vs 30m GDEM2	21.574	4	0.001	39.56[a]	4	0.940[a]	239.94	**3**	0.001
90m SUDEM vs 90m SRTM	1.38[a]	3	0.710[a]	22.047	4	0.001	0.802[a]	3	0.849[a]

[a] ***denotes geomorphon features where terrain features are not similar across observed DEM surfaces, i.e., failed to reject H_o***

3.3.4.2 Geomorphons and soil texture (clay fraction)

The spatial distribution and variability of soil texture (coarse sand %, fine sand %, silt % and clay %) is a particularly significant soil physical property since it influences structural and hydrological soil functions. These include drainage, water holding capacity, aeration, susceptibility to erosion (Veronesi, 2012), nutrient retention, organic matter content, cation exchange capacity, pH buffering, and soil tilth (Piikki *et al.*, 2013). Soil texture fraction and related particle size distribution greatly influence the rate of water flow through soils and so has particular value in DSM (McKenzie & Austin, 1993; Arrouays *et al.*, 1995; McBratney *et al.*, 2003) and hydopedology (Lin, 2003). Furthermore, its increasingly recognised as a critical soil covariate in a wide range of ecological, hydrologic, climatic, and other environmental models related to DGM (Greve *et al.*, 2012). For this reason, it serves as a universal indicator for a complete soil-landscape model (De Bruin & Stein, 1998). Much of the soil physical and chemical property data from the TBS is legacy data collected to qualitatively characterise, classify and simplify the soil resources in the survey area. Defining soil factors, such as soil texture, into aggregated classes rather than numerically representing specific soil texture fractions naturally presents limitations to thoroughly exploring the predictive influence of terrain on soil texture (Greve *et al.*, 2012). Evaluating soil texture with the help of terrain units, such as geomorphons, may be a cost-effective strategy for further developing a generalised understanding of localised pedotransfer functions (PTF) (Bouma, 1989; Jahn *et al.*, 2006), given that the basic premise of PTF is translating data we have into what we need (Kidd *et al.*, 2020). However, given the limitations of the TBS data, the results of the presented soil texture analysis are not intended to serve as inputs into simulation or decision-support models for explicit and detailed PTF for the prediction of other soil properties from soil texture.

Additionally, as Minasny & Hartemink (2011) point out, developing new PTFs can be a toilsome task since it requires an extensive soil database containing many soil measurements unavailable for the TBS dataset. Secondly, it is beyond the scope of the present research objective(s). Increasingly, in DSM, soil variables like, among other things, soil texture are increasingly mapped using geostatistical modelling approaches such as the regression-kriging framework leveraging ancillary data sources as well as providing a suitable framework for error propagation and prediction mapping. For a comprehensive review of PTF applications and the principles of predictions, users are referred to works of (McBratney *et al.*, 2002; Guber *et al.*, 2006; Minasny & Hartemink, 2011).

The presented interpretations of soil texture and geomorphon unit associations presented hereafter are intended to leverage the catenary approach as an analytical framework to understand the conceptual relationship of soil texture patterns over a landscape (Chiang & Peterson, 1970). A typical sequence of soil properties is associated with specific terrain units that coincide with evaluating soil properties related to geomorphon units in catenary soil development. Moreover, there is still a worldwide need for qualitative and spatial soil information for environmental monitoring and resource management (Lagacherie & McBratney, 2006; Ahrens, 2008). This approach lends itself well to a knowledge-based system which is still considered the most common instrument used to support soil mapping at the regional scale (Boehner *et al.*, 2001).

Typically, representing soil-landscape attributes such as soil texture requires two major processes: (1) first, conceptually organising soil properties or parameters into acceptable taxonomic classes and (2) defining the variation of these taxonomic classes over space (Zhu & Band, 1997). Consequently, the presented results predicate the clay fraction associated with the classified soil complex properties following the South African

Bionomical Classification System (SABCS) (MacVicar, 1977), as was understood in 1977. The SABCS provides a simplified classification of soil series description based primarily on clay content and grade of sand, yielding 501 soil series (Schulze, 2007) based on a Land Type unit (MacVicar *et al.*, 1974) which is considered a class of land over which macroclimate, terrain form and soil pattern each display regularity.

This regularity is such that there would be little advantage in defining, on a countrywide basis, smaller and more uniform landscape entities for potential agricultural determination (Schulze, 2007). Following the work set out by Hutson (1983), a silt content of approximately 10% is assumed as a general value for most of the South African soils. The SABCS clusters soil series by classes of clay content, and each soil series may be further grouped, in general terms, into a texture class, using an expected *median* value for each clay fraction class (Schulze, 2007). For reasons of convenience and practicality, a similar approach to Schulze (2007) was used to classify the vertical distribution of topsoil (≈ 0 to 150 mm) clay content values extracted from the modal profiles of the TBS survey into four textural classes (**Table 3.13**).

Consequently, these texture classes correspond well to the widely accepted discrete class breaks for infield measurement of soil texture by soil survey practitioners (Smith, 2006). The clay content and geomorphon analyses have been presented as histograms indicating frequency class and ordinal clay content category. This study adopts the same clay content class breaks outlined by Van den Bergh *et al.* (2009)(**Figures 3.13a to 3.13d**). Histograms are particularly useful at characterising the distribution of summary-statistic topographic attributes across a range of landform elements (Gessler *et al.*, 1995). Moreover, these clay texture classes are consistent with the National geo-referenced database comprising land capability, land suitability, agricultural ecological zones, land use data, and information metadata standards (EnviroGIS, 2016b).

Table 3.13: Clay fraction class used to describe clay content and associated point observations and coverage for each class

Clay fraction class	Texture class	Clay content	Number of observations	Approximate coverage
0 – 5%	loamy sands	low	10	0.15 ha
6 – 15%	sandy loams	medium	1 158	17.4 ha
15 – 35%	sandy clay loams	high	3 554	53 .3 ha
36 – 55%	sandy clay	very high	27	0.45 ha

There are several important features to note from the soil texture histogram plots. The main observation is that all DEM surfaces appear to have similar generalised clay texture distribution across the respective geomorphon features. Given the known varied topographic characterisation of the study site, I would expect to see a more varied description of clay content between geomorphon elements or at least between DEMs of different pixel resolutions. Especially since soils vary widely due to their position in the landscape, particularly at a regional scale (Wilson & Gallant, 2000; Luo *et al.*, 2014). Also, it is well-known that human activities at the field level influence soil properties. Many studies highlight the considerable variation of soil properties between fields with different land use and management strategies, even on the same soil type (Yemefack, 2005). Secondly, all geomorphon features (the entire landscape) appear to be dominated entirely by medium to high clay fractions.

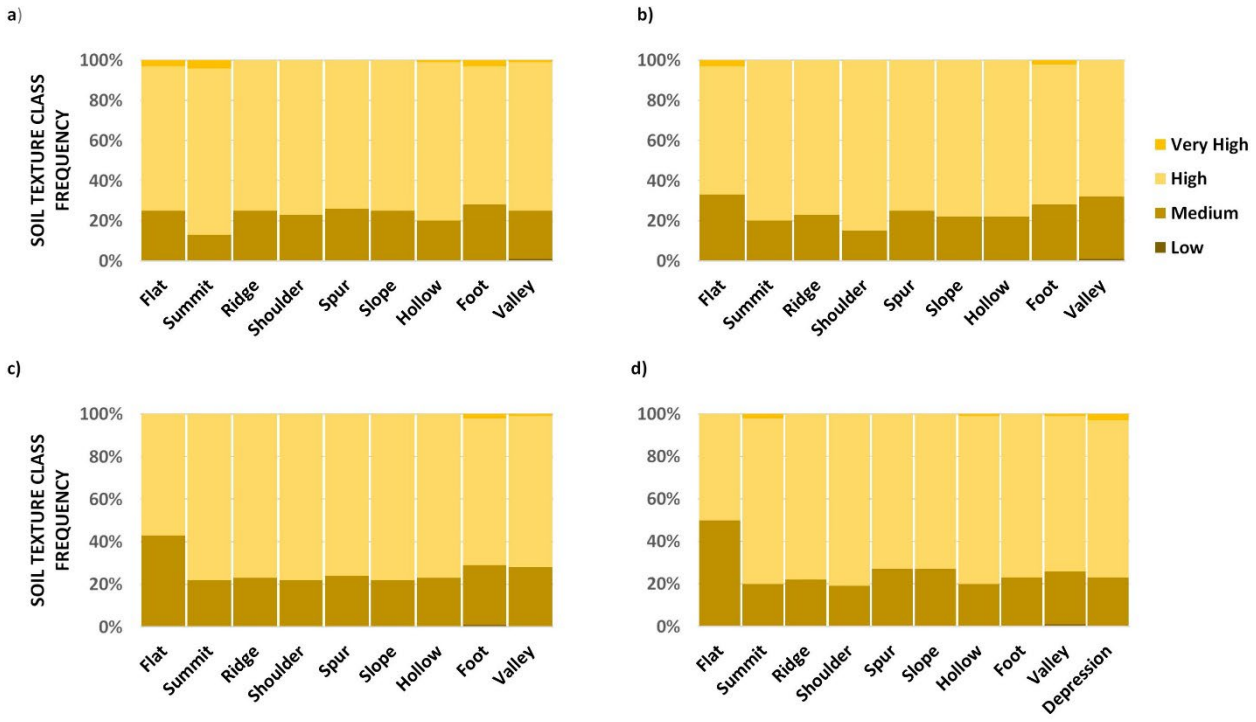


Figure 3.13: Histogram showing frequency and pattern of soil texture (clay %) for each geomorphon feature in the a) 30 m SUDEM b) 90 m SUDEM c) 90 m SRTM d) and 30 m ASTER GDEM2.

Notwithstanding, despite the uniform characterisation of clay content, the results could still detect a generalised catenary association/ similarity for selected geomorphon features linked to cognate topo-sequence underpinnings for the region. For instance, the lower-lying geomorphon features such as *flat*, *foot slope* and *valley bottoms* typically accepted to represent zones of colluvial material accumulation characterised by increased clay content were segmented by approximately 28% medium, 70% high and 3% very high clay content in the SUDEM 30 m, SUDEM 90 m and SRTM 90 m respectively. In contrast, the geomorphon features associated with higher landscape positions such as *summits*, *ridges*, *shoulders*, *spurs* and *slopes* reported higher frequencies of sandy clay loams with clay fractions between 15-35%. Specifically, the results of the histogram analysis provide a useful visual description of clay fraction association within the selected geomorphon features. In particular, the SUDEM 30 m, SUDEM 90 m, and SRTM 90 m were all able to represent *foot slopes* and *valley bottoms* as dominated by medium-high clay content.

However, similar clay fraction associations between these DEM surfaces were observed in the *flat*, *ridges* and *spur* geomorphons suggesting that these features too are characterised by medium to high clay fractions. These results are surprising considering that convex surfaces are typically well-drained, while concave surfaces, such as *valley bottoms* and *depressions,* are more likely to contain soils with hydromorphic features. So, we would indeed expect to see higher clay fractions associated with these lower-lying positions. For instance, Malo *et al.* (1974) reported that soil properties such as organic carbon content, clay content, and surface thickness increased from shoulder to lower-lying *foot slope* positions. Likewise, given characteristically that water flow determines different areas of erosion and deposition, Gessler *et al.* (2000) observed that higher values of calcium, magnesium and potassium, base saturation and clay content were recorded on the lower slope features, i.e., *flat*, *valley bottom*, due to the direct action of water surface flow by depositing material on these landscape positions.

Interestingly, however, Pierson & Mulla (1990) reported that soils on lower-lying landscape positions, such as *foot slope* and *toe slop*e classes, may have a higher organic carbon content, greater aggregate stability, and lower clay content than those on *summit* positions. At face value, these results appear to be too generalised and contradictory to the region's tacit or heuristic "rule-based" topo-sequence model. However, disregarding these results as erroneous must be treated with caution. Several key observations can be outlined, which may provide a supplementary explanation for the observed effects.

A significant shortcoming in using the clay content reported in the TBS relates to the estimated clay content methodology. Much of the soil texture (clay fraction) was estimated in-field and then "binned" into broad, predefined texture categories (Van der Eyk *et al.*, 1969; Smith, 2006). In fact, the data collected from conventional soil surveys often rely on qualitative analysis of the landscape, where it is assumed that properties of modal profile data apply to an entire mapping unit (Dent & Young, 1981; Schulze, 2007). This is not a trivial matter if we consider for a moment that the soils in this region are typically highly weathered, dystrophic and characterised by higher than the average accumulation of clay minerals within 80 cm rooting depth. Furthermore, the dominance of secondary minerals such as Goethite (α-FeOOH) and Hematite (α -Fe_2O_3), giving rise to the prominent reddish-brown hue down the soil profile, has been shown to mislead the in-field estimation of clay content, especially in highly weathered soils (Smith, 2006). While this may be a limitation of using historical and infield estimated data for detailed quantitative analysis of soil properties, especially for detailed PTF models, there is still information that can be gained from the ability to understand soil characteristics and properties when detailed point data are lacking or too expensive to obtain. This is especially necessary when the final application of the analysis does not require high temporal or spatial resolution accuracy (McBratney *et al.*, 2003). Incidentally, this is precisely the intention of the Land Type survey and remains the strategic prerogative of the National Land Capability approach (EnviroGIS, 2016a).

Secondly, the *prima facie* histogram results for soil texture need to be considered with the soil complex frequency for each geomorphon feature (**Figures 3.11a to 3.11d**). While the lower landscape features represent the hydromorphic soil complex pattern well, they do not define these landscape positions exclusively since each geomorphon can be associated with several different soil complex descriptions. For instance, the 30 m SUDEM *flat* geomorphon has the following soil complex combinations: 45% litho/ structured/ yellow-brown apedal soils, 15% yellow-brown apedal/ structured soils, 20% hydromorphic soils and 20% a variety of various minor soil association descriptions. Similarly, other geomorphons are characterised by their own set of soil complex variations for *valley bottoms* and *foot slope* features. Finally, the categorical class breaks used to represent the soil texture estimation values are coarse and known to be imprecise (Camp, 1995). This is intentionally designed to accommodate field scientists' estimation of soil texture in-field as a first approximation. Moreover, the geomorphon features may be too large, or sampling frequency unequally distributed (oversampled or undersampled in certain geomorphon features) to adequately represent detailed soil texture variation giving rise to the observed isotropic properties for each respective DEM surface.

These limitations suggest that the use of legacy categorical data (and associated analyses) for geomorphon characterisation for soil-landscape analysis, at minimum, must explicitly consider the scale of measurement for the terrain parameters under observation, i.e., discrete vs continuous (actual clay percentage values), the method of clay content analysis, i.e., in-field estimations vs actual clay percentage obtained through wet chemistry approaches as well as the sampling distribution to predict these soil properties and represent each geomorphon feature adequately. In other words, the spatial prediction or spatial interpolation of soil texture characterised by regional geomorphon units may benefit from applying geostatistical methods that rely on the statistical analysis of sampled field data for an area of interest (Hengl *et al.*, 2008). The key to applying these statistically-based systems is creating quantitative models (experimental variogram) generated from a model-

fitting process where the accuracy of model outputs is determined by the sampling distribution (strategy) of the data to be modelled. The models are intended to "explicate" the relationship (nugget, sill, lag size) between the predictor and target variables and are typically applied to environmental covariates to achieve the spatial predictions (Webster & Oliver, 1990; McBratney *et al.*, 2003; MacMillan & Shary, 2009; Lark, 2012; Oliver & Webster, 2014).

However, as Lark (2012) contends, many soil properties arise from deterministic processes, which can be modelled. However, it is rare, if at all, to deploy our understanding directly to address the practical needs of the land manager or regional planner who require pragmatic predictions of soil properties at unsampled sites. Interestingly, Paterson (2018) recently showed that the relationship between topography and soil texture, when evaluated using Kriging (Krige, 1951) to model the clay variogram, is strongly isotropic variograms did not fit well with small-scale landscape applications. A study conducted by Yemefack (2005) reported that a closer sampling density would be required to map regional variability, which is not due to land use, regional trend or environmental covariates when using geostatistical analysis. Their study highlighted that regional and local effects, and their interaction, accounted for approximately 70% of the clay fraction variation in the landscape. These findings highlight that soil properties exhibit a high spatial dependence, even at the plot level, with regional trends predominantly defined by elevation range, accounting for approximately 30-50% of the total clay fraction variation.

The non-parametric association of clay texture with geomorphon features could prove useful to a range of landscape-scale land use and natural resource planning functions for strategic agronomic purposes. Firstly, land planning practitioners can be assured that DEM source choice, and consequently, DEM resolution between 30-90 m, should not significantly influence the representation of soil texture across geomorphic features. Considering that much of the Land Type survey data for this region is derived from the TBS (Paterson et al., 2015), similar results could be expected when disaggregating the Land Type data by geomorphic features. Secondly, the correct representation of medium to very high textured soils in lower altitude features such as *foot slopes*, *valley bottom* and *flat* features typically characterised by deeper and higher textured soils assures that the modelled geomorphic landscape features align well with knowledge-based conceptual models of soil-landscape patterns for the region. Not surprising, the soil texture results for the GDEM2 provide a final *coup de gras* of why users should be cautioned when integrating elevation datasets of known inherent data deficiencies for soil-landscape and digital soil mapping analysis, especially under South African conditions. Despite the glaring radiometric inadequacies highlighted earlier in this study, the GDEM2 texture results are still positively represented compared to the other geomorphon DEM surfaces. Invariably, the ASTER DEM's endogenous error is not manifested in the soil texture results given the broad class categories used for soil texture. Users should carefully consider how the soil parameter investigated is influenced by and influences DEM surface choice.

.

3.3.4.3 Geomorphons and soil depth

Many soil and environmental variables vary horizontally and with depth and altitude (Hengl, 2009). Soil depth is strongly linked to landscape characteristics. It is an essential component for soil mapping and land use planning and management since it influences and benefits a wide range of environmental and ecological functions (Menezes *et al.*, 2014; Flynn *et al.*, 2019a). Soil depth has been shown to affect hydrological processes (Liang et al., 2004; Schenk & Jackson, 2005; Tesfa et al., 2009), impact soil quality and productivity (Power & Sandoval, 1981; Bunning & McDonaugh, 2011), influence the soil's capacity to store and hold moisture ((Boer & Del Barrio, 1996) as well as an integral component when determining soil carbon and other elemental stocks (Wiese & Ros, 2016a; Lu & Feng, 2019). Furthermore, any method that enables the accurate classification and prediction of soil depth can improve soil erosion and land degradation management (Hudson, 1990; Scarpone *et al.*, 2016). Yet, soil depth is still one of the least understood and challenging to measure soil variables of any landscape system since spatial patterns in soil depth result from the complex interactions of many factors, i.e., terrain variability, lithology, climate, biological, chemical and physical processes (Jenny, 1941; McBratney *et al.*, 2003).

Dietrich *et al.* (1995) point out that soil depth is highly variable spatially, laborious, and costly to measure infield, even for a manageable study region. Understanding its spatial relationship with landscape features that can be observed using remotely sensed approaches could be of great importance to many aspects of land resources management. Therefore, there is a need to define the relationship between soil depth with terrain morphological units using geomorphons since soil depth has been shown to vary as a function of many terrain factors. These include slope gradient, land use, slope curvature and upslope contributing area (Minasny & McBratney, 1999; Kuriakose *et al.*, 2009), with topography identified as a pedogenic factor that significantly influences the spatial distribution of soil depth (Florinsky *et al.*, 2002). It is acknowledged that simply describing the association between soil horizon depth for the study site with specific geomorphon patterns in divergent and convergent landscape positions does not provide a comprehensive alternative to conventional soil survey depth assessment methods. However, this study's approach provides a valuable method of assessing the non-stationarity of legacy soil depth data with functional terrain morphological units. Consequently, still a central operational strategy for much of the synoptic land resource planning in South Africa.

Furthermore, observing the soil depth patterns for each geomorphon feature will provide a knowledge-driven model validation approach for the predicted DEM geomorphons since much of the region's topo-catenary and soil relationships are well defined. There is a strong correlation with the spatial variability of soil depth over a landscape with deeper soils, higher clay content, and more extended periods of wet conditions associated with the base of slopes and flatter level regions. In comparison, shallower soils and drier conditions tend to be positioned on steeper slopes (Hiederer, 2013). It is understood that the use of digital soil morphometric approaches and the application of advanced techniques such as mass-preserving splines (Ponce-Hernandez *et al.*, 1986; Jones & McBratney, 2016) and depth-splines (Malone *et al.*, 2009) to derive continuous depth function distributions (Hartemink & Minasny, 2014) offer improved harmonisation with contemporary global methods. These mechanistic approaches provide necessary empirical results of lithostratigraphic discontinuities, signifying changes in a landscape that influence soil properties and their distributions in space or time or both (Zinck, 2013). These spline functions respect soil properties' average values and assume continuous variations of soil properties with depth and entail fitting splines to available soil profile data and then integrating the splines to generate harmonised depths of each observation (Malone *et al.*, 2009; Vaysse & Lagacherie, 2015). However, as Malone *et al.* (2009) point out, these novel procedures are not without drawbacks.

While continuous soil depth functions such as equal-area splines are advantageous for predicting soil properties at specific depths, from a spatial context, a collection of spline functions for individual site observations will only provide point observation data sets. Such data, especially in the context of the present study, will be of little use when continuous "fuzzy" estimates of soil property variation across defined study areas or landscapes are necessary. For a full review of the application of equal-area splines, readers are referred to the foundational usage and modelled mathematical expression of soil depth functions detailed in Ponce-Hernandez *et al.* (1986); Bishop *et al.* (1999) and Malone *et al.* (2009).

This study has not attempted to model continuous soil depth function variability related to terrain position due to field point sampling deficiencies, precision, and accuracy limitations imposed by the legacy TBS. Instead, the results presented provide an exploratory analysis of the average solum depth for each DEM-derived geomorphon surface. While this approach may lead to discontinuous or stepped profile representations of specific soil attributes for future studies, the present method is considered the most rapid and pragmatic first approximation since the aim is to describe the relationship between geomorphon features and soil depth for 2D mapping applications. This explicitly limits the application to be applied to soil depth and other soil properties by horizon depth (3D soilscape applications), i.e., soil texture, soil organic carbon. Furthermore, the interpretation of descriptive soil observations, which can subsequently be interpreted and linked to soil types, soil classes and terrain features which are then assumed to be characterised by a limited and definable range of soil properties, is a cost-efficient alternative to soil analytical approaches (Bouma *et al.*, 1998; Kempen *et al.*, 2011).

Using the legacy soil modal data, the depth representation has been presented as three categorical classes given the TBS initial modal soil data representation and based on the depth class specifications defined by Van den Bergh *et al.* (2009). The theoretical solum depth is down to the soil's parent material. For practical reasons, field depth observations were limited to the effective rooting depth/ depth to impeding layer to a maximum diagnostic depth of 1.5 m. Like soil texture, geomorphon features are characterised by the frequency of gridded sample points per depth class. Soil depth values initially described by the original five soil depth interval classes, i.e., 300, 600, 900, 1200, > 1200 mm, were then aggregated to the following three soil depth descriptions: 0 to 450 mm: *shallow*; 451 to 850 mm: *medium*; 851 to 1200 mm: *deep* (**Figure 3.14**). While these described soil depth intervals are not readily compatible with the specifications outlined by international systems for soil data harmonisation, such as the GlobalSoilMap.net project (Arrouays *et al.*, 2014), they are in accord with soil depth specifications at a national level (EnviroGIS, 2016a) and comparable with selected international studies using the same soil depth intervals (Brubaker *et al.*, 1993).

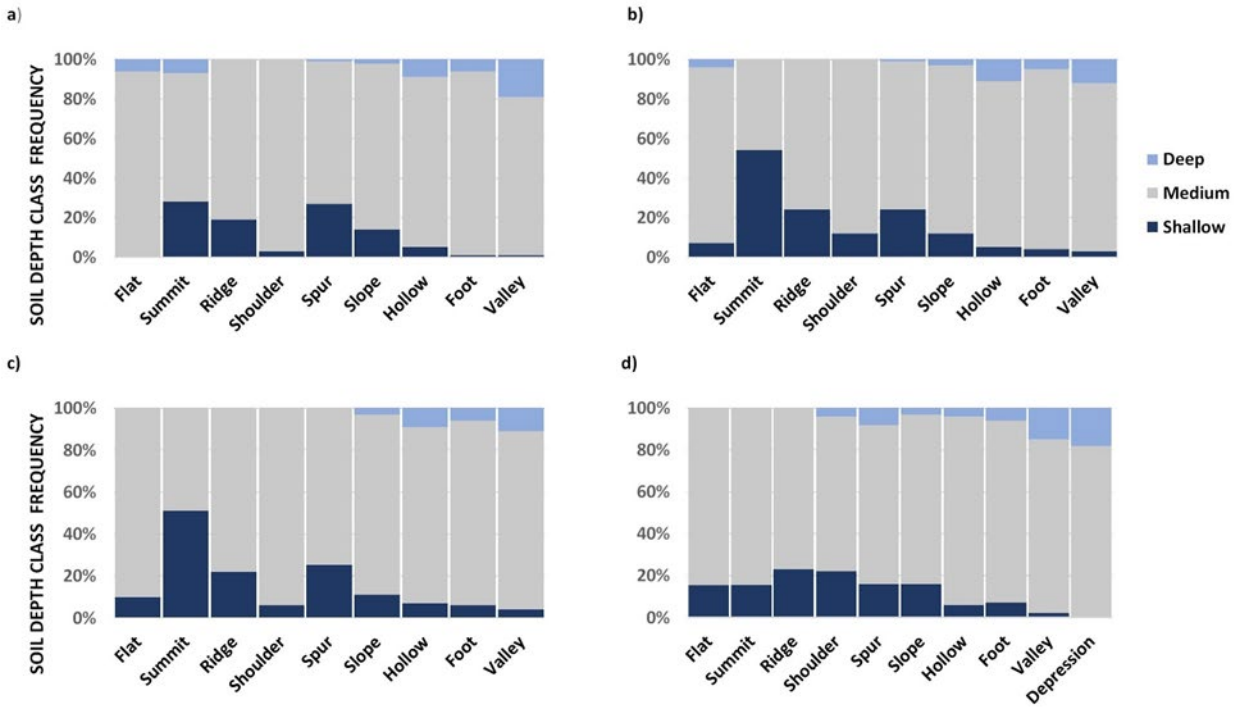


Figure 3.14: Histogram showing frequency and pattern of soil depth for each geomorphon feature in the a) 30 m SUDEM b) 90 m SUDEM c) 90 m SRTM d) and 30 m ASTER GDEM2.

The initial six-class diagnostic soil depth intervals were aggerated to three broad qualitative depth description classes. The geomorphon soil depth histograms present no interpretational challenges to characterising regional landscape-scale soil depth patterns. Notwithstanding, an explicit limitation of the depth and geomorphon association presented in this format is the depth intervals' observed discontinuity representation. Tesfa *et al.* (2009) point out that averaging the depth components limits quantification of soil attributes' variability within each class. These results manifest this since soil depth values use the *medium* soil depth range as the functional class boundaries. This is most likely to generalise the soil depth spatial pattern, absorbing small scale variability into larger class units (Moore *et al.*, 1991; Zhu, 1997). Consequently, the results indicate that much of the geomorphon features for all DEM surfaces are dominated by medium soil depths varying from 451-850 mm. Interestingly, a recent detailed field survey of the region evaluating 150 soil pits confirms that the average effective rooting depth for the study area is 700-750 mm (Barichievy & Botha, 2017).

While the approach may be rudimentary, the outcomes should not be too hastily critiqued. The results for *summits, spurs,* and *ridge* geomorphon features – typically erosional terrain features characterised by shallow depths – were detected in the SUDEM 30 m, SUDEM 90 m, SRTM 90 m and intermittently in the 30 m GDEM2 DEM as dominated by *shallow* and *medium* soil depths, i.e., 0 to 450 and 451 to 850 mm respectively. Moreover, both the 90 m SUDEM and SRTM recorded between 50 to 60% of the geomorphon *summit* features dominated by *shallow* depths with the absence of any *deep* soils detected. Likewise, lower-lying geomorphon features such as *hollows, foot slopes*, *valley bottoms* and *depressions* (GDEM2) contained between 10 to 20% *deep* and 80 to 70% *medium* depth soils. In effect, the cumulative frequency for all lower-lying geomorphon features is estimated to be between 0.75 and 1% for the entire study region for all DEM surfaces.

Interestingly, the 90 m DEM surfaces represented the steeper, shallower terrain features better, possibly even better than the SUDEM 30 m surface. The latter showed more refined representations of the flatter, narrower and in-depth features such as the *valley bottoms* and *hollows*. Additionally, histogram patterns for both 90 m DEM surfaces were almost identical in their representation of soil depth across similar geomorphon features within the study site, encouraging the interoperability of SRTM and generalised SUDEM datasets for similar terrain applications. Of particular interest are the geomorphon features, which appear to show high similarity in soil-depth pattern across all DEM surfaces, albeit with variable frequencies. These include *summits, ridges, shoulder, spur, slope,* and *hollow* geomorphonsThis result is encouraging since it would integrate many open-source synoptic spatio-temporal strategic and geospatial planning resource databases. In particular, the South African Atlas of agro-hydrological and climatology (Schulze, 2007) provides comprehensive regional coverage of climatic parameters essential in agro-hydrology and agro-climatology for the application in the catchment management, land use planning, and agriculture sectors to be useful in regional and local decision making. Of particular interest, the summits modelled with the SUDEM 30 m were able to represent 7% of the area as deep soils, suggesting that the use of fine-resolution DEMs with coarse resolution soil covariate datasets such as the Tugela soil depth are still able to represent better selected soil-landscape relationships of narrow terrain features than their coarse 90 m counterparts.

When overlaid with soil texture, these soil depth results could benefit operational land use planning from an agronomic perspective. Superimposing the soil-covariate datasets with the geomorphon features may better define ontological geomorphon traits for the area and representation of soil-landscape composition and information. Since the methodology allows for rapid assessment of soil-covariate associations with geomorphons features, I see no limitation to extending future investigations to the national level. This would encourage the testing of standards for creating and implementing these models for consistent sampling and validation refinements for unsurveyed regions, especially when using legacy soil data. Moreover, regional adaptations must be incorporated into developed models to represent better soil variability and geometric signatures in the specific area of interest (Silva *et al.*, 2016b). This could add value to a range of ecological and conservation objectives to meet the growing demand for sustainable resource management.

These soil depth results could benefit a broad spectrum of coordinated natural resource management outputs within the region when overlaid with soil texture. Superimposing the soil-covariate datasets with the geomorphon features may better represent soil-landscape composition and information, providing better defining ontological geomorphon traits for the area. Moreover, regional adaptations must be incorporated in developed models to represent soil variability and geometric signatures in an area of interest for a range of ecological and conservation objectives to meet the growing demand for sustainable resource management (Silva *et al.*, 2016b). The results from this study have shown that, as a first approach, geomorphons can successfully stratify the landscape into units that may have more homogeneity in terms of soil-landscape properties.

3.3.4.4 Geomorphons and terrain surface classification

There remains a need for tangible landscape objects as pressure increases on land managers to adopt comprehensive landscape planning, nature conservation and resource management tasks (Drăguţ & Blaschke, 2006). According to Pike (2000), many digital geomorphology approaches aim to delineate homogenous landscape objects or landform elements using standard procedures or a specific set of algorithms (Gardner *et al.*, 1990) to define slope gradient and slope curvature characteristics.

The development of spatial techniques for evaluating three-dimensional conceptual models for open and closed systems on hillslope landforms began early, with Ruhe & Walker (1968) being amongst the first to describe a nine-unit terrain geometry model. Ruhe (1975) further described the convex-concave associations in slope curvature with a matrix of nine primary forms and varying three components: (1) slope gradient, (2) slope length, and (3) slope width. Building on these early concepts, Huggett (1975) tested surface flow on three-dimensional slope shapes for a conceptual area, defined as a soil-landscape system. His major contribution is relating this soil-landscape system to the drainage divide of a watershed. This allows complete integration of slope curvature models to hydrologic, geomorphic, and pedologic models (Grunwald, 2016).

By describing terrain surface classifications (*TSC)*, i.e., the distribution of convex and concave areas, we can evaluate gravity-driven material propensity to converge or diverge as it moves through the landscape (Greve *et al.*, 2010). Concave slopes are expected to concentrate more water and sediment material and are generally considered soil accumulation zones. Convex slopes show an inverse effect, dispersing flow and limiting material accumulation, leading to a lesser quantity of soil accumulating than concave slopes (Greve *et al.*, 2012). Therefore, the slope gradient directly influences this material's movement rate, with flow acceleration directly affected by the slope gradient's steepness. The ability of geomorphons to accurately represent these patterns and processes is key to application in DSM and soil-landscape analysis. In this study, *TSC's* leading utility to geomorphon mapping is to characterise the morphons into convex (erosional) or concave (depositional) features in the context of slope gradient (Van Westen, 2016) and determine if these associations are aligned with knowledge-based landscape interpretations typically adopted by land use practitioners.

Applications of *TSC* are diverse. The introduction of point-cloud methods to generate grid DEMs for landscape discretisation has also rapidly developed using Lidar or UAVs and is gaining operational popularity (Hu *et al.*). However, many contemporary global terrain classification data sets still rely on the work by Ruhe & Walker (1968), Dikau *et al.* (1991), Meybeck *et al.* (2001), Jenness (2007), Drăguţ & Eisank (2012) and Sayre *et al.* (2018). These datasets are, however, predominantly primed for coarse-resolution (1 km) landscape applications. They do not adequately represent certain terrain features important for soil-landscape mapping (Iwahashi *et al.*, 2018), particularly for this study region. Therefore, in this study, The results for *TSC* follow the classification used to describe local slope convexity proposed by Iwahashi & Pike (2007). This relies on an unsupervised landform classification method, optimised in the SAGA platform based on only three terrain attributes: slope gradient, surface texture and local convexity. The unsupervised approach's advantage is that it treats topography as a continuous random surface, independent of any spatial or morphological orderliness imposed by fluvial activity and other geomorphic processes (Doumit, 2018). This original method can segment the landscape into 8, 12 or 16 *TSC* classes with a physical meaning of landscape properties. For this localised analysis of *TST,* the original method was modified to four categories: *steep slope, high convexity (SSHC), steep slope, low convexity (SSLC)* and *gentle slope, high convexity (GSHC), gentle slope, low convexity (GSLC).* Similar to Drăguţ & Blaschke (2006), the choice of reducing the overall categories to only four is to minimise the combination of landform profile elements that are likely to be similar in a natural landscape. Therefore, all other classes with different degrees of membership are assigned to one or more main classes.

A visual analysis of the frequency histogram of *TSC* (**Figure 3.15**) highlights a well-defined and comparable general similarity between the SUDEM 30 m, SUDEM 90 m and SRTM 90 m DEM surfaces. The geomorphon classifications of lower-lying terrain features such as *valleys*, *foot slopes*, *hollows* and *flats* were strongly represented as being dominated (> 81%) by *gentle slopes* with *low convexity*. Results for similar terrain features were reported by Doumit (2018) and Regmi & Rasmussen (2018) and with the landscape dominated by depositional processes, as expected for regions commonly associated with water accumulation and eroded material. These *valley bottom* regions are naturally likely to have greater potential wetness than the *steep slope,*

high convexity geomorphon features (King *et al.*, 1983; Regmi & Rasmussen, 2018) with depressional units generally composed of Gleyed Orthic or Gleyed Elluviated soils. In contrast, the results for the GDEM2 showed patterns for the same geomorphon features classified as *gentle slopes* with *high convexity,* i.e., areas of dispersion rather than regions of accumulation. **Figure 3.15** further highlights the influence of the implicit surface artefacts present in the GDEM2 and the impact of these anomalies on the slope gradient calculations for the modelled *TSC* results. The 2D profile description presented in **Figure 3.10** further outlines the "jagged" and broken GDEM2 terrain surface resulting in the over-prediction of *steep slopes* with *high* and *low convexity.* Interestingly, while the GDEM2 histogram suggests that these errors are not limited to any specific geomorphon feature but rather systemic across the entire image scene (landscape), we can see that the effect of the DEM errors on *TSC* appears to be less prevalent in geomorphon features with distinctive elevation transitions such as *ridges, summits, valley bottoms* and *hollows.* Notably, geomorphon features, such as *slopes* and *foot slopes,* can have similar ternary patterns and transitions between ternary elements (Jasiewicz & Stepinski, 2013) and therefore expected to be more typical, spatially contiguous and contain the most *TSC* variability. Recall that the final geomorphon map is a compilation of the ten most commonly recognisable terrain forms summarised from a whole set of 498 possible patterns with *slopes* and *foot slope* features typically accounting for almost 52% (256 geomorphons) of all reclassified geomorphons (Jasiewicz & Stepinski, 2013).

Let us consider the soil texture and soil depth results for the same geomorphon features. There is a direct link between the pedogenic and geomorphic processes since these low-lying regions are characterised by highly textured (clay > 35%) and deep profiled (ERD > 800 mm) soils which are expected for zones of material accumulation, rather than dispersion, within the landscape. Hall (1983) provides a thorough description of soils on specific landscape positions, emphasising the catena concept's contribution as the primary soil-landscape unit. Fundamental to the catena concept is the movement and distribution of water on slopes. The catena is understood to, *strictu sensu,* represent the primary cause of material movement on slopes and is considered the principal reason for differences in soils on the landscape. For this reason, lower-lying landscape positions such as *foot slopes* and *toe slopes* are dominantly constructional and relatively unstable. Hall (1983) outlined that material in these landscape positions is derived from up-valley positions and somewhat from superjacent footslope and backslope positions. This landscape feature concavity then results in upslope-deposition of particulate material and material carried in solution with seepage zones, with high water retention common. Drainage is typically lower than on higher positioned landscape features and may range to well-drained on higher microrelief positions. Also, soil pedon depths are expected to be variable but tend to increase downslope as longitudinal gradients decrease, and drainage channels debouch onto lower-gradient slopes.

Another noteworthy observation is that lower-lying geomorphon features are represented by 5 to 20% *gentle slopes* with *high convexity*. This may suggest the possible delineation and representation of certain discontinuous terrain features, typical of steep or abrupt breaks in the landscape such as riverbanks or levees, fluvial terraces, artificial road cuttings or even degradational features, e.g., incised gullies. These features are typical for the region, with most stream channels having steep gradients (Macfarlane *et al.*, 2014). The *valley bottom* features are typically very narrow with young valleys, and steep, well-drained slopes often extend to the channel bank. Furthermore, hill profiles are often stepped and particularly evident in the study region. The parent rock lithology causes this effect; erosion gives rise to sheer scarp face slopes interrupted at intervals by gently sloped terraces/foot slopes associated with hill contours (Bainbridge & Marneweck, 1996).

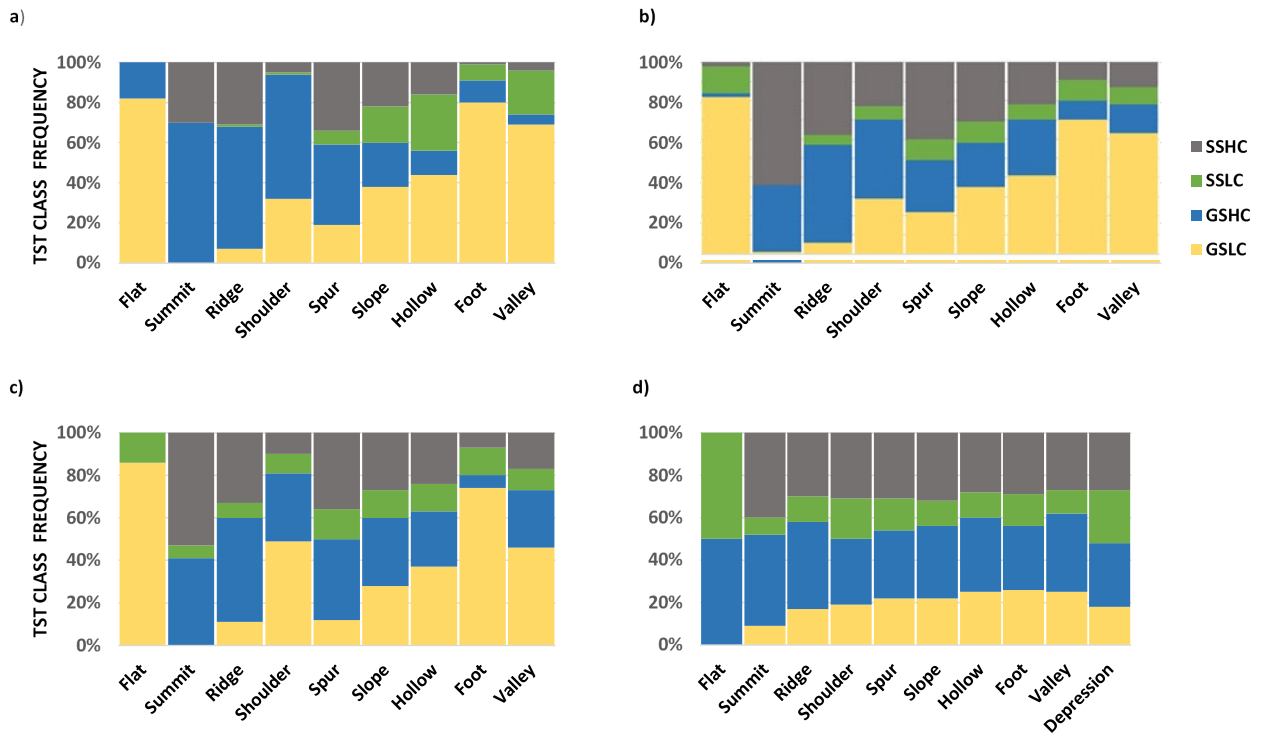


Figure 3.15: Histogram showing frequency and pattern of terrain surface classification for each geomorphon feature in the a) 30 m SUDEM b) 90 m SUDEM c) 90 m SRTM d) and 30 m ASTER GDEM2

Typically, in sloping areas and other ablation landscapes, the relationships between geoform and soil are more complex than constructed landscapes (Zinck, 2013). However, the geomorphic "signature" for the higher altitude features such as *summits, ridges, shoulders,* and *spurs* are characterised by a combination of *steep* and *gentle high convexity* slopes with *gentle* slopes more prevalent (30 to 60%). These results align well with the similar terrain morphologies described by Regmi & Rasmussen (2018) and indicate the overall gradual, undulating topography of the study site detailed in **Figure 3.16,** which provides a functional visual perspective of *TSC* derived from the respective DEM surfaces. In this region, these higher slope areas contribute to either shallow soil's pedogenic development with limited rooting depth within up-slope convex areas or relatively young soils with little or no profile development. Typically, these soils have clay enriched-subsoil located at the basal region of *ridges* and *slope* transitions where there may be a sharp gradation between deep and shallow soils (King *et al.*, 1983). Similarly, these areas usually include a combination of *Acrisols, Lixisols, Arenosols, Ferralsols, Planosols* and *Leptosols* (**Figure 3.11 and Table 3.11)** (Van der Eyk *et al.*, 1969; Camp, 1995).

Reflecting on these findings, including those of *soil complex*, *soil depth*, and *soil texture*, several critical drawbacks of geomorphons to soil-landscape applications are worth noting. First, it is evident that each geomorphon unit is associated with multiple soil property patterns, which means that while geomorphons may represent the generalised soil-landscape associations within the study region, there is extensive overlap in these associations. This suggests that these observed soil-landscape relationships are contributive rather than absolute. Put differently, multiple geomorphons share common soil-forming conditions (Zhang *et al.*, 2016). While much of the variability between the Tugela Basin's observed soil properties may readily be explained

by reference to specific DEM terrain parameters, factors other than those considered by geomorphons invariably drive soil differentiation within the study region.

Secondly, as Erskine *et al.* (2007) point out, the variability of geometric signatures derived from different DEM surfaces will result in varied geomorphon descriptions (**Figure 3.3** and **Figure 3.16**). For this study, the I set the same thresholds for local convexity to all DEMs to extract the standard set of geomorphon features, with no priority or importance for any DEM surface. This resulted in geomorphons modelled at 30 m and 90 m geomorphon representing similar landscape-scale trends in *soil depth*, *soil texture* and *terrain surface convexity*, confirming the scale-invariable value of the geomorphon approach utility to regional soil-landscape analysis. However, further optimised threshold values will likely output better classification results for other purposes or specific terrain units when coupled with other DEMs. Finally, selecting the optimal model parameters for each classification method is dependent on spatial resolution, quality of DEM, characteristics of relief in the study area and spatial variability of pedological units, which are to be predicted (Zhang *et al.*, 2016). In this case, the GDEM2 geomorphon surface has proven inadequate in the systematic application to soil-landscape analyses for this region due to its pervasive data quality issues, making it unsuitable for mapping larger areas outside the region. The findings of this research further highlight that DEM selection must consider the intended purpose of the application. The use of DEMs "not fit for purpose" could severely compromise the digital representation and extraction of geomorphometric properties. Stated differently, while the GDEM2 may be a suitable option for selective visualisation of a three-dimensional surface, it may not be a convenient option for analysing derived terrain attributes, particularly in mountainous regions of the South African interior.

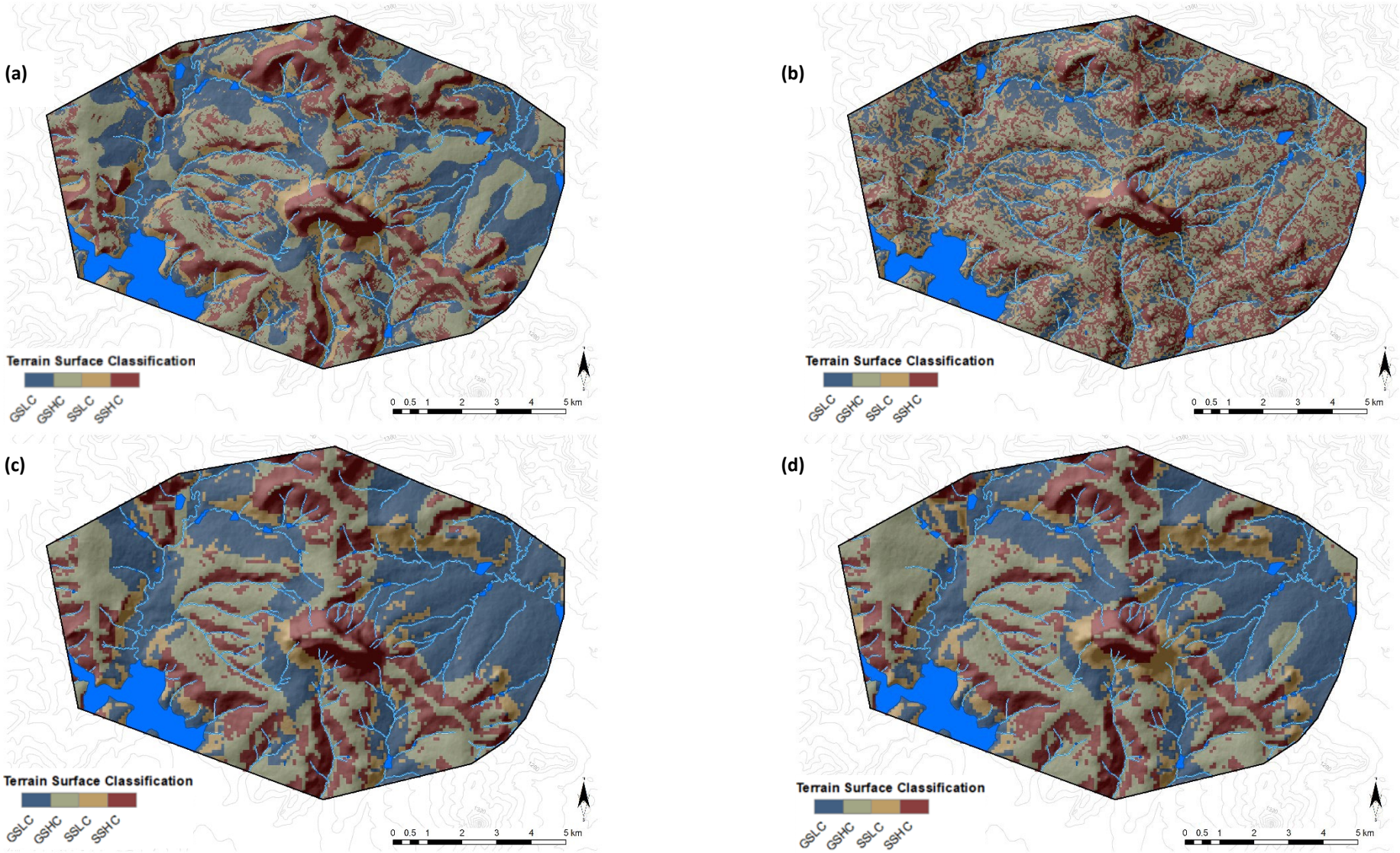


Figure 3.16: Modelled *TSC* for the Bergville Study site, a) 30 m SUDEM b) 30 m ASTER GDEM2 c) 90 m SUDEM d) and 90 m SRTM.

3.4 CONCLUSION

Much reliance has been placed on pattern recognition for terrain discontinuity segmentation, classification, and mapping. It is further supposed that geomorphological maps that can better define the soil-landscape relationship can be improved using the medium to high-resolution DEM datasets. Terrain classification by geomorphons has been a practical approach to significantly enhancing soil-landscape characterisation in the Drakensberg interior region. In this study, the use of DEMs, GIS, and the open-source geomorphon approach has been fast, feasible, and user-friendly for landform classification in Central Drakensberg's highland region. Continual advances in quantitative modelling of surface processes, combined with new Spatio-temporal and geo-computational algorithms, have revolutionised the auto-classification and mapping of landform components through the automated analysis of high-quality DEMs. This chapter provides an appealing outlook of geomorphon characterisation across varying DEM resolutions to highlight how different DEM source and spatial resolutions influence the representation of soil-landscape pattern and processes phenomena. A significant part of this research involves identifying the relationship between geomorphon surfaces and terrain character and the similarities in terrain products across these scale-specific geomorphon surfaces.

This research provides a comprehensive exploratory assessment of digital terrain representation and relief classification using an automated geomorphometric mapping approach by evaluating three different digital surface models (SUDEM, SRTM, ASTER GDEM2) and different spatial resolutions (30 m & 90 m) for an 11200-ha catchment in KwaZulu-Natal, South Africa. The study further tested the notional scale-independence property of the geomorphon approach for selected soil (*soil complex*, *soil depth* and *soil texture*) and terrain parameters (*elevation, slope* and *terrain surface classification*) with varying DEM resolution using 4 750 gridded terrain samples. Here, I set out to quantitatively analyse how the terrain model and scale influence the extraction, generalisation and representation of digitally derived terrain attributes such as slope, elevation and terrain unit feature extent. Evidently, it is shown that the variation in resulting terrain unit representation is limited by spatial resolution discontinuities of selected elementary *soil complex* distribution, *soil texture* and *soil depth*. The regional geomorphon-soil relationships' findings are presented in a readily interpretable and qualitative manner, providing a “quasi-landscape signature” for potential localised geomorphons

What is promising for representing geomorphic associations with soil properties is that despite the coarse feature presentation of properties like *soil texture*, *soil depth*, and *terrain classification* derived from infield assessment, the observed soil-landscape trends across the geomorphon features are predominantly well-aligned except for the GDEM2 surface products. Moreover, the results for both the *soil texture* and *soil depth* are highly co-linear to their respective broad soil complex units. The results suggest that geomorphons may have suitable application at a landscape and regional level rather than a local farm level. While the categorical assessment of soil features in this study may not definitively characterise the process of each geomorphons surface for complete landscape character assessment, there may be a benefit to the synoptic delineation of a geomorphic pattern signature for terrain units used for defining agricultural land potential and categorisation for a variety of land use planning functions. DGM practitioners should, therefore, be aware of the implication to land use planning when integrating elevation datasets of known inherent data deficiencies for soil-landscape and digital soil mapping analysis, especially under South African conditions.

The presented results and discussion compare the modelled geomorphons surfaces, i.e., 30 m ASTER GDEM2, 90 m SUDEM and 90 m SRTM DEM, to a reference 30 m DEM surface derived from high resolution 5 m SUDEM, digital terrain model. A pragmatic yet visually compelling similarity index (BK composite measure)

is tested in this study to evaluate the association between the derived geomorphon surfaces. The BK metric's main advantage is that it allows two or more generalised maps to be compared, regardless of spatial resolution, with either serving as a reference map with the obtained values representing the degree of similarity between the two maps. Further statistical analysis included applying the chi-squared goodness of fit independence test for soil complex similarity between the different geomorphon surface products. The results showed positive associations between selected geomorphic features and soil-landscape covariate layers. Specifically, the main findings of the study worthing noting include:

- The application of medium and coarse resolution DEMs for geomorphon feature discretisation holds promise for soilscape studies in South Africa, particularly in the central regions where open-source DEM datasets such as SUDEM and SRTM are still a primary source of quality terrain data.
- These results indicate that some replication, albeit at an individual geomorphon feature level and not complete landscape level, is possible between a predicted and reference surface even with coarse DEM resolutions of 90 m, even if the DEM surface contains inherent artifactual inconsistencies
- Geomorphons represented the expected altitudinal and slope gradient properties of known terrain features characterised by higher altitudes, i.e., *ridges, spurs*, and lower-lying features, i.e., *valleys, foot slopes*, albeit in varying extent and frequency of occurrence across all DEM surfaces.
- However, this study has also shown that while most geomorphon features are detectable across DEM surfaces of varying resolution, the geomorphon characterisation accuracy is not linearly correlated with pixel resolution, *i.e.,* finer resolution DEMs do not necessarily produce “improved” geomorphon surface representation.
- Geomorphon users need to reconcile the choice of optimal scale for simple static terrain visualisation versus the representation of spatial features and associated environmental phenomena or processes across the landscape.
- Geomorphon feature relevance for defining landscape structure and terrain spatial heterogeneity must be framed in the context of landscape or terrain detail, soil covariate membership, DEM pixel resolution and user preference.
- The findings highlight how the representation of legacy soil-landscape covariates within geomorphons may contribute to the concept of present landscape pattern-process character assessment, i.e., the distinct, recognisable and consistent pattern of elements in the landscape that differentiates one landscape from another.
- These results further suggest that when using legacy soil data for geomorphon characterisation and comparisons for soil-landscape analysis, users must consider the measurement scale for the terrain parameters under observation, i.e., discrete vs continuous.
- The results for the respective GDEM2 geomorphons are not aligned to accepted generalised soil-landscape patterns for the region, which presents a challenge in defining the utility of the GDEM2 for other geomorphon soilscape applications regionally.
- Additionally, while the GDEM2 may be a suitable option for selective visualisation of a three-dimensional surface, it may not be appropriate for analysing derived terrain attributes, particularly in mountainous regions of the South African interior.
- The findings highlight that DEM selection must consider the intended purpose of the application. The reliance on DEMs that are not fit for purpose could severely compromise the digital representation and analysis of geomorphometric properties.

Information and methods discussed in this chapter will be valuable for landscape and suitability studies, especially at the regional level. The ability to quantify the overall similarity between different landscapes by merely evaluating the descriptive data of the associated and derived geomorphon surfaces is a critical element

of the current methodology. The accurate representation of the study site's contemporaneous soil and terrain properties provided the first approximation for guiding the description and segmentation of the geomorphon "fingerprint" for each DEM surface evaluated for this study. The SUDEM is perhaps the most accurate and readily available localised high-resolution DEM for South Africa. This study has shown why it remains a popular choice for various terrain analysis, DSM and hydrological modelling applications. The ease of automation, swift replication and acceptable representation of an expanded set of morphometric classes using geomorphons have shown to be a suitable alternative to the basic 5-class Land Type discrete relief models for South Africa. These geomorphons coincide well with actual terrain geomorphological entities and present better opportunities for soil-landscape characterisation and topographic quantification with considerable potential for future digital geomorphological applications for the central regions of KwaZulu-Natal.

CHAPTER 4

4 UNRAVELLING REGIONAL GEODIVERSITY: A GRID-BASED MAPPING APPROACH TO QUANTIFY GEODIVERSITY IN THE UTHUKELA DISTRICT, KWAZULU-NATAL[4]

"The interaction between geomorphic and ecologic landscape components has been largely conceptualised without attention to soils. The diversity of soils and landforms has marked quantitative and qualitative effects on the landscape. Geodiversity (geomorphic and lithological diversity) and pedodiversity together have strong influences on the architecture of an ecosystem. In order to assess landscape diversity, it is necessary to estimate the diversity of individual landscape components as the constitutive elements of landscape complexity and heterogeneity. Only through amalgamating the component diversities will an overall measure of landscape diversity be achieved."[3]

[3] Ibáñez, J.J., Vargas, R.J., Vázquez-Hoehne, A. (2013). Pedodiversity state of the art and future challenges. *Pedodiversity, CRC Press (Taylor and Francis Group) Boca Ratón. California*, 1-28.

[4]This work was published as a book chapter: Atkinson, J.T., de Clercq, W.P., 2020. Unravelling regional Geodiversity: a grid mapping approach to quantify Geodiversity assessment in the uThukela District, KwaZulu-Natal. In: Adelabu, S., Ramoelo, A., Olusola, A., Adagbasa, E (Ed.), Remote Sensing of African Mountains – Geospatial Tools Towards Sustainability (*In Print*)

[4]The work was also presented at the Conservation Symposium, 02-09 November 2020, Pietermaritzburg, South Africa (Virtual).

4.1 INTRODUCTION

The concept of geodiversity has gained significant international recognition with holistic decision-making application(s) in the last decade (Betard & Peulvast, 2019). The interpretation and integrated application of geodiversity have systematically undergone an evolutionary adaptation in the context of geoscience since its genesis between 1991 and 1993, admittedly to the benefit of a broader scientific audience (Araujo & Pereira, 2018). First introduced at the International Symposium on the Conservation of Geological Heritage (Panizza, 2007) and then further unpacked during the Malvern Conference of Geological and Landscape Conservation (Sharples, 1993; Sharples, 1995) and (Serrano & Ruiz-Flaño, 2007); the concept of geodiversity has undergone significant theoretical and applied undulation in an attempt to establish its own identity beyond its initial generic application linked to geological diversity and biodiversity conservation (Pellitero *et al.*, 2015). It is well accepted that both geodiversity and biodiversity are concepts that emerged from the 1992 World Biodiversity Convention in Rio de Janeiro, Brazil, and have since been embraced by numerous countries (Brilha, 2005; Gray, 2008). Undeniably, however, geodiversity (abiotic complexity) has not received the same level of attention as its biological (biotic complexity) counterpart despite its intrinsic and indivisible linkages to both ecosystems and landscape richness characterisation (Pellitero *et al.*, 2015). Natural ecosystem diversity is understood to be the combination of these two constituents: both biological and physical complexity of nature is necessary to determine the individuality of the landscape of a region, country and even continent (Kostrzewski, 2011; Santos *et al.*, 2017). Manosso & de Nóbrega (2016) further points out that like biodiversity, geodiversity is not a constant but is instead adapted to a given moment, place or region.

While it is essential to acknowledge the synergies between bio- and geodiversity, it is equally important to recognise that geodiversity is of inherent value in itself, as a significant driver of many environmental processes that require further affirmation on its characterisation, spatial distribution and systematisation of mapping techniques to quantify better and evaluate the full scope of its capability (Manosso & de Nóbrega, 2016; Zwoliński *et al.*, 2018). Recently, Atkinson *et al.* (2020) outlined that geomorphic classification and delineation need to be better represented in discrete landform mapping endeavours in South Africa. Central to this directive is the shift away from an ad-hoc to a more formal and systematic approach to landscape character assessment (Wascher, 2005), using modern geographic technologies and incorporating better base maps of topography into the landform mapping process (Miller & Schaetzl, 2014). Here landscape character is defined as the distinct, recognisable and consistent pattern of elements in the landscape that distinguish one landscape from another (Swanwick, 2002). Therefore, this definition underlines the explicit recognition of individual landscape elements that constitute the landscape, thus allowing for a systematic comparison of areas based on their landscape character (Galatowitsch *et al.*, 2009).

Owing to its usefulness as a driving mechanism for habitat variation and how widely it can be employed, the convenience of the geodiversity concept has gained reputable exposure beyond that of its inaugural application centred on geological site (geosite) heritage conservation (Sharples, 1993) and pedodiversity (Ibáñez *et al.*, 1995; Alba Alonso *et al.*, 1998; Thwaites, 2000). Most notably, recent geodiversity studies highlight additional contributions in the disciplines of biodiversity and ecosystem development (Parks & Mulligan, 2010; Gordon *et al.*, 2012), geomorphodiversity (Zhang *et al.*, 2003), geotourism (Serrano & González Trueba, 2011), geoconservation (Gray, 2004), environmental policy management and even geodiversity action plans (Burek & Potter, 2006; Prosser *et al.*, 2011). Of particular relevance to the present study are the works by Serrano & Ruiz-Flaño (2007), Benito-Calvo *et al.* (2009), Hjort & Luoto (2010), Araujo & Pereira (2018) and Betard & Peulvast (2019), highlighting how the combination of geological, geomorphological, climatic and hydrological information can be used to quantify geodiversity at a regional scale.

Given the broad scope of geodiversity application, it is not surprising that there is limited *lingua franca* for defining the concept and its methodological assessment. Several pragmatic and mainstream definitions for geodiversity, notably that of Sharples (1993), Dixon (1996), Kozlowski (2004) and Zwoliński (2008), dominate the progressive geodiversity pedagogy. However, this study follows the definition(s) of the concept of geodiversity endorsed by Gray (2004) as a departure point for scope of assessment: "*The natural heterogeneity (diversity) of geological (rocks, minerals, fossils), geomorphological (landforms, topography, physical processes), soil and hydrological features. It includes their assemblages, structures, systems and contributions to landscapes*", and Ruban (2010) considers the assessment of geodiversity "*as a numerical expression of geocentric entity diversity*". For reasons of clarity, consistency and simplicity, it is necessary to acknowledge the interrelationship of geodiversity with the following commonly used concepts (adapted after (Crofts *et al.*, 2020):

- *Geoconservation,* the conservation of geodiversity for its intrinsic, ecological and geoheritage value;
- *Geoheritage* is an element of geodiversity, whether singular or combined, that has significant value for intrinsic, scientific, educational, cultural, spiritual, aesthetic, ecological or ecosystem reasons and therefore deserves conservation.
- *Geotope,* Geomorphological or geologic associations within a particular spatial area. A geotope is the geologic equivalent of an ecotope

Knight *et al.* (2015) caution that the use of these different terms may lead to a cause of confusion and limitations in potential geodiversity application in the context of digital geomorphological mapping. While the nature of 'heritage' differs considerably among people, between places and over time, it invariably stems from nature and culture, and our attachments to it are universal (Lowenthal, 2005). This concept is well defined in the Digne Declaration (Crofts *et al.*, 2020) - *Our history and the history of the Earth cannot be separated. Its origins are our origins, its history is our history, and its future will be our future.* The pluriform manifestations of geoheritage and geoconservation appear universal and straightforward. However, the myriad tensions that can arise when defining heritage, where values, objectives, and expertise belong, and whom the intended beneficiaries are for protecting it can be challenging to resolve (Graham *et al.*, 2000; Rassool, 2013). Generally, it is understood that geoheritage features must have unique geological, cultural or geomorphological values (Gray, 2004). Although a descriptive version of geodiversity is useful for geoconservation, all relevant management activities require a numerical expression of geodiversity quantification (Ruban, 2010). This is the second major limitation of geodiversity assessment in South Africa. No formative approach still exists to characterise priority patterns or exclusion patterns numerically or geospatially (Cocks et al., 2018), i.e., areas characterised by rich or poor geodiversity and low or high significance. Recently Kori *et al.* (2019) piloted the application of a Geomorphodiversity Index in the Soutpansberg Range, South Africa, using the geodiversity assessment criteria proposed by Zwoliński *et al.* (2018). Their geospatial quantification of geodiversity provides a baseline for the present study, with substantive semblances regarding the recognition of geodiversity factors. The present and aforementioned referenced study diverge; however, in the application concerning the quantification of geodiversity, i.e., geodiversity based on feature richness and abundance.

This divergence of methods remains a challenge to both geodiversity's applied and theoretical unanimity and further extends to its valuation, both numerically and descriptively. Pereira *et al.* (2013) conveniently diagnose several key-methodological points that remain unresolved: these include which criteria to use to assess geodiversity? How to rationalise scale-factor dependencies, i.e., how does the size of the area under analysis influence the type of criteria, and how should the results for the given methodology then be presented?

Encouragingly, however, Piacente (2005) stress that the equivocation regarding geodiversity and the concepts that underline it should not be seen as the "stick" but rather the "carrot" for expanded dialogue and continual exploration in light of its perspectives and interpretations.

Following Ruban (2010), numerous quantitative methods for geodiversity assessment have been proposed in the last decade. According to Zwoliński *et al.* (2018), geodiversity assessment procedures are relatively subjective and informed by the observer's knowledge and experience. Equally, these procedures can be selected and adapted to the object or phenomenon being analysed. Church (2011) further cautions that geodiversity studies are beset by the same methodological issues analogous to scientific observation in geomorphology: sampling, measurement, scales, scaling, classification and residual errors. Research out of South America (Serrano & Ruiz-Flaño, 2007; Benito-Calvo *et al.*, 2009; Pereira *et al.*, 2013; dos Santos *et al.*, 2020) and Europe (Zwolinski, 2018) present the most state of the art interdisciplinary approaches for geospatial, mathematical and statistical geo-indexed geodiversity assessment and geo-visualisation. The benefits of employing quantitative and geospatial methods to geodiversity assessment should be apparent. Melelli *et al.* (2017) outline these benefits: First, quantitative approaches are both repeatable and objective, allowing for useful comparisons of areas in different geographical backdrops. Second, ranking geodiversity using a Geodiversity Index (*GDIx*) facilitates practical gridded planimetric geo visualisation and cartographic approaches in a geographical information system (GIS), allowing areas of similar *GDIx* values to be harmonised. By the same token, representing *GDIx* as a numerically ranked grid-based digital dataset would allow overlaying the *GDIx* with other spatial covariate information of identical grid size, reaching a *GDIx* value per areal sample unit and allowing the comparison of *GDIx* values in different cells of the grid as well as identifying areas of low or high geodiversity.

The point of departure for this research is the understanding that geodiversity is increasingly recognised to underpin and deliver essential ecosystem services (landscape functioning) linking people, landscapes and their heritage to the broader benefit of the natural and built environment as outlined in the Millenium Ecosystem Assessment (MEA, 2015). Comer *et al.* (2015) showcased that land units with high geodiversity are relatively resistant to degrading living conditions and well equipped to cope with disturbances such as climatic changes compared to low-geodiversity land units. Therefore, most important to decision-makers is prioritising the endogenous linkages and mapping the spatial variability between geodiversity and geological, hydrogeological, topographic, climatic, geomorphological and pedological covariates. This presents a far more holistic approach to mitigating cross-cutting issues of current concern, including habit loss and food insecurity in the face of increased natural and anthropogenic pressures linked to economic development (Grab & Knight, 2015).

In fact, Lilburne *et al.* (2020) poignantly reaffirm this didactic outlook stating that the growing demand on land-based industries and land managers to balance the need for economic prosperity with a greater focus on sustainable landscape management is best achieved through the use of comprehensive land information systems that can assess the benefits, impacts and trade-offs of land use decisions at varying temporal and spatial scales. While the I am aware of the numerous geodiversity studies conducted internationally (see Zwolinski et al., 2018 for a thorough review), only a select few have attempted to quantify South Africa's geodiversity importance, notably the work by Kori *et al.* (2019). Typically, most national studies on geodiversity have focused heavily on the importance of geosite (Ruban, 2010) diversity for heritage preservation (Knight *et al.*, 2015). Therefore, the present work aims to perform a semi-quantitative assessment of geodiversity in the entire uThukela District Municipality located in KwaZulu-Natal, South Africa. The specific application of quantitive geodiversity assessment is yet to be fully explored in the Southern Africa context, where various landscape pattern-process functions inevitably influence a unique set of soil-landscape

and pedo-hydro-geomorphic processes (Partridge *et al.*, 2010; Grab & Knight, 2015; Holmes *et al.*, 2016). Foremost this work seeks to provide a first approximation methodology for geodiversity assessment adapted to the regional context with possible scaling-up potential to a National Level. A functional application of the concept is presented by producing a *GDIx* map highlighting the richness of selective (limited) abiotic elements at a landscape scale. The *GDIx* digital dataset would then be a flexible decision-support tool, even for the most agnostic users, allowing straightforward interpretation regardless of specialist background. I further explore a modified methodology initially designed by Pereira *et al.* (2013) and later refined by Araujo & Pereira (2018). Here, the methodology describes the degree of geodiversity as objectively as possible and then maps its distribution from a panoptic context. More specifically, a two-pronged approach is investigated as follows: first, I will leverage a suite of geospatial analytical tools offered by GIS platforms to overlay a regular grid onto different thematic covariate categories that take into account hydrographic, geological, pedological, climate, topographic, atmospheric and geomorphological information to obtain a final *GDIx* calculated from these partial thematic indexes. A key consideration of this research is thus ensuring synoptic replicability across the landscape. In this regard, all attempts are made to control model simplicity at minimal expense to represent each input element (partial index) diversity. After that, an exploratory approach is used to highlight the capability of the derived *GDIx* dataset as a decision-support tool for the entire uThukela District Municipality. The resulting outcome relates the consequent diversity distribution to several strategic land use planning spatial coverages currently in active practice, namely: *National Land Cover (NLCov)*, *National Land Capability (NLCap)*, *Provincial Bio-Resource Units (BRU)*, *National Biodiversity Conservation Status (NCS)*, *National Terrestrial Biomes (NTB)* and *Biome Protection Status (BPS*).

Beyond the relative regional merits for decision-makers and practitioners, the novelty of this research is possibly its contribution to unravelling certain application enigmas for the discipline. The intention is to cement specific affirmations regarding optimisation gained with GIS, outline limitations due to feature scale, define appropriate abiotic feature selection and highlight ease of cartographic interpretability. I believe that the study's findings are pursuant to the broader international discussion on the evolution of quantitative geodiversity assessment methods as a valuable, holistic, and replicable spatial decision support tool.

4.2 MATERIALS AND METHODS

4.2.1 Regional settings

The study area is the uThukela District Municipal area located between the Kingdom of Lesotho in and the western boundary of the coastal Province of KwaZulu-Natal (KwaZulu-Natal), South Africa, with the seat for the Municipality situated in the town of Ladysmith (*28°33'35"S 29°46'50"E*) (**Figure 4.1**). The uThukela District Municipality (UTDM) derives its name from one of the major rivers in KwaZulu-Natal, the uThukela River, that rises from the Drakensberg Mountains and supplies water to a large portion of KwaZulu-Natal and as well as Gauteng Province (Henning *et al.*, 2014). Approximately 11 500 km^2 in extent, the UTDM provides a valuable case for designing a geodiversity assessment map in a mesoscale context. The region presents an eclectic combination of high-value natural resources with competing demands at different scales. Cox *et al.* (2015) highlight the competition for using resources between local users (livelihoods) and agricultural production and international demands from the tourism sector and biodiversity conservation necessary for maintaining sound natural systems.

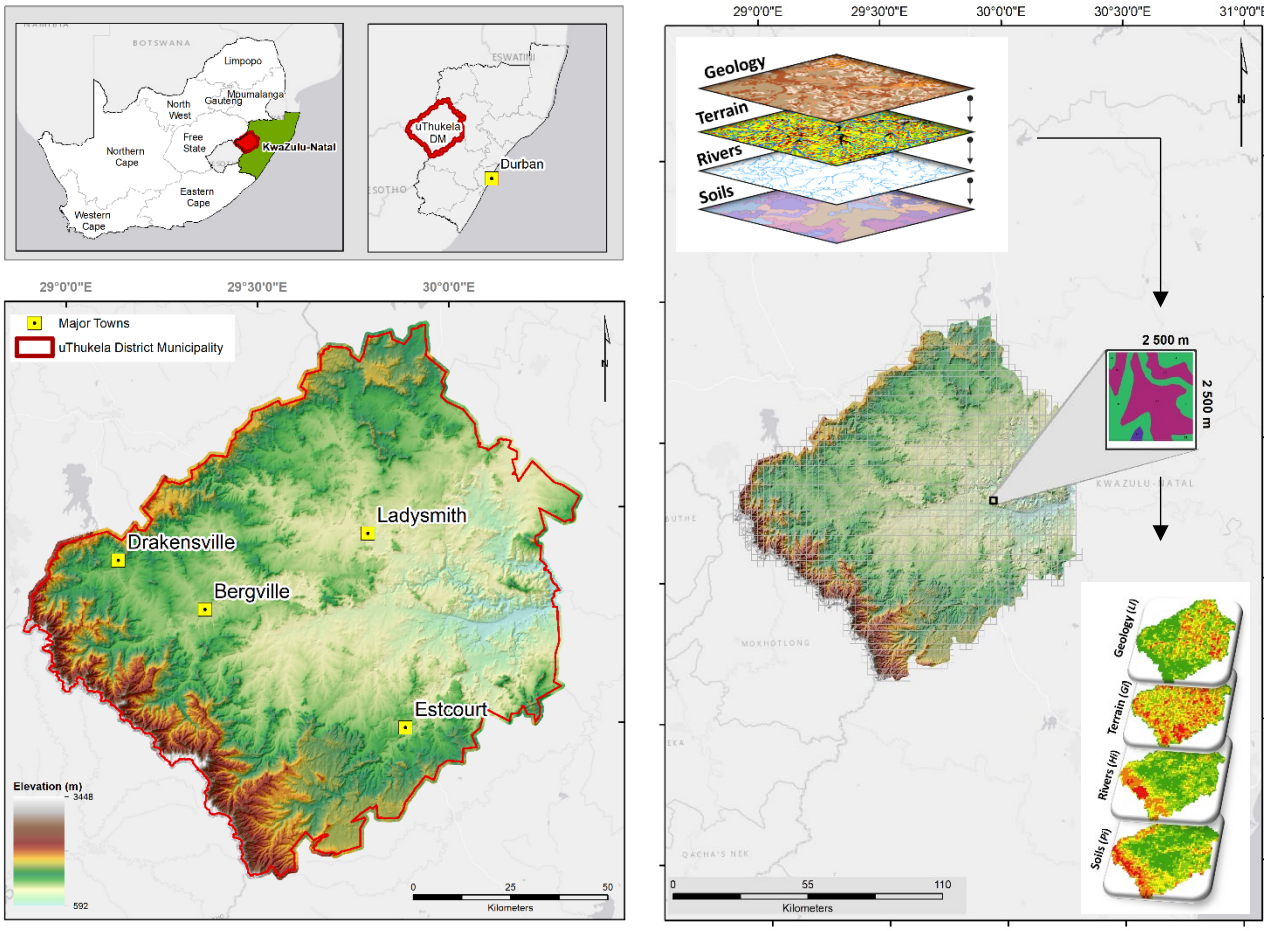


Figure 4.1: (a) Map showing the Uthukela District Municipality (UTDM) situated in the western region of KwaZulu-Natal, South Africa (DEM Source: CGIAR, 2014). **(b)** The conceptual workflow uses gridded datasets to derive geodiversity sub-index data layers for the UTDM region overlaid with the 2.5 x 2.5 km sample grid.

Furthermore, the area also includes a rich cultural heritage left by the indigenous San People. The rock art of the Maloti-Drakensberg Park is the largest and most concentrated group of rock paintings in Africa, south of the Sahara. It is outstanding both in quality and diversity of subject (UNESCO, 2000). The region is almost entirely rural. The dominant land use in private and tribal lands is commercial and small-scale agriculture, respectively (EKZNW, 2010). Most small-scale farmers practise extensive livestock grazing, dryland cropping, and some vegetable gardening, while private agricultural operations are large scale, more diverse, more productive, and strongly commercially oriented.

Varying climatic conditions dominate the District, with the region characterised by a combination of two climate classifications: temperate, dry winter, warm summer rainfall (Cwb) and arid, steppe, hot climates (BSh) (Conradie, 2013). Mean Annual Precipitation (MAP), predominately thunderstorms and snow, ranges from 576 mm to 1923 mm, with large scale variations over relatively short distances not uncommon. Typical Mean Annual Temperature (MAT) range from approximately 4ºC to 17°C, with mean monthly maximum and minimum temperatures ranging by up to 42.5°C for December and June in parts, highlighting significant temperature deviations between summer and winter periods (Elliot & Escott, 2015)

The topographic expression of the UTDM is broadly diversified from high mountains in the south-western region of the Province, becoming more undulating as it forms part of the Thukela catchment and drains to marine estuaries and coastal dunes towards the Indian Ocean in the east. The Drakensberg, the highest lying component of the study domain, includes three altitudinal zones (the Montane zone, the Sub-Alpine zone, and the Alpine Zone) extending from approximately 1 300 m to 3 500 m above sea level, which encompasses the steepest altitudinal gradient in the District. Elliot & Escott (2015)characterise the landscape as an assortment of gently undulating hills through to a rolling and partly broken landscape across a notable altitudinal gradient extending through the District (600 m to 1 300 m). In the higher altitude regions, typical terrain morphological features include rocky, rugged slopes and terraces, including mountainous areas incised by river gorges, a variety of narrow and broad valleys, plateaus with sharp hills, steep ravines and escarpment slopes, high mountain ridges separated by cavernous valleys, prominent cliff faces, complex mountain topography, and steep basalt rock faces and terraces. In the lower declivity regions meandering stream channel and paludal floodplains with drier ephemeral stream and semi-arid alluvial fans and aeolian dunes are most prevalent (Botha & Singh, 2012).

The Drakensberg Mountain Range encompasses a wide range of geological formations. These include a diversity of Karoo Supergroup rocks, including Stormberg basalts, tillite of the Dwyka formation, as well as Ecca and Beaufort Group formations. Geological formations are typically either sedimentary or igneous in origin. The region is mainly dominated by various sandstone, shale, mudstone, dolerite, quartzite, dolomite, granite, diabase, and basalt saprolitic material. Moreover, altitudinal variations highly influence geological exposure, especially for the well-layered Karoo Supergroup (Elliot & Escott, 2015). The geological and geomorphological diversities are reflected in the pedodiversity of the area in terms of pedon-depth, topsoil depth, drainage, fertility and soil texture. Typical soil groups include, in varying spatial assemblages and without limitation: *Histosols, Gleysols, Vertisols, Umbrisols, Fluvisols, Luvisols, Lixisols, Planosols, Plinthosols, Ferralsols, Arenosols, Leptosols, Cambisols* and *Stagnosols*.

EKZNW (2010) highlights the exceptional habitat heterogeneity in the UTDM, which translates to bounteous vegetation diversity, comprising lower altitude dense bushveld, savanna and grasslands, extending to high altitude montane and alpine grasslands, including significant pockets of indigenous forests. The District includes four biomes and 22 vegetation types, including the only Afro-alpine vegetation in Southern Africa, which borders South Africa and Lesotho. The District is also strategically significant as it makes up the

principal catchment area for the Thukela River. Two major impoundments occur along the Thukela River within the central-western regions: Woodstock Dam and Spioenkop Dam. Another critical natural impoundment, namely Wagendrift Dam, occurs along the Boesmans River within the southern area of the UTDM. This rich biodiversity and cultural value of the Drakensberg is the basis for its World Heritage Status and includes both the uKhahlamba National Park and the Maloti-Drakensberg Trans-frontier Peace Park between Lesotho and South Africa, not to mention wetland areas of Ramsar Conservation status as well. It is worth noting that all World Heritage sites are designated as Protected Areas, which means that mining or prospecting are not permitted within the property or proclaimed buffer zone. Further, any unsuitable development that may adversely affect the property is prohibited by the South African and Lesotho Ministers responsible for Environment and Culture (UNESCO, 2000).

4.2.2 Geodiversity classification

Digital inventories of landscape resources are vital for landscape management, as one cannot protect that which is unknown (Jackson *et al.*, 2019). Geodiversity quantification is calculated using the *Geodiversity Index* (*GDIx*) methodology developed by Araujo & Pereira (2018) with minor modification. The approach is based on the definition of partial numerical indices estimated from selected thematic geo-informational coverages considered to represent the main components (direct and indirect) of geodiversity segmented within a regular grid overlay. The grid-based framework provides a simple solution for assessing the degree of geo-richness. It avoids any element's ordinal classification and overrating by considering each part's quantification as equally important aggregates to the final *GDIx*. Holistically then, the *GDIx* is estimated from the sum of the partial indices calculated from the discrimination of occurrences (count) in each grid cell. Hjort & Luoto (2010) further emphasise the benefits of a grid-based *GDIx* approach in that it enables an objective subdivision of geospatial richness, creating units of identical dimensions, allowing direct comparison of results to explore the relationship of geodiversity with a diverse range of spatial abiotic or biotic variables. The implementation of the *GDIx* assessment is further explained in the forthcoming subtopics.

4.2.2.1 Grid resolution

Geodiversity cannot be understood without the proper definition of scale: regarding the area under observation and the elements targeted for assessment (Serrano & Ruiz-Flaño, 2007). For this study, I applied a regular cartesian grid of 2 500 x 2 500 m resolution using ArcGIS Desktop 10.5 (ESRI, 2021)– hereafter ArcMap, generating 2 042 grid squares. This approach is considered the most pragmatic parametrisation for optimal differentiation of maximum range between the highest and lowest partial indices and ultimately *GDIx* values for the UTDM (Zwolinski, 2018). Moreover, the grid resolution of the *GDIx* map was defined considering the map scale(s)/ resolution of the input geospatial layers used to perform the assessment and enabled the inclusion of coarse-scaled thematic map data (**Table 4.1)**. Due to the significant diversity within a landscape, finding an objective measurement unit of geodiversity quantification while at the same time giving comparable results remains an ongoing subject in the study of geodiversity. Given the multiplicity of approaches used to assess geodiversity, studies that provide empirical and taxonomic standardisation methods for assessing geodiversity must be acknowledged. Zwolinski (2018) provides a valuable look-up summary of landform taxonomic hierarchy based on the area under observation with recommendations on suitable sample grid resolution.

Similarly to Kozlowski (2004) & Benito-Calvo *et al.* (2009), the mesoscale (1:750 000 map scale) application of data was based on prioritising the minimum detectable feature range of elements of geodiversity as well as the overall extent of the UTDM study area. It is worth noting that the gamut of scale dependency in geodiversity, and its commonly used landscape correlates, is still very much underdeveloped conceptually, spatially and empirically (Hjort & Luoto, 2010). While some universal conclusions have been drawn by authors such as (Serrano & Ruiz-Flaño, 2007; Benito-Calvo *et al.*, 2009; Araujo & Pereira, 2018), without any formalised ontology on the scale-diversity dependencies, primary considerations are for the discernment of the end-user. Typically, small-scale maps are not suitable for detailed local studies. The grid-resolution selected for quantifying geodiversity in the UTDM follows the recommended grid resolution of 100 to 2 500 m (with a map scale of between 1:500 000 and 1:1 000 000) for regions between 2 500 and 10 000 km^2 (Zwolinski, 2018).

In contrast, the detail of large-scale maps is not appropriate for use in national or regional studies since the latter would not cater to the effective cartographic representation of landscape geodiversity. The primary importance of scale consideration is that it directly influences the complexity of geodiversity interpretation, i.e., geodiversity of objects, elements, places or landscape. Beyond providing a set of standard guidelines for defining these geodiversity-scale dependencies, Zwolinski (2018) further proposes a 9-level scale to evaluate the taxonomic hierarchy of geodiversity assessment: microform, landform, set of landforms, type of landform, morphological landscape, geomorphological region, morphogenetic Province, morphogenetic zone and geomorphic zone. Finally, Pellitero *et al.* (2015) highlight a central component for appropriate scale consideration beyond data extraction, generalisation and representation, namely application. Quantitative assessments of geodiversity for areas above 300 000 km^2 on scales as high as 1:250 000 or 1:500 000 present unique implementation and application challenges for veritable management as “hotspot” subject to prioritisation turn out to be entire regions or squares as extensive as several municipalities. In this instance, this is ideal for the present study focusing on the regional assessment of geodiversity in the entire UTDM.

Table 4.1: Fundamental components of input data, and their sub-types, used for quantitative assessment of geodiversity in UTDM

Partial Geodiversity Index	Sub-partial Data Layer(s)	Resolution/ Scale	Source	Format	Geodiversity relevance	Determination of Diversity
Hydrographic (***Hi***)	Groundwater recharge	1.7 km	Schulze *et al.* (2007)	Raster	Variability in annual baseflow	Modal zonal statistic of mean annual baseflow (mm) per grid cell
	Stream density	1:500 000	Dept. of Water Affairs, DWAF (2020)	Vector	Stream density	Sum of stream length per grid cell
Lithological (***Li***)	Lithostratigraphic	1:250 000	Council for GeoSciences, CGS (2000)	Vector	Geological richness	Count of different lithostratigraphic formations per grid cell
	Paleo-sensitivity	1:250 000	Council for GeoSciences, CGS (2000)	Vector	Geological sensitivity	Count of different paleo-sensitivity classes per grid cell
Pedological (***Pi***)	Land Type Soil Association	1:250 000	Institute for Soil, Climate and Water, ARC (2003)	Vector	Pedological richness	Count of different soil classes per grid cell
	Profile soil depth	1:250 000	Institute for Soil, Climate and Water, ARC (2003)	Vector	Soil depth > 800mm	Count of features > 800mm soil depth
Climatic (***Ci***)	Rainfall intensity	1.7 km	Schulze & Maharaj (2007)	Raster	Variability in monthly rainfall concentration	Modal zonal statistic of rainfall concentration index (%) per grid cell
	Rainfall variability	1.7 km	Schulze & Maharaj (2007)	Raster	Variability in mean annual precipitation	Modal zonal statistic of the coefficient of variation (%) of precipitation per grid cell
	Diurnal temperature variability	1.7 km	Schulze & Maharaj (2007)	Raster	Variability in daily temperature	Modal zonal statistic of the mean monthly temperature range (T_{max}-T_{min} ^{o}C) per grid cell
Topographic (***Ti***)	Terrain ruggedness index	90 m	Created for this study using SRTM 3 arc-second DEM	Raster	Topographical variability	Modal zonal statistic of terrain ruggedness values per grid cell
	Terrain surface classification	90 m	Created for this study using SRTM 3 arc-second DEM	Raster	Relief richness	Modal zonal statistic of terrain surface classification values per grid cell
Geomorphometric (***Gi***)	Geomorphon	90 m	Created for this study using SRTM 3 arc-second DEM	Vector	Landform richness	Count of different Geomorphon features per grid cell
Solar morphometric (***Si***)	Solar radiation variability	1.7 km	Schulze & Chapman (2007)	Raster	Variability in daily solar insolation	Modal zonal statistic of mean daily solar radiation (MJ/m^2/day) per grid cell

4.2.2.2 Geodiversity partial index classification

After using the vector (polygon) grid design, an overlay function was used to calculate seven theoretical numerical partial indices based on the following geodiversity elements: hydrographic, geological, pedological, geomorphological, climatic, atmospheric and topographic diversity within the UTDM region. Broadly, a bilateral approach was implemented to separately quantify each sub-partial index on the format of each geospatial data layer, i.e., raster or vector format. This step is mandatory considering the differences in quantifying phenomena of either a discrete or continuous nature. All datasets analysed in the diversity quantification were re-projected to Transverse Mercator Projection with a 31° meridian and WGS 84 Datum. For the vector-based sub-partial datasets, the quantification is based on the variety (richness) and count (abundance) (Ruban, 2010) of the different elements representing geodiversity (Serrano *et al.*, 2009; Araujo & Pereira, 2018). The primary assumption is that a higher occurrence or variety of features per grid cell will represent a corresponding degree of geodiversity. Importantly, I assessed the geodiversity classes directly by computing the primary indicators of geodiversity, such as geological classes and landforms and using surrogate factors that are considered reflections of geodiversity, such as topographical or climatic variability (Pellitero *et al.*, 2015). The utility of the grid quantification approach is to express, in the most harmonious way possible, all of these aspects without overrating any particular sub-partial element (Araujo & Pereira, 2018). For the analysis of geodiversity, it was decided that only the *count* of relative frequency and the *variety* of each element within the sample area and avoided qualitative ranking, or weighting, based on the intrinsic, or perceived, importance of each landscape component. In doing so, I minimised the subjectivity of the final geodiversity output, a common caveat with many other evaluation methods (Silva *et al.*, 2013).

Furthermore, this provides a segway for the final geodiversity map to provide an objective and unprejudiced regional interpretation of diversity within the UTDM. For all contributing raster-format datasets, the challenge is defining a central measure of quantification to represent each sample unit, i.e., grid cell. Therefore, a *Zonal Statistic* geoprocessing approach was implemented in ArcMap Desktop to determine the most commonly occurring (*mode*) aggregated value for each grid cell. This data-generalisation approach provides a simplistic yet still regionally representative solution for aligning the various multi-resolution raster datasets to the unified geodiversity sample grid area. The mode statistic's selection to represent the *Zonal Statistic* is purely a pre-processing computation choice to ensure that values are represented as absolute integer values rather than decimals, allowing a more elegant calculation of the resulting sub-partial index values.

While all elements are considered equal shareholders in the final diversity assessment, each sub-partial index constitutes a distinct set of variables (**Table 4.1**): *hydrographic index* (two sub-partial covariates), *topographic index* (two sub-partial covariates), *lithological index* (two sub-partial covariates), *pedological index* (two sub-partial covariates), *geomorphological* aspects (one sub-partial covariates variables), *climatic index* (two sub-partial covariates) and *solar morphometric index* (one sub-partial covariate). Therefore, each partial index combines either one, or several geospatial layers normalised, using a simple normalisation formula (**Equation 4.1**), to a scale of 0-0.5 or 0 and 1.0 where single covariates are analysed.

$$X_{norm} = \frac{(X_i - X_{min})}{(X_{max} - X_{min})} \quad \textbf{(4.1)}$$

A brief description of each of the seven partial indices, and their contributed sub-partial indices, including the geoprocessing steps, are described in detail below and graphically represented in **Figure 4.2 and Figure 4.3**:

- The *hydrographic index (**Hi**)* is evaluated by considering both surface and subterranean hydrological contributions to geodiversity. For the sub-partial index diversity calculation of surface waters, I used the 1:500 000 National vector dataset (DWAF, 2020) of the perennial stream network. I augmented this coverage with a higher resolution 1:50 000 National vector river network (NGI, 2018) to include non-perennial streams into quantifying drainage density (km/km^2) for each grid cell. Drainage density is useful for describing drainage basin morphometry by estimating the stream channel's length per unit area, indicating how well or how poorly a watershed is drained by stream channels (Horton, 1932). Drainage density for the UTDM is represented per grid cell area (6.25 km^2), and then, using **Equation 4.1.**, all values are normalised to a scale of 0 and 0.5. A score of 0 is assigned to squares, in which no hydrological elements are represented, while a score of 0.5 indicates high drainage density. To complement drainage density, *mean* annual baseflow (mm), i.e., the dry weather and non-rainy season streamflow which are sourced from the groundwater stores, are also assessed. The baseflow coverage has been empirically generated for South Africa on a 1` x 1` (1.7 x 1.7 km) spatial resolution using 50 years of historical climate data. I resampled the baseflow coverage to 90 m resolution using a bilinear interpolation approach in ArcMap to allow interoperability with the other datasets, specifically the topographic, derived from the 3-arc second Shuttle Radar Transfer Mission (SRTM) DEM. The ArcMap *Zonal* S*tatistic* function was then used to calculate the modal raster baseflow within each 6.25 km^2 grid cell. Like drainage density, baseflow values were normalised to 0 and 0.5. The final ***Hi*** partial index value is calculated as the sum of the partial drainage density diversity and baseflow diversity with values ranging from 0 to1.
- The *lithological index* (***Li***) is calculated as the sum of the lithostratigraphic and paleo-sensitivity sub-partial indices; both derived using a digitised 1:250 000 geological series map courtesy of the South African Council for Geoscience (CGS, 2000) with minor modifications. The lithostratigraphic map is a special-purpose vector polygon dataset depicting the 12 dominant lithostratigraphic units and geologic strata in the UTDM. Following the method outlined by Araujo & Pereira (2018), the lithological units are first homogenised to eliminate duplication of polygons of the same classification within each cell. This is done using a unique numerical ID assigned to each lithological unit and then 'exploding' the dataset before intersecting the layer with the spatial grid, with each grid cell itself having a unique identifier, finally 'dissolving' on unique grid ID and summarising on *Count* to calculate the sum of differentiated lithological features within each cell. The paleo-sensitivity and geological heritage (Lavin, 2013) sub-partial index provides a valued assessment for tentatively outlining paleontologically sensitive areas in the UTDM and adds a necessary quantification to the geological sub-index calculation and thus final *GDIx*. The paleo-sensitivity sub-partial dataset is based on the same 12 lithological classes as above but ranked according to their paleo-sensitivity using a six-class classification (unknown to very high). The paleo-sensitivity diversity is computed by counting the number of occurrences of different sensitivity classes within each grid cell, following the same procedures used to calculate the lithological heterogeneity. While this geoprocessing analysis is wholly automated in a GIS, a crucial step is the manual inclusion of same-ID and non-contiguous features as part of the feature count. This approach is considered a "goldilocks" approach to avoid over or underestimating feature count per grid cell. Since the ***Li*** comprises two sub-partial vector layers, each layer is normalised to 0 and 0.5 and then summed to a scale of 0 to 1.0 for the final lithological index. **Figure 4.2** outlines the geospatial approach used to derive the sub-partial indices for all vector datasets. Simultaneously, **Figure 4.2d** represents the optimised estimation based on the goldilocks concept for each grid cell's feature count and variety.

- For the *pedological index (**Pi**),* two sub-partial indices: soil pattern and soil depth > 800 mm, were analysed using the 1:250 000 Land Type Survey (LTS) digital soil inventory dataset (ARC, 2003). The land types of South Africa have been comprehensively mapped over South Africa for agricultural purposes. Each land type shows an accepted degree of uniformity concerning terrain form, soil pattern and climate. In total, the LTS contains information for twenty-eight broad soil pattern groups (association) classified according to the Soil Classification: A Binomial System for South Africa (MacVicar *et al.*, 1977). In the LTS, soil data are provided as inventories, of which there are some 7 200 comprising about 65 000 soil components. Supporting memoirs for 2 500 modal profiles, for which comprehensive soil analysis data are available for the various horizons, including a description of overall soil profile depth (Turner, 2007). The soil pattern sub-partial index, therefore, facilitates the tracing of the pedological diversity in the UTDM. It should be noted that the South African Soil classification has seen two subsequent revisions since the Binomial System: first the Soil Classification: A Taxonomic System for South Africa (SCWG, 1991) and the third revised addition Soil Classification: A Natural and Anthropogenic System For South Africa (SCWG, 2018). Neither of these revised classifications has been used to update the LTS memoirs and corresponding polygons. Following the same grid-geoprocessing procedure, as outlined for ***Li**,* the information for 15 soil patterns for the entire UTDM was analysed with the count of soil class per grid normalised to 0 and 0.5. The soil depth sub-partial index was derived by selecting all soils with a depth > 800 mm (given to a depth of 1 500 mm) in the LTS dataset. Soil depth here refers to total profile depth and not effective rooting depth highlighting limitations due to physical rather than chemical impediments (Schulze *et al.*, 2007). This definition, therefore, denotes the depth to which plant roots can be active. Polygons where soil depth is > 800 mm were all equally weighted and assigned the maximum value of 0.5. All remaining polygons were assigned a value of 0 and then summed with soil pattern index scores for the combined final ***Pi*** values ranging from 0 to 1.
- The *climatic index* (***Ci***) assessed the contribution of two sub-partial indices: rainfall and temperature diversity. Rainfall diversity was evaluated by rainfall concentration and rainfall variability, while daily max-min temperature variability was considered a suitable surrogate contributor to geodiversity in the UTDM. Rainfall concentration, instead of MAP, is a good determinant of how concentrated (severe) rainfall events are throughout the year, which provides a complete perspective of the influence of events on the landscape. Schulze & Maharaj (2007) represent rainfall concentration, calculated per quaternary catchment, for South Africa on a 1` x 1` (1.7 x 1.7 km) spatial resolution using 50 years of historical data. The raster coverage delineates high rainfall regions as locations that receive all their rainfall in one month. In contrast, low rainfall concentration regions receive the same rain for each month of the year. Therefore, the higher the concentration index, the less spread the rainfall season is over time (whether it is a high or low rainfall area or in a winter or summer rainfall region). To complement the effects of rainfall concentration, I likewise quantified the coefficient of variation (CV) of annual precipitation for the UTDM. MAP maps do not showcase the natural year-to-year rainfall variability within a region. The CV (%) is a more suited index to express precipitation with high/low CV values corresponding to high and low inter-annual rainfall variability, respectively.

  The sub-partial temperature diversity is represented by the regional diurnal temperature range (Schulze & Maharaj, 2007). Very simply, the diurnal temperature range is the difference between the maximum and minimum temperatures on a given day. Unlike MAT, which is a reasonable indicator for continentality, the selection of mean monthly diurnal temperature range is a convenient index to assess the influence of temperature diversity related to a broad range of local anthropocentric and natural process phenomena. These include, *inter alia,* land cover/ use change, climate change, thermo-periodism and photosynthesis, i.e., the response of plants to diurnal temperature ranges (Schulze & Maharaj, 2007). As with ***Hi***, both ***Ci*** sub-partial indices are resampled to 90 m and using the *Zonal Statistics* to calculate

the modal rainfall concentration per grid cell. The rainfall concentration and rainfall CV are first normalised to a scale of 0 and 0.25 and then totalled to an index of 0 to 0.5. The geoprocessing of the diurnal temperature range follows suit with *Zonal Statistics* values rescaled to 0 and 0.5. Both rainfall diversity and temperature diversity sub-partial indices are tallied resulting in a final ***Ci*** diversity score between 0 and 1.

- *Topographic index (**Ti**)* is calculated as a collective of the terrain ruggedness index (TRI) and terrain surface texture (TST) sub-partial indices with both datasets derived from two mosaiced and void-filled 3-arc second SRTM DEM V4.1 (Jarvis *et al.*, 2008) scenes with a resolution of 90 m for the entire UTDM region. The simplest definition of TST (Bennett & Mattsson, 1999) is that it represents the amplitude and frequency of topographic relief features of a surface and provides one of many useful surrogates for assessing topographic diversity (Benito-Calvo *et al.*, 2009). Consequently, large roughness diversity values indicate a substantial dispersion of slope gradient and slope aspect with short slope lengths. In contrast, small values indicate a homogenous orientation of slopes with minor variations and long slope lengths (Hjort & Luoto, 2010). At the same time, Riley *et al.* (1999) consider TRI an expression of terrain heterogeneity, quantitatively indicating how undulated, broken, rugged or dissected the landscape is between a grid cell and its eight neighbouring cells and an ideal indicator for assessing land surface ruggedness. The inclusion of TRI and TST is supported by the decisive role of both parameters on the energy flows (exposure to sunlight, humidity) and material flows (water, sediments on the slopes), and the diversity and distribution of landforms and processes (Serrano & Ruiz-Flaño, 2007). These two sub-partial indices collectively provide a far more pragmatic contribution of topographic variability to geodiversity than only evaluating solitary altitudinal and declivity gradients. To quantify the resulting ***Ti***, the *Zonal Statistics* was calculated for each sample grid resolution as outlined before. Using **Equation 4.1**, I rescaled the modal values for TRI and TST to 0 and 0.5 respectively. Both sub-partial indices are then summed representing the final ***Ti*** diversity between 0 and 1 per grid cell.
- For the *geomorphometric index (**Gi**),* the utility of the geomorphon (geomorphological phonotypes) mapping approach was leveraged to provide a general-purpose geomorphometric map for the UTDM using the previously processed 90 m resolution SRTM DEM. Moving beyond simple topographic heterogeneity measures, the geomorphon approach's utility (Jasiewicz & Stepinski, 2013) is the low-cost characterisation of landform features, which I consider a reasonable standard of landform richness. Geomorphons are analogous to textons (Julesz, 1981) of a landscape. Their extraction from a DEM comes at a small computational cost considering that they simultaneously represent quantitative and stratified terrain attributes and landform types. The geomorphon approach's product is the stratification of the landscape into ten unique but recognisable landform elements: *peak, ridge, shoulder, spur, and slope, hollow, foot slope, valley, pit* and *flat* terrain morphological units. Recently, Atkinson *et al.* (2020) successfully applied the geomorphon approach to soil-landscape characterisation to a smaller region within the UTDM.

For this reason, I decided to implement the scale-independent approach to a larger regional area such as the UTDM. Similar to Bailey *et al.* (2017) and Atkinson *et al.* (2020), I used the geomorphometric algorithm '*r.geomorphon'* extension in GRASS GIS 7.4.1 (GRASS Development Team, 2016) to delineate the landform coverage data from the SRTM DEM using the pattern recognition approach. To optimise the geomorphon surface, I adopted a similar approach to Luo & Liu (2018) and Gruber *et al.* (2015) by iteratively selecting model parameters that offered a reasonable balance between the accuracy of capturing landform elements and model computational cost. The final model parameters were defined by a 30-cell outer search radius and a six-cell inner search radius with a flatness threshold of 1.2°.

A four-cell majority filter in ArcGIS *Spatial Analyst* was applied to the resultant 10-class raster geomorphon surface to eliminate non-contiguous "single-pixel" features for a more generalised or "smooth" landscape representation. Geomorphon features were then vectorised to polygons and then further homogenised to eliminate duplication of the same feature with multiple polygons of the same geomorphon unit 'dissolved' to eliminate redundant features from the diversity quantification. The quantification of the geomorphon sub-partial index followed the same methodology as previously outlined for vector datasets (**Figure 4.2**), namely that of soils and geological material, except for the single sub-partial ***Gi*** values normalised directly to 0 and 1.

- Finally, I evaluated *solar morphometric diversity (**Si**)* through the contribution of monthly mean solar radiation variability (MJ/m^2/day), using the solar radiation raster grid developed by Schulze & Chapman (2007) at 1` x 1` (1.7 x 1.7 km) resolution using 50 years of historical data for South Africa. Typically, the surface (slope) aspect is used to determine incident solar energy (insolation) reaching the earth at a particular location, especially when considered with slope gradient (Wilson & Gallant, 2000). Insolation influences critical local climatic and morphographic trends, including near-surface temperatures, evaporative demand and soil moisture content (Prasannakumar *et al.*, 2011). Knowing the amount of radiation at different geographic locations at different times contributes to the climate, biosphere, and external geodynamics (i.e., erosion, transportation, and sedimentation of materials) within the landscape (Petrasova *et al.*, 2015). Despite the significance of the slope aspect to the soil-plant-atmosphere continuum, being a circular variable, there are numerical limitations associated with its calculation and broader application. Solar radiation, the exogenous component of the surface thermal regime, is a suitable morphometric replacement for slope aspect with added computational benefits. First, the estimation of solar radiation is more sophisticated. It considers inputs such as latitude, atmospheric properties, and temporal frequency, which accounts for differences in topographic shadowing (from adjacent hills) and slope gradient (Thompson *et al.*, 2012). Second, Van Niel *et al.* (2004) demonstrated that solar radiation derived directly from a DEM is less affected by DEM error than aspect and slope. Finally, unlike the slope aspect, solar radiation considers the measurable quantity of direct sunlight rather than merely the cardinal representation of solar insolation. Schulze & Chapman (2007) record solar radiation as monthly means of daily solar radiation variability. To derive a yearly average, I used the ArcMap *Raster Calculator* (Spatial Analyst Tools) to cascade the twelve-monthly datasets and calculate the annual average for the UTDM region. The solar radiation sub-partial index then followed the same geoprocessing protocol applied to the previous raster datasets: first generalising the solar radiation dataset from 1 700 m to 90 m and then summarising the sample-grid values by *Zonal Statistics* to calculate the modal radiation variability per 2 500 m grid cell. the solar radiation sub-partial index was then normalised directly to 0 and 1.0 with grid cells approaching 1.0 displaying higher ***Si*** diversity.

The final *GDIx* score for each grid square is the sum of the seven sub-partial diversity indices (**Equation 4.3**). Following Araujo & Pereira (2018) I reclassified the resultant seven-class *GDIx* grid values to an equal interval five-class model using, **Equation 4.2**, to distinguish the following *GDIx* classes: Class 1 (0 to 1) = very low; Class 2 (1 to 2) = low; Class 3 (2 to 3) = medium; Class 4 (3 to 4) = high; Class 5 (4 to 5) = very high. These values are based on the minimum, and maximum values obtained (Betard & Peulvast, 2019) (**Figure 4.4h**).

$$Y_{new} = \left(\frac{(X_i - X_{min})}{(X_{max} - X_{min})}\right) n_{upper\ limit} \quad \textbf{(4.2)}$$

$$G_{diversity} = H_i + P_i + L_i + C_i + T_i + G_i + S_i \quad \textbf{(4.3)}$$

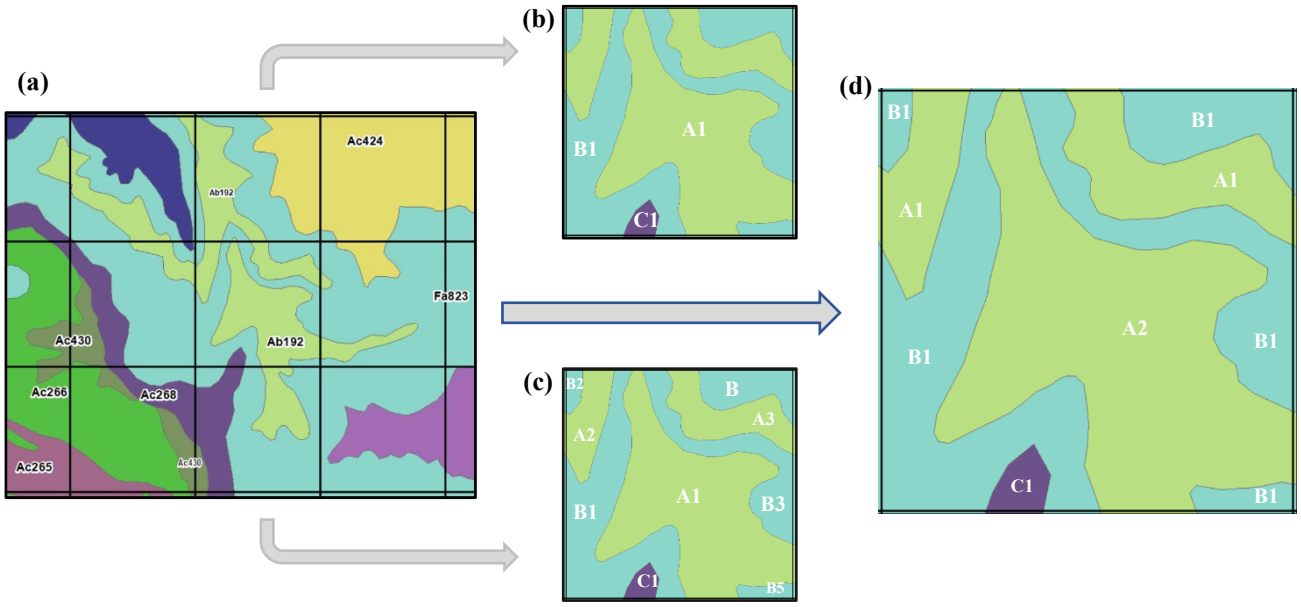


Figure 4.2: Concept of the geodiversity quantification of vector datasets in the 2.5 x 2.5 km grid. a) overview map showing vector soil association(s) in a 3 x 3 window range. b) Counting the number of soil forms per grid cell. Approach too conservative with only three soil form features counted and full abundance of soil features not quantified c) Counting the number of vector features per grid cell. Approach too liberal with a total of 9 soil form features counted. Duplication of similar features resulting in an overestimation of geodiversity per grid cell. d) counting the number of vector features per network cell per spatially contiguous element.: *"Goldilocks Case"* of quantifying vector feature diversity resulting in 4 soil form features per grid.

.

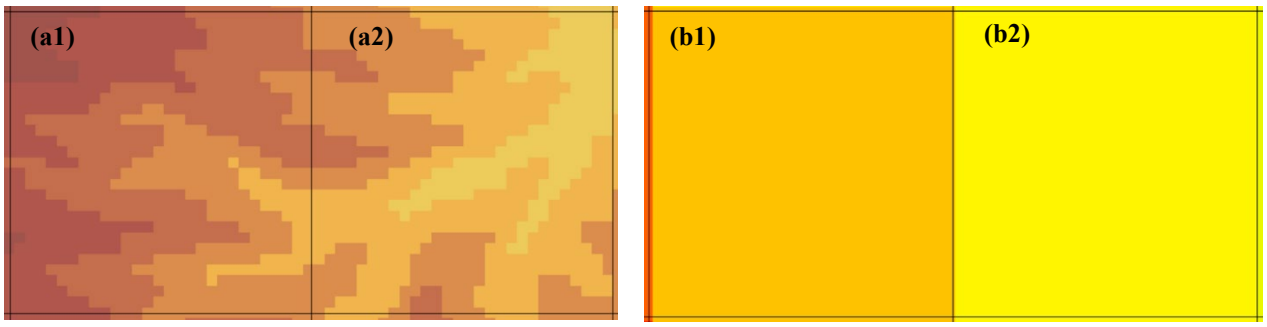


Figure 4.3: Concept of the geodiversity quantification of raster datasets in the 2.5 x 2.5 km grid. a) SRTM 90 m elevation for two grid cells at with a1) having an elevation height range of 413 m and a2) having an elevation height range of 345 m. Grids highlighted in b1) and b2) have been generalised using the Zonal Statistical modal value for elevation calculated for each grid cell resulting in new elevation grid values of 375 m and 305 m, respectively, for the Geodiversity Index calculation.

4.2.2.3 Geodiversity mapping

While the final grid map may be readily interpreted as a regional representation of geodiversity, the tessellated grid results do not offer further opportunity to interrogate the geodiversity-landscape relationships within the UTDM. To better visualise the *GDIx* as a smooth continuous surface, I followed a similar protocol to Argyriou *et al.* (2016) and Araujo & Pereira (2018) with modifications as follows:

1. The original 2 500 m grid was resampled to 500 m, generating a centroid for each new grid cell. This resulted in 25 centroids at 500 m compared to 1 centroid at 2 500 m grid resolution.
2. Then the centroid points were used in an optimised ordinary kriging interpolation method (with standard model parametrisation) in ArcMap *Geostatistical Wizard*, using a maximum of five and a minimum of two neighbours within a four-sector range to generate intermediate values between the geodiversity centroid values.
3. The final output is a more detailed spatially continuous raster surface of geodiversity distribution for the UTDM (**Figure 4.4**).
4. To further capitalise on the spatially detailed interpolated geodiversity surface, I generated a histogram to highlight the distribution, range and class intervals of the *GDIx* values for the UTDM.

4.2.3 Geodiversity of landscape resources

The evaluation of landscape inter-connectedness is an integral part of environmental planning, management and decision-making. Furthermore, regional landscape assessment for planning, conservation, and decision-making is a central component to sustainable land management success (Mander *et al.*, 2005). Landscape metrics (indicators) have become increasingly popular in determining landscape characteristics, both structurally and functionally. The diversity of landscape indicator models for evaluating performance functions of landscapes has received due consideration, e.g. (McGarigal *et al.*, 2009; Malinowska & Szumacher, 2013; Uuemaa *et al.*, 2013). Dale & Kline (2013) deliberately reiterate that when landscape indicators are combined with land productivity measures, biodiversity, greenhouse gas emissions, soil, water, and air quality, they can contribute to sustainable land management.

Furthermore, landscape indicators help address questions about land availability and capacity for providing ecosystem services and urban and industrial development (MEA, 2005). The juxtaposition of abiotic partial-indices, i.e., ***Hi, Pi, Li, Ci, Ti, Gi*** and ***Si,*** used to quantify geodiversity in this study provides reliable evidence for applying the *GDIx* as one among many landscape indicators for sustainable land management in the UTDM region. As an original contribution, I conducted an exploratory analysis of *GDIx* and land quality (including selected components of biodiversity) with the aim of mapping the association between high and low geodiversity regions with several prevailing land resource information datasets considered functional representatives for land use, environmental and conservation planning and management (**Table 4.2**). The exploratory analysis does not extend to evaluating geodiversity and any specific inherent or dynamic land resource properties or ecosystem services. The investigation's primary focus is to highlight the capability of the *GDIx* as a resource to support policy development and decision-making that covers a wide range of issues relating to productivity and environmental outcomes. Quantifying the overall association between different land units and *GDIx* by evaluating the associated and derived surfaces' descriptive data is a critical element of

the applied methodology. For this reason, I include the histogram distributions for each land resource concerning *GDIx* values (very low to very high) (**Figure 4.5**).

4.2.4 Statistical analysis

For this study, some measures of association between the geodiversity and land resource datasets were investigated. Measures of association provide a means of summarising the size of the association between two variables. An added benefit of these non-parametric tests includes an estimate of significance for association measures (Gingrich, 1992). One such test, a chi-square independence analysis, can provide a rapid exploratory and descriptive measure of spatial association between several spatial datasets (Bonham-Carter, 2014). This is achieved by cross-tabulation to obtain the expected number of cases under the assumption of no relationship between the two variables. The chi-square statistic value tests whether there is a statistical relationship between the cross-classification table variables (Gingrich, 1992).

The statistical program SPSS Inc. (v25) was used to perform the chi-square analysis to derive two measures of association for testing the statistical association between geodiversity and the following geospatial layers: *National Land Cover (NLCov)*, *National Land Capability (NLCap)*, *Provincial Bio-Resource Units (BRU)*, *National Biodiversity Conservation Status (NCS)*, *National Terrestrial Biomes (NTB)* and *Biome Protection Status (BPS*). The first of these measures is the Cramer's *V* coefficient (**Equation 4.4**). Cramer's *V* designates the statistical association among the compared categories by calculating the variance between the variables. Values of Cramer's *V* lie between 0 and 1. The closer V's value is to 1, the higher the association between the variables (Ott *et al.*, 1992). Also, a chi-square value (**Equation 4.5**) was determined to test Cramer's V's significance.

$$V = \sqrt{\frac{x^2/N_{Obs}}{min_{(i-1)(j-1)}}} \quad \textbf{(4.4)}$$

$$X^2 = \sum_{i=1}^{K} \frac{(Observed_i - Expected_i)^2}{Expected_i} \quad \textbf{(4.5)}$$

Table 4.2: Landscape resources and their relationship to Geodiversity

Landscape resources	Attribute description	Resolution/ Scale	Source	Geodiversity relevance to ecosystem service types
Bio-Resource Unit	Defined as an ecological unit within which factors such as soil type, climate altitude, terrain form and vegetation display enough homogeneity. Appropriate land use practices and production techniques can be defined for each unit.	1:50 000	(Camp, 1995), (Schulze *et al.*, 2007)	Provisioning & Supporting
Land Capability	Land capability is the most intensive long-term use of land for purposes of rainfed farming determined by the interaction of climate, soil and terrain.	1:250 000	(EnviroGIS, 2016a)	Provisioning & Supporting
Land Cover	South Africa National land-cover generated using multi-seasonal 20m resolution Sentinel 2 satellite imagery. The imagery used represents the entire temporal range of available imagery acquired by Sentinel 2 from 01 January 2018 to 31 December 2018. Over 30 land cover classes have been mapped, including wetlands, forests, mines, agriculture, built up, erosion, and vegetation.	1:250 000	Geoterra Image, (GTI, 2019)	Provisioning, Supporting, Regulating & Cultural
Vegetation Ecosystem Protection Level	The ecosystem protection level dataset outlines whether ecosystems are adequately protected or under-protected. Ecosystem types are categorised as not protected, poorly protected, moderately protected or well protected, based on the proportion of each ecosystem type that occurs within a protected area recognised in the Protected Areas Act. In addition to protection status, conservation status and vegetation biome were also evaluated.	1:250 000	South African National Biodiversity Institute (SANBI, 2012)	Supporting & Regulating

4.3 RESULTS AND DISCUSSION

4.3.1 UTDM Sub-partial diversity quantification

The final geodiversity map resulting from the interaction of the seven partial diversity index datasets evaluated for the entire UTDM is presented in **Figure 4.4**. The methodology used for quantifying the partial index diversity for both vector and raster datasets proved to effectively highlight the richness and abundance of features for the respective partial sub-indices. In particular, the topographical, pedological and geological diversity estimation and the geomorphometric diversity quantification all showed strong associations with regions in the northern, southern and western areas of the UTDM. Incidentally, these are characterised by the Drakensberg mountain range, higher altitude, complex terrain and associated increased diversity of soil and geological material. Consequently, these partial-index covariate datasets strongly influence the final *GDIx* map with medium partial-index diversity areas randomly distributed throughout the UTDM. Notably, despite using a moderate-resolution spatial layer of soil association, the pedological partial-index displayed a surprisingly high pedological diversity. The eastern and western regions of the UTDM are known to be pedologically diverse, predominantly due to topographic and climatic influence.

These findings are aligned with similar findings by Kori *et al.* (2019), who found that their final geomorphological diversity map was strongly influenced by geology, slope and soils in the Soutpansberg range in South Africa. This would suggest that despite differences in the quantification approaches to geodiversity estimation, i.e., factor maps (Kori *et al.*, 2019) or partial-index maps, as presented in this study, the influence of selected covariates for driving geodiversity remain relatively systematic regardless of regionality. These results present inspiring opportunities for further testing the quantification approach to geodiversity mapping in other country regions to further validate the influence of physical environmental covariates. Identifying landscape units is an integral part of understanding a given portion of land's behaviour, functioning, and dynamics and how these interrelationships drive geodiversity in the landscape. Specifically, this study has shown that the topographical (terrain character), pedological (soil character) and lithostratigraphic (geological and paleo-sensitivity character) diversity, as well as the geomorphometric (landform character) diversity quantification, are all highly invested in the final *GDIx* valuation, having a distinctive spatial clustering in the northern, southern and western areas of the UTDM. Moreover, the partial-index quantification represented the catenary diversity of *Cambisols*, *Leptosols* and *Planosols* in the most dominant topographic regions nearest the Drakensberg. Simultaneously, sandier soil associations related to *Ferrasols* and *Arenosols* dominated in the central area of the UTDM. Finally, *Plinthosols* rich in iron and manganese oxide formed under hydromorphic conditions governing the lower *foot slope* and *valley bottom* morphogenic parts of the UTDM.

The climatic and solar morphometric partial indices exhibited a categorical clustering of high diversity regions closest to the Drakensberg with a gradual decrease in diversity clustering in an easterly direction towards the lower altitude homogenous coastal region. These trends are entirely expected considering the topographic and altitudinal dominance of the Drakensberg in the western part of the UTDM. Indeed, the influence of the Drakensberg mountain range on the micro/ meso-climatic variability in the western region would be more pronounced than the lower altitudinal and more uniform undulating eastern coastal terrain region. The pockets of high solar and climatic diversity in the central and eastern areas of UTDM may be attributed to the known seasonal climatic variability associated with the humid coastal region and the adjacent interior Mist belt, which forms a band through the entire KwaZulu-Natal Midlands, and consequently a central region of the UTDM (Zaloumis, 2013). To further highlight the climatic variability, the District has an approximate annual

minimum and maximum precipitation of 255 mm and 1 558 mm, respectively, while yearly mean minimum and maximum temperatures are in the region of -15ºC and 42ºC, with temperatures and rainfall generally declining from coastal to inland areas. A recent study by Lakhraj-Govender & Grab (2019), evaluating decadal climatic temperature data for KwaZulu-Natal for the past 85 years, has shown evidence of oscillatory interior warming and coastal cooling by magnitudes of between 0.04 ºC and 0.06 ºC. Their findings may support the present study's high climatic "gradient" diversity detected in the central UTDM region.

Surprisingly, the hydrographic partial-index layer did not perform as expected, considering the high stream density calculated for the region. An average stream density of 2.3 km per grid (6.25 km^2), with a maximum density of 7.46 km, was recorded for the UTDM with the highest values conditioned by the Drakensberg catchment basin(s) and fluvial drainages within closed valleys in the central and eastern regions of the study site. One possible cause for these lower than anticipated ***Hi*** results may simply be the quantification approach of combining the medium-resolution river network coverage with the low-resolution groundwater coverage for the UTDM. The ***Hi*** for the UTDM was evaluated based on the method proposed by Pereira *et al.* (2013) and followed a similar approach to Betard & Peulvast (2019) in considering both surface and underground waters for the hydro diversity partial diversity calculation. Similar to Betard & Peulvast (2019) findings, surface waters' contribution dominates the diversity estimation in areas of high stream density and baseflow and show less influence in low baseflow regions. Likewise, Manosso & de Nóbrega (2016) highlight that, indirectly, sets of elements with the highest number of variables significantly influence the overall quantification index process. This is an unavoidable caveat of the numerical method of counting feature richness and abundance using a grid-based approach.

By default, features with the most occurrence (count) will register as more diverse while other elements become reflections of conditions contributing to the overall diversity. Panizza (2009) argues that there is no consensus that numerical calculations of indices offer enhanced representation(s) of geodiversity. Erikstad (2013) suggests that numerical methods must balance the assessment's subjective and objective elements. The major drawback of numerical approaches is the lack of distinction between the description and the evaluation of nature. The valuation of nature must be defined in its own right. Likewise, indicators of diversity may be of great interest as value criteria. Additional value criteria should be included if these indicators are designed to measure a part of the total variety. For quantitative approaches, it is particularly problematic when the absence/ presence of a limited set of value indicator(s) are exclusively used to represent the lack/ abundance of geodiversity value.

The results in **Table 4.3** highlight how the number of sub-partial indices for each partial-index calculation significantly influences the overall *GDIx* calculation. For example, the estimation of ***Hi,*** which comprises two sub-partial indices, requires the normalisation of each sub-index to 0 and 0.5. Then each sub-index score is further combined to form the final hydrographic partial index score between 0 and 1.0. However, since the scores are not weighted, this results in the more abundant stream density areas having a more considerable influence on the maximum sub-partial index of 0.5, including an overall ***Hi*** diversity of 0.5 in regions where the baseflow is low or absent. The same principle would apply to areas of maximum baseflow and low stream density. Furthermore, this principle would apply to all partial indices comprised of two or more sub-index datasets. The lower diversity estimations would be expected as more sub-partial indices are used to define each partial-index and minimal coincident features per grid cell. Therefore, an optimal result would be expected where sub-indices are not combined but instead comprise a single partial index. **Table 4.3** further highlights that each sub-partial index's contribution to the maximum expected normalised partial diversity score would decrease by a factor equivalent to the maximum normalisation threshold for each sub-index. For instance, in the case of a partial index comprised of four sub-partial indices where only two sub-partial indexes contribute

to the grid's diversity, the overall representation of diversity will be 50%, or 25% if three sub-partial indices are absent.

A pragmatic solution for future geodiversity estimations using a gridded approach may need to avoid using sub-partial indices altogether to derive partial indexes of geodiversity and instead use individual datasets, i.e., soil form, geology, baseflow, and stream density, independently as partial-index layers. This may resolve the shortcomings of the obligatory normalisation approach between datasets of varying scales (spatial and measurement) whilst maximising the representation of diversity features at a partial-index level for overall geodiversity representation. Similar methods were adopted by Hjort & Luoto (2010), who evaluated 74 different types of elements of geodiversity in Northern Finland and Manosso & de Nóbrega (2016), who considered 15 features of geodiversity to calculate the geodiversity from landscape units of the Cadeado Region in Paraná, Brazil. Alternatively, future analyses may further benefit from introducing a weight factor for each sub-partial index, particularly for those sub-index features used for ***Hi***, where there is an apparent spatial bias of feature richness and abundance that influences feature diversity. The introduction of an attribute weighting factor may also address the "edge-effect" limitations expected from coarse resolution datasets that do not necessarily show much diversity over small distances within the landscape, i.e., Koppen Geiger climate classification (Conradie, 2013). In these instances, the diversity quantification is limited to the adjoining climatic-class boundaries with a minimal intra-feature variation. However, as Santos *et al.* (2017) point out, introducing a weighting factor for each element may be challenging without introducing subjectivity into the diversity analysis leading to an imbalance in the final results. Qualitative assessment methods based on knowledge and expert experiences such as descriptive documentary (Panizza, 2009), expert classification (Kozlowski, 2004) and values and benefits (Gordon & Barron, 2013) have successfully been used for systematic geodiversity evaluation, albeit with limited applicability and repeatability beyond the spatial regions of interest. The solution to these limitations is to ensure maximum data representation and a more detailed spatial variation of parameters at minimal expense to data quality.

Regarding estimating climatic diversity in this study, rather than introducing a "weighting's factor to the methodology, I admittedly applied a rudimentary but effective downscaling approach, suggested by (Vajda & Venäläinen, 2003; Hjort *et al.*, 2012). This led to resampling the 1.7 x 1.7 km resolution climatic sub-partial indices to a 90 x 90 m resolution using a bilinear interpolation approach with an overall R^2 prediction of 96%. Rescaling the sub-partial climatic index to a more acceptable resolution allowed improved climatic diversity representation with higher-density natural class breaks (Jenks) for each sub-partial climate variable.

Table 4.3: Contribution of increasing sub-partial indices to overall partial index diversity quantification (Partial-decay). Min-max defines the upper and lower limit of normalised scores for each sub-partial index. The min-max detection of features, combined with the normalisation to a scale of 0 to 1 within each grid cell, directly influences the final partial index of feature diversity.

Sub-partial index with 1 diversity element		Sub-partial index 2 diversity elements		Sub-partial index 3 diversity elements		Sub-partial index 4 diversity elements		Partial index Normalised Score	Expected Partial index Diversity
Max								Normalised: 0-1.0	Max
Min									Min
Max	Normalise: 0.5	Max	Normalise: 0.5						Max
Min		Max							50%
Max		Min							50%
Min		Min							Min
Max	Normalise: 0.33	Max	Normalise: 0.33	Max	Normalise: 0.34				Max
Min		Min		Max					33%
Max		Min		Max					66%
Min		Max		Min					33%
Min		Min		Min					Min
Min		Max		Max					66%
Max	Normalise: 0.25	Max	Normalise: 0.25	Max	Normalise: 0.25	Max	Normalise: 0.25		Max
Min		Max		Max		Max			75%
Max		Max		Max		Min			75%
Max		Min		Max		Max			75%
Max		Max		min		Max			75%
Min		Min		Min		Min			Min
Max		Max		Min		Min			50%
Min		Min		Max		Max			50%
Min		max		Min		Max			50%
max		Min		Min		max			50%
Max		Min		Min		Min			25%
Min		Max		Min		Min			25%
Min		Min		Max		Min			25%
Min		Min		Min		Max			25%

4.3.2 Geodiversity index map for UTDM

Despite particular methodological challenges, which represent similar issues of the current discourse on geodiversity quantification (Santos *et al.*, 2017), the final geodiversity map presents an exciting product that highlights the integrated analysis of the physical environment and the richness in abiotic elements distributed throughout the UTDM area. The components of geodiversity were analysed using two sets of informative diversity indices: feature *richness* representing the number of fundamental elements within the feature space or sample grid, and feature *abundance* representing the frequency of sub-partial diversity elements within each grid cell (Ferrer-Valero, 2018). The *GDIx* score of each grid square is the sum of all the previously outlined partial indices: ***Hi, Pi, Li, Ci, Ti, Gi*** and ***Si,*** with each index representing an element of geodiversity with values of the final *GDIx* in the range of 1 and 5 (**Figure 4.4**) based on the minimum and maximum values categorised into 5-class equal intervals.

The areas with the highest geodiversity concentration are in the eastern, central and western parts of the UTDM, with smaller pockets of high diversity scattered throughout the region. In the west region, the Drakensberg mountain range stands out as the dominant feature of ***Li***, ***Gi***, ***Ti***, ***Pi*** and surprisingly ***Si*** diversity, where there are diverse landforms and many exposed lithostratigraphic features. More precisely, the highest values of the topographic, geomorphometric, pedological and solar morphometric partial indices (3 to 5 points) are associated with the ensuing morphological contrast between the central region and the diversified relief of the Drakensberg mountain range in the west.

Similarly, the diverse lithostratigraphic and paleo-sensitivity variety justifies the high geodiversity values east of the UTDM. A total of seven of the eleven lithostratigraphic elements were identified in the eastern region of the UTDM. These comprised a mix of alluvial and colluvial material and agglomerates of Masotcheni Basalt and Karoo Dolerite with sedimentary surface intrusions of shale, sandstone, and siltstone and mudstone of the Vryheid, Estcourt, Volksrust formations, respectively. In such areas, the regions of high paleo-sensitivity (fossil heritage) were positively associated with the corresponding lithological diversity signalling that the eastern part of the UTDM is characterised by diverse fossil heritage from geological formations.

Large portions to the north and east as well as south and east regions of the UTDM are characterised by low to medium geodiversity, with the primary diversity influenced by the ***Gi*** partial index of varying geomorphon abundance. Consequently, this same region is marked by a moderate to intermittent high surface water sub-partial index normalised score (0.4 to 0.5) and a low underground sub-partial index normalised score (0.1 to 0.25). This is an apparent cause for the low-medium final ***Hi*** partial index score due to the scale-abundance-quantification paradox of selected sub-partial indices outlined in the previous section. A key observation regarding the *GDIx* is an association of medium-high index values (> 3) for higher altitude regions above 1 450 m a.s.l. With lower index values characterised by lower-lying areas below 1 220 m a.s.l. For these regions, the results suggest that the *GDIx* distribution is closely linked to the topographic arrangement of the UTDM: where the amplitude of terrain relief is high, *GDIx* shows the highest values; in contrast, the lowest *GDIx* values uniformly define the dominant flat areas of the region. Pellitero *et al.* (2015) observed similar trends in glacial landforms and attributed this relationship to the increasingly active landform, geological and soil processes at higher altitudes.

The entire UTDM region of approximately 11 500 km^2 was divided into 2 042 separate 6.25 km^2 grids (**Figure 4.4h**). However, while the grid format may be useful for block diversity quantification, it is not readily applicable to visualise the final *GDIx* results as a continuous surface or interoperable with other spatial datasets for supplementary spatial analysis. Following the methodology outlined by Santos *et al.* (2017), I performed grid-to-point conversion of the *GDIx* values, after which these points were interpolated using an ordinary kriging approach, resulting in a predicted continuous surface (R^2= 0.96) of *GDIx* class distribution throughout the UTDM region (**Figure 4.4i**). The results presented in **Table 4.4** have been derived from the interpolated *GDIx* surface and provide the most meaningful outcomes of key descriptive metrics related to the frequency distribution and selected fragmentation class metrics for each *GDIx* class. The first significant observation is that low-medium *GDIx* areas cover approximately 600 000 ha (or 60%) while high-very high *GDIx* occupies almost 400 000 ha (14%) of the study area. This result suggests that while there are specific features of high diversity - geodiversity hotspots or loci - the region is dominated by low-medium overall geodiversity. These results are further contextualised by considering the outputs of selected patch metrics (McGarigal *et al.*, 2009) calculated for each *GDIx* class. A total of 131 low diversity patches compared to 77 high and 70 very high *GDIx* classes were detected within the UTDM.

Interestingly, despite the lower patch number for the high-very high *GDIx* classes, the higher *GDIx* classes recorded larger mean patch sizes, by a factor of 2, compared to the lower-valued *GDIx* class. Moreover, the mean perimeter-to-area ratio, which represents shape complexity, suggests that the high-very high *GDIx* patches are less complex in shape. Alternatively, the patch area exposed to edges is lower and less pronounced than the low-medium *GDIx* class. The perimeter-to-area ratio values for the low-medium GDIx values also suggest that these regions are characterised by *GDIx* patches with elongated shapes or indented and are generally small patches and therefore have higher perimeter–area ratios than the larger-shaped and unbroken edges of the high *GDIx* patches (Helzer & Jelinski, 1999).

Table 4.4: Categorisation of frequency of selected landscape metrics summarised by Geodiversity class

Geodiversity index	Geodiversity class	Area (ha)	Mean Patch Size (ha)	Mean Perimeter-area ratio (shape complexity)	Number of Patches	Area Ratio (%)	Cumulative frequency (%)
0-1	Very low	92 452	1 566	80.4	59	8.1	8.1
1-2	Low	252 339	1 926	166.7	131	22.2	30.4
2-3	Medium	343 644	3 818	150.9	90	30.3	60.7
3-4	High	288 287	3 743	70.6	77	25.4	86.1
4-5	Very high	156 387	2 234	49.8	70	13.8	100

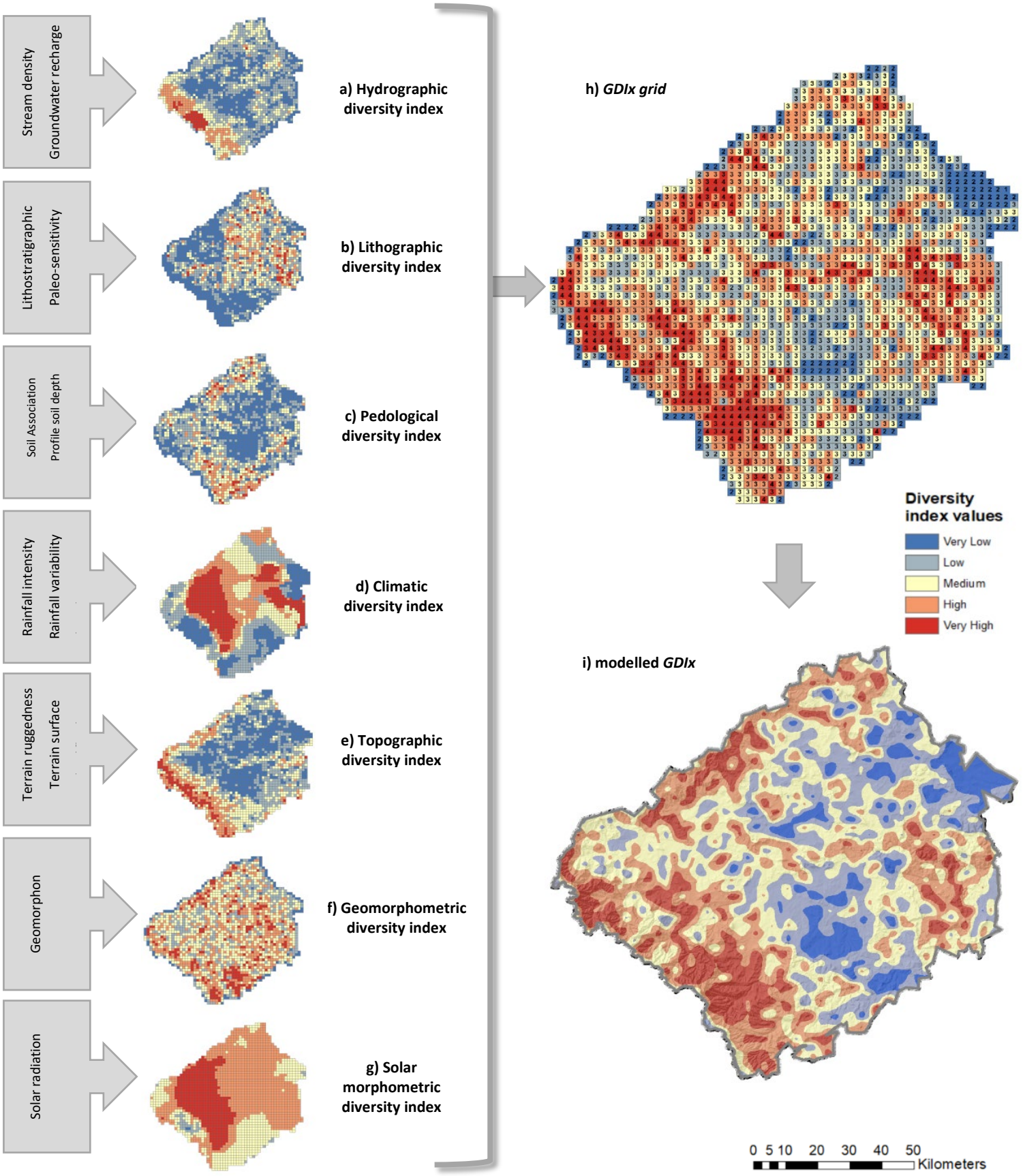


Figure 4.4: Calculation of geodiversity index as the sum of seven sub-indices: **Hi, Pi, Li, Ci, Ti, Gi** and **Si**. The final thematic geodiversity layer is interpolated using a geostatistical interpolation approach to provide a smooth surface of diversity representation and rescaled and classified to a five-class scale of theoretical geodiversity index for UTDM.

4.3.3 Linking Geodiversity to natural resources management

The influence of land use and landscape structure on biodiversity has been intensively studied (McKinney, 2004; Flynn *et al.*, 2009; Parks & Mulligan, 2010). However, less is known about how land use changes affect diversity patterns of abiotic features such as geodiversity (Tukiainen *et al.*, 2017). Furthermore, according to Hudson & Inbar (2012), landscapes that exhibit high degrees of geodiversity have combinations of distinct physical qualities, if not unique, across a given landscape. A better understanding of the connections between biodiversity and the abiotic environment along changing land use gradients is essential in developing sustainable measures to conserve biodiversity under changing regional and global vulnerability scenarios. Therefore, sustainable land resource decision-making is becoming more holistic and broader in scope than the earlier focus on biophysical impacts on sustainable productivity and economic return (Lilburne *et al.*, 2020). The benefit of including geodiversity measures in regional landscape assessments is that they are well suited to landscapes with limited biodiversity data. However, sufficient geographic data is available to characterise geodiversity and relative landscape intactness (Beier & Brost, 2010). Moreover, regional assessments of landscape and resource character benefit from comparing landscape diversity through a better understanding of landscape feature representativity rather than landscape feature rarity in a landscape that considers one landscape different from another instead of better or worse (Erikstad, 2013).

Here I ask: what is the spatial relationship between geodiversity and other forms of land resources within the UTDM? Furthermore, how do these relationships characterise the regional landscape from a resource diversity perspective? The histograms presented in **Figure 4.5** illustrate the frequency distribution results for the UTDM *GDIx* and *BRU*, *NLCap*, *NLCov*, *NCS*, *NTB* and *BPS* (**Table 4.2**). The BRU and *GDIx* analysis results show that areas of medium-high geodiversity are concentrated in upland regions. These results coincide well with the ***Gi*** and ***Ti*** partial index results on the relationship between geodiversity, altitude and topography. Surprisingly, the lowland *BRUs* predominantly characterised by low-altitude, homogenous terrain and fluvial geomorphology and most prioritised for anthropogenic purposes show very little geodiversity.

Furthermore, the Cramer's *V* and chi-squared test-scores (**Table 4.5**) are used to evaluate the following statistical hypothesis: *Geodiversity distribution is not associated with the distribution of land resource features in UTDM*; that is, the predicted *GDIx* regions do not show any statistical association with the land resource coverages (H_0:$\mu 1 \neq \mu 2$..,). The null hypothesis findings would explain that geodiversity hotspots are spatially dissimilar/ independent from land resource variables. The alternative hypothesis would then evaluate the following: *Geodiversity distribution is associated with the distribution of land resource features in UTDM* (H_1: $\mu 1 = \mu 2$.,). The null hypothesis for this categorical analysis is evaluated in a contingency table to compare the frequency of distributions between geodiversity and the various resource layers with the chi-square test for independence (α=0.05) and the Cramer's *V* coefficient used to interpret the effect size of the strength of association between the variables. In a more general sense, the contingency table and chi-square tests determine whether distributions of categorical variables differ from each other. The hypothesis test results for the relationship between geodiversity and land resource covariates show a medium effect-size of association between the *GDIx* and *BRU* regions with the null hypothesis rejected (α=0.05), suggesting that there is indeed a statistical association between *BRU* regions and *GDIx* values in UTDM.

The histograms for *NLCov* and *NTB* concerning *GDIx* frequency indicate that natural Grassland and natural Wooded Land and the Grassland and Savanah Biomes account for much of the region's geodiversity. The

uKhahlamba-Drakensberg range is the most conspicuous upland-montane region in Southern Africa and a significant landscape feature of the UTDM. Moreover, the Tugela River's primary catchment area is known to be dominated by fire-climax grasslands (Elliot & Escott, 2015). Therefore, a positive correlation of high *GDIx* values within the endemic Grassland Biome in the Upland and Montane areas in the UTDM is expected. The Grassland Biome of the UTDM forms part of the Montane region of the Drakensberg Alpine Centre (Brand *et al.*, 2019), which is composed of three main floral compositions: uKhahlamba Basalt Grassland and Highland Basalt Grassland within the upper montane region and Drakensberg Afroalpine Heathland within the lower alpine region. Brand *et al.* (2015) further identify four central vegetation communities for the area, which includes wetland grass and forblands, sheetrock grass and forblands, high altitude alpine grassland, and high-altitude alpine fynbos grassland. What is encouraging is that both within *NLCcov* and *NTB* layers, the highest geodiversity loci are concentrated in natural Grassland Biomes. These findings tie in with the *GDIx* results obtained for the upland and montane *BRUs.* These grassy systems are extensive and support a rich and distinct diversity of plant and animal life. Studies have shown that plots of 100 m^2 can contain up to 40 species. Within a 1000 m^2, a natural grassland state can support an average of 82 different species making the grassland biome the second largest biome in South Africa, after the Savanna Biome, and second only to the Fynbos Biome in respect to its endemism of floral vegetation biodiversity (Cowling *et al.*, 2004; Zaloumis, 2013).

Table 4.5: Chi-Square and Cramer's V results for geodiversity and land resource distribution

***H₀**: Geodiversity distribution **is not** associated with the distribution of land resource features in UTDM*
***H₁**: Geodiversity distribution **is** associated with the distribution of land resource features in UTDM*

Feature	Cramer's *V*	Interpretation of effect	Chi – X^2	Degrees of Freedom	Critical Value	p-value	Null hypothesis
BRU	0.247	medium	82.7	12	21.03	0.04	*Rejected*
Land Capability	0.175	small	54.8	20	31.41	0.47	*Rejected*
Land Cover	0.1	small	97.8	20	31.41	0.52	*Fail to reject*
Conservation	0.091	small	7.45	8	15.51	0.01	*Rejected*
Biome	0.075	small	7.56	12	21.03	0.01	*Rejected*
Protection	0.238	Medium	50.9	8	15.51	0.56	*Fail to reject*

Given the medium to high-frequency distribution of *GDIx* values in the *NLCov* and *NTB* resource layers, the corresponding results of statistical association in the hypothesis test are therefore unexpected. The statistical analysis indicates a small effect size of the association between the *GDIx* and *NLCov* resource boundaries, suggesting that *Geodiversity distribution is not associated with the distribution of land resource features in UTDM.* Conversely, though the association effect is small for *GDIx* and *NTB*, H_0 is rejected (p=0.01), indicating that *Geodiversity distribution is associated with the distribution of National Terrestrial Biomes.* We would expect to see both *NLCov* grassland and *NTB* grassland exhibit similar statistical associations, given their similar spatial associations, to the *GDIx* values. However, these differences in results may largely be attributed to the limitations imposed on using datasets of varying scales to evaluate geodiversity at a regional level. The *NLCov* is a highly detailed National 20 m resolution land cover product (GTI, 2019). This detail is necessary to accurately detect and classify land use and land cover discontinuities over short distances within the landscape. By comparison, the *NTB* resource layer is a broad-scale vector dataset intended to define the nine biomes at a national scale with limited micro-scale applicability. Furthermore, biomes broadly correspond with climatic regions as moisture and temperature strongly influence plant establishment and survival.

This may also explain why the association between *GDIx* and *NTB* is better represented, considering that climatic and solar morphometric were used as partial indices to derive the final *GDIx* values.

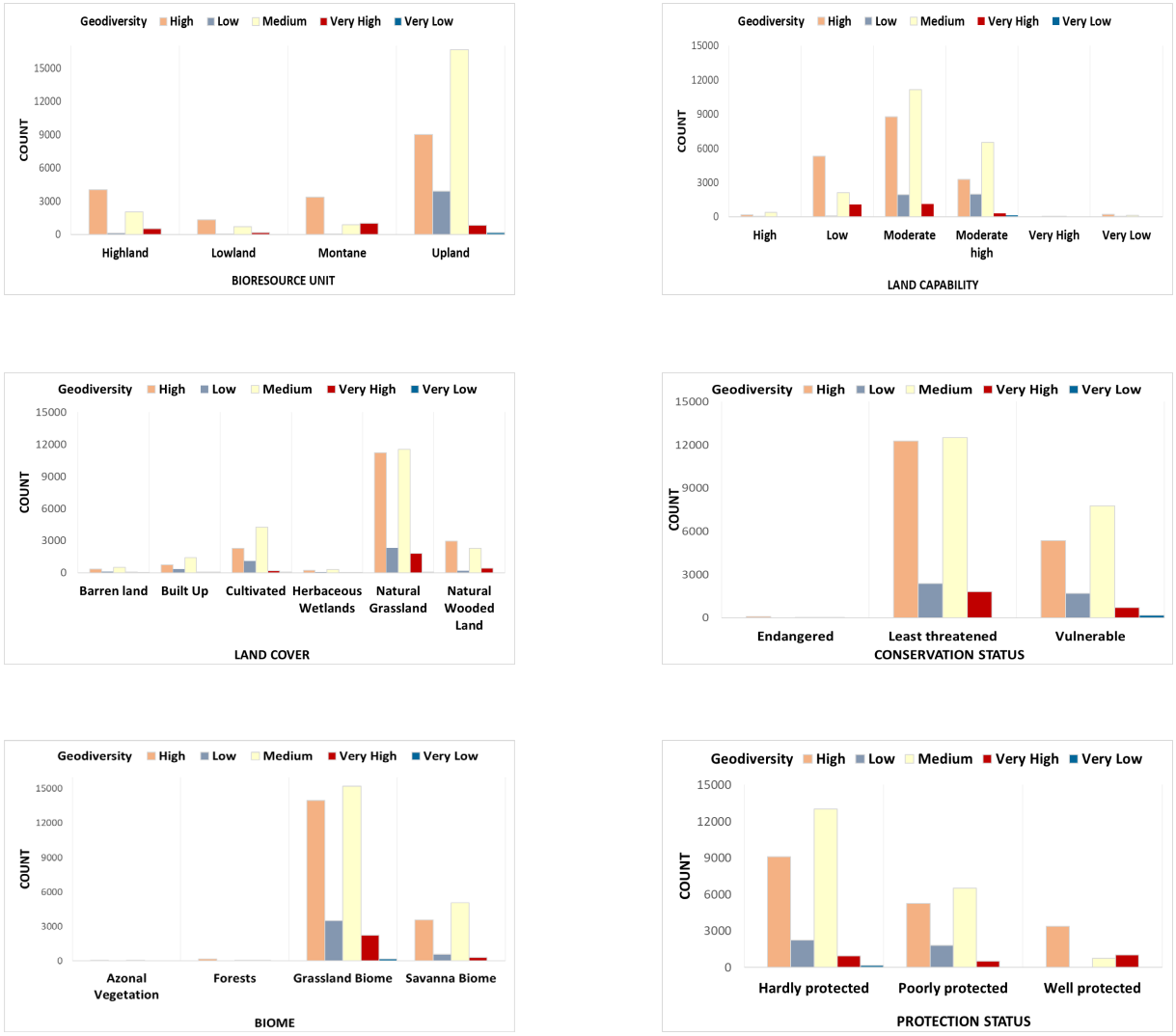


Figure 4.5: Histogram description of landscape resources categorised by geodiversity index class.

The remaining land resource associations with geodiversity provide perhaps the most significant implications for geodiversity conservation in the UTDM. First, areas classified as having *Vulnerable* or *Least Threatened* conservation status harbour the highest geodiversity within the UTDM. Second, defined medium-high geodiversity areas are either *Poorly* or *Hardly Protected* based on *GDIx* and NPS's relationship. One of the biggest threats to landscape-scale conservation is land use change: modifying natural and semi-natural environments to more anthropogenic landscapes (Vitousek *et al.*, 1997). Given that *NLCov* and *NTB* prioritise grasslands with high geodiversity, the findings between *GDIx* and *NCS* and *NPS* help prioritise regions of dual bio- and georichness importance. Driver *et al.* (2012) estimate that close to 18 per cent of South Africa's natural land cover is transformed, mainly by cultivation (10.4 per cent), degradation of the natural cover (4.5%), urban land use (1.5%), and forestry (1.4%). A further 0.7 million hectares of land are collectively degraded and left bare by soil erosion (sheet and gully erosion).

In comparison, 4.6 million hectares of natural vegetation are degraded, mainly in indigenous forests, woodlands, and grasslands. Therefore, it is interesting that the *GDIx* and *NCS* results for the UTDM focus on the *Least Threatened* and *Vulnerable* categories whilst, nationally, the outlook is on *Vulnerable* or *Endangered* priorities. In fact, Matsika (2007) confirms that South Africa's Grassland Biome has been identified as critically endangered and the biome most requiring conservation attention by implementing efficient, sustainable systematic conservation plans. Not much convincing is necessary to appreciate the value of these biomes. If well managed, these grasslands maintain the vegetation cover that mitigates the loss of topsoil during precipitation events, effective water absorption into the soil, and the slow attenuation of water into the hydrological system. These extensive grasslands are also crucial for carbon sequestration, primary production, carbon storage and nutrient cycling, and reducing soil erosion and sedimentation. All the same, the results for *GDIx* and *NCS* in UTDM, concerning *Least Threatened* and *Vulnerable* status, are corroborated by Elliot & Escott (2015), who also identified a total of three grassland vegetation types as vulnerable and eight types as least threatened with only the Eastern Mist belt Forest as being endangered. Finally, regarding the *NPS* regions' geodiversity, it is worth reiterating that the *GDIx* values have been derived from a total of seven partial index diversity scores. Given the multi-modal richness of abiotic factors in the *GDIx* quantification and the interoperability of selected natural resources, it may be necessary to reprioritise the protection status of medium-high geodiversity areas from *Poorly* or *Hardly Protected* to *Well Protected.* The results of this study suggest that geodiversity is indeed a significant driver of the distribution of biota. This view is supported by Comer *et al.* (2015), who highlight that the greater the assortment of environmental conditions, the higher the total biological diversity the environment can support. Identifying high geodiversity areas enables the recognition of niches within the same environment, allowing a higher degree of biodiversity to co-exist (Parks & Mulligan, 2010). Fortunately, there is no shortage of international agreements or conventions and national regulations and guidelines for heritage and geo-conservation management in South Africa.

Many organisations, including the International Union for Conservation of Nature and Natural Resources (IUCN), the International Union of Geological Sciences (IUGS), and UNESCO, contribute to landscape conservation in various ways, including a geodiversity component. Specific to South Africa, two key parliamentary statutes aim to regulate the utilisation and ownership of geological resources, i.e., the Geoscience Act (GSA, 2010) and the National Heritage Resource Act, 1999 (NHRA, 1999). The NHRA, in particular, aims to address South Africa's imbalances in the representation and management of its heritage following apartheid and introduce a heritage management system that is reflective of its cultural diversity and constitutional democracy (Deacon, 2015). That being said, while the NHRA coordination - through the South African Heritage Resources Agency (SAHRA), serves to protect South Africa's valuable cultural heritage resources, two notable caveats remain unresolved: First, If geological specimens cannot meet the criteria of being cultural heritage or rare artefacts then ideally they cannot fall within the ambit of the NHRA (Cairncross, 2011). This, of course, limits the interdisciplinary contributions of geodiversity beyond archaeological and paleontological considerations. Disciplines such as sustainable land management, natural resource management or even integrated environmental management would instead consider pedological and geomorphological diversity rather than cultural value and archaeological rarity high on the developmental agenda. In fact, S24(3) of the National Environmental Management Act (NEMA, 2008) has already identified and managed geographical areas of importance in South Africa. Currently, NEMA only prioritises geographical regions for renewable energy production. However, suppose a readily accessible, accurate and repeatable approach is demonstrated to provide a pragmatic solution to quantify geodiversity. In that case, a similar strategy may be considered a primer for prioritising regions with high-value agricultural land, water catchment conservation, or even ecosystem-based services and resilience. For a synthesis of legislative and management context of landscapes and geoheritage sites in South Africa, readers are referred to (Grab & Knight, 2015).

4.3.4 Geodiversity and agricultural land use management

South Africa is characterised by a significant variance in its natural resource potential (EnviroGIS, 2016a). Land users are increasingly expected to maintain necessary food production levels with a minimal negative impact on these resources in the process. Therefore, medium-high value agricultural potential land is considered a limited non-renewable resource that must be managed based on the most suitable land-use option. Consequently, it is paramount to the sustainability of present and future land use activities to identify and delineate (zone) agricultural land, based on its inherent capability and suitability, to be preserved for exclusive agrarian use. Here I recognise the definition of "land capability" to imply the most intensive long-term use of land for purposes of rainfed farming, determined by the interaction of climate, soil and terrain (EnviroGIS, 2016a). This further emphasizes the concept of the agro-ecosystem, or more specifically, the agro-ecological zone (AEZ) principle developed by the Food and Agricultural Organization (FAO)(Fischer *et al.*, 2000) of the United Nations, as a functional land planning unit. An AEZ is defined as a land resource mapping unit with a specific set of climate, soil, landform and landcover factors, limiting the range of potential land uses. AEZ's are thus areas with similar characteristics concerning land suitability, production potential, and environmental impact. Such an AEZ can form the basis for agricultural land use planning but requires collecting empirical data at a relevant scale for local-level planning, including soil, vegetation, climate, hydrological and degradation data. Incidentally, this definition fits well within the scope of geodiversity assessment.

Not surprising, the development of the KwaZulu-Natal Agricultural Land Potential Categories Demarcation Tool is based on the above AEZ approach. The land categories are instrumental in providing development controls for agricultural land categories indicating permissible development intensities. A further set of development controls applies to a combined designation called Agro-biodiversity. This designation is spatially integrated into a provincial *Agricultural Land Categories Dataset* (Elliot & Escott, 2015) that is used in conjunction with KwaZulu-NatalDARD's Bioresource Program (BRW) (Camp, 1995; KZNDARD, 2015). The Bioresource Programme consists of Bioresource Groups (BRG's representing vegetation zones) and Units (BRU's representing relatively homogeneous agro-ecological units)(Scotney, 1987; Camp, 1995). Bioresource units are defined as zones where the soil type, vegetation, climate, and to a lesser extent, terrain type are similar. Consequently, the BRW provides a framework for uniform recommendations regarding land use, natural resource management, and agricultural practices and helps land users make informed decisions. The BRW is the primary operational geospatial decision support resource for KwaZulu-Natal. It readily provides a reconnaissance appraisal of the natural resources for environmental impact assessments and the agricultural potential. It would benefit from including geodiversity resources in the framework, particularly to the advantage of delineating AEZs.

Agriculture (cultivated land and stock farming) is the predominant land use activity in the UTDM (Cox *et al.*, 2015), with commercial and subsistence agriculture accounting for nearly 13% of the district. This provides an essential contribution to the economy in production and provision of resources for up and downstream processing, and manufacturing and job creation (COGTA, 2017). Therefore, the importance of the correct use of natural resources to the overall economy of the UTDM and the Province of KwaZulu-Natal can not be over-emphasised. The results for the chi-square test to assess whether *Geodiversity distribution is associated with the distribution of Land Capability* in the UTDM are therefore strategically relevant to national interests that aim to preserve and promote sustainable use and development of agricultural land for the production of food, fuel and fibre for the primary purpose to sustain life (PDALB, 2016). The results for *GDIx* and *NLCap* were found to be statistically significant, with H_0 rejected in favour of H_1. Considering the Cramer's V coefficient of 0.175, these results indicate that the distribution of *medium-high* geodiversity zones shows a small effect

size of association with *medium* and *medium-high NLCap* categories. In South Africa, land capability is further linked to agricultural potential. The Preservation and Development of Agricultural Land Bill (PDALB, 2016) define agricultural potential as the measure of potential productivity per unit area and unit time achieved with specified management inputs for a given crop or veld type and management level. The agronomic potential is primarily determined by the interaction of climate, soil and terrain. The overlap of similar abiotic variables to geodiversity quantification suggests that there may be concurrent linkages between direct and latent agri-potential and landscape geodiversity representation. The findings of the statistical analysis indicate that *moderate* and *moderate-high NLCap* classes exhibit the highest geodiversity. In contrast, areas of *high-very high NLCap* contain minor regions of increased diversity. These results support earlier studies by Barichievy & Botha (2017), who comprehensively evaluated soil ecotopes' distribution in KwaZulu-Natal, including yield for selected perennial crops by land capability. Their results indicate that moderate potential land is often underestimated in the diversity of soils and production potential. While these intermediate potential lands typically require significant interventions to achieve viable and sustainable food production, these moderate potential lands are still the majority of land use in the rural landscape in the absence of high-value agricultural land. However, from a land use perspective, agricultural conservation motives often conflict with those of ecological preservation. Regardless, the acknowledgement that the quantification and evaluation of geodiversity can guarantee the variability of an ecosystem and the survival of unique landscapes and, above all, of different forms of biodiversity is the starting point for a range of sustainable development applications and adaptive landscape decision making. Zunckel (2013) further adds that where natural infrastructure, including geodiversity, is prioritised, maintained and managed to deliver a range of ecosystem services, a suite of additional benefits will be realised. These benefits include improved agricultural productivity, improved landscape integrity, securing cultural benefits, improved resilience to natural disasters, and increased adaptive capacity to the impacts of climate change.

While the overall *GDIx* has been quantified based on the contribution of multiple partial abiotic variables, it is necessary to specifically recognise the contribution of the functional value of soil diversity as a vital natural resource and, in fact, the "substrate" on which concepts such as land capability, landscape pedodiversity and more precisely landscape agroecology, the ecology of a productive landscape (Wojtkowski, 2003), are reliant. Why soil specifically? Soils are not uniform. Given the vital role of soils for human survival, soil diversity (pedodiversity) preservation merits exceptional attention, more so than other natural resources (Ibáñez & Bockheim, 2013). The value of soils' functional diversity must, therefore, be recognised. Pedologists have recently started paying attention to soil diversity using the same mathematical tools that ecologists use and reaching exciting relations between the spatial patterns of soil and landscape characterisation (Ibáñez & Gómez, 2016). Soils are the vital link between the inanimate world (geosphere) and the living world (biosphere) and one of the thinnest and most vulnerable human resources upon which, both deliberately and inadvertently, humans have had a significant impact (Bridges, 1994; Goudie, 2018). Brady & Weil (2008) were amongst the first to recognise the need for a global perspective on soils - changes in soil productivity in one area affect food security and food prices, biodiversity and water quality in nearby and distant places. Framing soil diversity in a multi-functional global perspective is paralleled by the increasing appreciation of the ecosystem concept as the prime basis for decisions on natural resource management.

To maximise agricultural production, Ramamurthy (2018) highlight the importance to know the basic facts of soils, their potential and limitations and the management strategies needed to overcome potential limitations. Likewise, there are continual advances recognised in state-of-the-art approaches to both the dynamic (e.g. soil health, organic matter content, aggregation, density) and inherent (topsoil depth, soil texture, soil colour) modelling of soil properties in the surface horizons where the effects of land management are most expressed in the landscape (Bünemann *et al.*, 2018). Indeed, the prioritisation of soil diversity, including geodiversity, to

inform mainstream evidence-based landscape policy agenda requires relevant, reliable, and applicable soil-based information.

Therefore, synoptic DSM approaches are *sine qua non* as a solution to increase the effectiveness of modelling and visualising soil-based phenomena (McBratney *et al.*, 2003), typically based on ancillary data. These approaches incorporate methods of prediction accuracy assessment to infer patterns of soils across various spatial and temporal scales (Grunwald, 2010). The overall *GDIx* map product presented in this study is an aggregation of inter alia, pedological, geomorphic, and topographic diversity that exhibit strong relations to overall landscape geodiversity. These components are equally important since they form part of the conceptual framework that states that soils can be described by the primary environmental soil-forming factors: climate, organisms, relief, parent material, and time (CLORPT) (Jenny, 1941). Recently, McBratney *et al.* (2003) have extended this framework to the *scorpan-based DSM* where covariates are chosen to represent climate ('c'), organisms ('o'), relief ('r'), parent material ('p'), and time ('a'); known soils ('s') are used for calibration, and neighbourhood ('n') (Zinck, 2013) approach for predictive modelling (and mapping) of soil, which in itself is rooted in earlier works by Jenny (1941) and Russian soil scientist Dokuchaev. There is a natural synergistic opportunity for geodiversity and pedodiversity, as a primer for DSM and mapping of soil attributes at scale, to contribute to landscape planning approaches that are flexible, responsive and adaptive to landscape characterisation and diversity quantification within the UTDM. The importance of recognising geodiversity, including its partial indices thereof, as a cornerstone or its own inherent sake should constitute a significant component for present and future landscape character assessments in the UTDM. This would likewise require further exploration into geodiversity assessment approaches and products into provincial land and agro-ecological planning tools such as the Bioresource Program or even National planning tools such as the Land Capability Assessment. However, international discourse outlines that this integration will likely require collaborative inputs from multi-stakeholder agencies to consider whether practical approaches to conserving geodiversity can support and complement their core species-centred mandates (Comer *et al.*, 2015).

Holistic interventions are needed to protect and rehabilitate agricultural land through coordinated planning of zoned agricultural land. South Africa is known for its progressive and innovative environmental legislation that reflects the natural, physical, economic, and psychological importance of the environment to humans (Grab & Knight, 2015). Specific interventions already provide an opportunity to comprehensively and systematically integrate geodiversity into adaptive decision making. The South African regulatory framework on spatial planning, land use, land use management and land development is well defined in the Spatial Planning and Land Use Management Act 16 of 2013 (SPLUMA) – and all land use planning legislation has to be consistent with its provisions (Letuka, 2016). One of the central frameworks governing the preservation of high value and best available agricultural land provincially is the requirement to conduct a comprehensive agricultural impact assessment (AIA) before the subdivision or rezoning of agricultural land is authorised. In particular, the AIA outlines the specific soil investigation procedures to collect and present harmonised findings on natural agricultural resources, agricultural land capability and agricultural land suitability for each landholding (DAFF, 2017).

Additionally, under the PDALB, the State must protect agricultural land, particularly high-value agricultural lands, for agricultural production in permanency. An exciting approach to geodiversity importance may need to consider how an ecosystem-based approach to resource management can provide a framework for better integrating geodiversity, biodiversity, and landscape management. This will provide a means of realising geodiversity's broader values and benefits through its contribution and functionality in delivering ecosystem services. From a national perspective, including areas of high geodiversity into systematic agro-ecological planning may also provide a valuable platform to address a range of contemporary policy, strategy and

landscape-action sustainability issues: What geodiversity does a country have, and can we map it? What is the condition of geodiversity across the landscape? Where and how should a country act first to manage and mitigate geodiversity importance? The concept of "conserving natures stage" is defined by Lawler *et al.* (2015) as a concept that encapsulates the idea that maintaining a varied physical landscape will enable diverse ecological processes to operate, protecting and promoting landscape diversity. Failure to acknowledge such responses may cross essential geomorphic and ecological thresholds that initiate land degradation and increase human risk and vulnerability to natural hazards.

4.3.5 Future work

The methodology presented in the study is, in fact, a sub-national adaptation of the method(s) proposed by Araujo & Pereira (2018). The findings, therefore, confirm the application of their approach within a regional southern Africa, something which has not been tested before. A significant advantage of this study is that it introduces additional sub-index covariates to detail a richer perspective of local geodiversity in the UTDM. The proposed methodology simultaneously considers abiotic feature richness and abundance that can be readily developed and deployed in many proprietary or open-source GIS platforms. While this chapter aims not to describe the full geodiversity of the UTDM, the approach developed in this work has one main disadvantage: The quantification and prediction of geodiversity do not consider the combined geovalues and products of the region's physical resources. This study has focused on limiting any considerations of geodiversity's s intrinsic value, i.e., the belief that some things (in this case, nature's geodiversity) are of value simply for what they are rather than what they can be used for by humans (utilitarian value). Both Gray (2004) and Wilson (1994) point out that is the most challenging value to describe since it involves ethical and philosophical dimensions of the relationships between society and nature. Instead, this study aimed to highlight the utility of adopting an objective quantitative approach to geodiversity valuation based on the functional value, or utilitarian interpretations, of the supporting surrogates that contribute to landscape diversity. Nevertheless, to fully meet end-users needs, it may be necessary to translate this kind of quantitative geodiversity information into qualitative maps that explicitly address the importance of geoheritage and geoconservation of geotopes or geomorphological features within UTDM. Therefore, in addition to functional and scientific values of geodiversity, future studies may consider evaluating other overlapping potentialities such as cultural, aesthetic and economic values of geodiversity, which are equally as important and may provide a full range of local and regional geodiversity (Gray, 2004; Araujo & Pereira, 2018). Nevertheless, the evaluation of geodiversity should focus on the conservation of resources and their management, which are vital to the development of human activities (Benito-Calvo *et al.*, 2009). In this context, the method presented herein provides a reliable and user-friendly approach to unravelling geodiversity in KwaZulu-Natal, but there is still much left to do.

A caveat to the present study is that since no other attempt has been made in KwaZulu-Natal, until now, the geodiversity results for the UTDM cannot be readily compared with those of different areas. The UTDM is considered an archive of geoheritage and biodiversity importance. Future research in the region may, therefore, present new findings of geodiversity significance. In this case, the current geodiversity assessment and associated maps for the UTDM will evolve as better data and processing capabilities are explored, allowing improved quantification of landscape heterogeneity and homogeneity. It must be noted that future refinements of geodiversity assessment may benefit from systematic improvements to the present methodology. Firstly, the inclusion of other surrogate elements of intrinsic geodiversity importance, such as detailed lithostratigraphic resource data or improved soil classifications for the region or further includes climatic

variables, beyond those applied in this study, since increasingly more studies on climate change vulnerability assessment and adaptive management are being conducted at the scale of landscapes.

Further exploration of the utility of the *GDIx* in micro-landforms may benefit from the extraction of landscape and topographic data from localised, high-resolution digital elevation data sources. Stellenbosch University's Centre for Geographical Analysis provides a 5 m Stellenbosch University Digital Elevation Model (SUDEM), which offers significant-resolution and radiometric improvements to the SRTM 1-arc second (30 m) surface product, or the 2 m Digital Surface Model (DEMSA), available for anywhere in South Africa with a mean vertical error of 35 cm. Both products will allow the low-cost quantification of terrain features related to geodiversity assessment at pragmatic spatial resolutions as high as 20 m across South Africa. A final consideration beyond the present study's scope is the possible application of hexagonal grids, rather than the standard fishnet grid, for the numeric quantification of the *GDIx*. A significant advantage of using hexagonal grids is that they offer reduced edge-effect. They provide the lowest perimeter to area ratio of any regular tessellation of a sampled plane. Moreover, in a hexagonal sampling frame, all neighbouring features are identical. Hexagonal grid cells have six similar adjacent cells, each sharing one of the six equal-length sides in contrast to square cells with two classes of neighbours: those in the cardinal directions that share an edge and those in diagonal directions that share a vertex. Finally, hexagonal sampling may offer increased benefit to users' intent to use vector-based abiotic variables for geodiversity quantification since hexagons can better fit curved surfaces than squares across large areas such as landscapes.

4.4 CONCLUSION

To date, only one other landmark study has successfully investigated the influence of environmental factors on geodiversity mapping in South Africa (Kori *et al.*, 2019). This work was the first attempt in KwaZulu-Natal to produce a full extent geodiversity map for the UTDM using a geospatial index approach to objectively quantify the influence of physical environmental factors on landscape geodiversity. The proposed contribution of partial environmental indices, readily obtained from various open-source platforms and harmonised to a spatial resolution of 90 m to successfully produce a 90 m geodiversity map for UTDM, can become a reference for inclusive geomorphological mapping for KwaZulu-Natal.

This study assessed and mapped geodiversity by quantifying seven primary components of geodiversity importance based on the hydrographic lithostratigraphic, pedological, climatic, topographic, atmospheric and geomorphological diversity across the landscape. The semi-quantitative methodology for assessing geodiversity relies heavily on the approach outlined by Araujo & Pereira (2018), with minor modifications related to sample-grid resolution (resolution of 2500 m x 2500 m) and the specific quantification of feature richness and abundance from varying spatial data formats, using a data scale of 1:250 000. The methodology used in this study relies on the input of numerous spatial datasets of varying data format, scale and measurement scale. The results of this study provide a valuable resource to evaluate the limitations of data normalisation necessary for geodiversity score standardisation when constructing the geodiversity analysis.

This research set out to provide a first approximation *GDIx* map by adapting previous methods for the entire UTDM. Second, to evaluate the selected surrogate partial-index diversity's spatial utility as explanatory variables for overall contribution to *GDIx* ranking(s). The methodology used for quantifying the *GDIx*, using both vector and raster datasets, proved to be effective in highlighting the richness and abundance of features for the respective partial sub-indices. The main results of the geodiversity analysis were categorised and graphically represented into five classes from very low to very high geodiversity. The topographical, pedological and geological diversity estimation, as well as the geomorphometric diversity quantification all, showed strong associations with regions in the northern, southern and western areas of the UTDM characterised by the Drakensberg mountain range, higher altitude, complex terrain and associated increased diversity of soil and geological material. A key finding is that these partial-indices consequently strongly influence the final regionalisation of geodiversity with areas of medium partial-index diversity randomly distributed throughout the UTDM. These high-index value regions of geodiversity indicate that these areas should be prioritised from a land management perspective. On this note, I summarise critical reviews regarding explicit linkages between the importance of *GDIx* and pedodiversity. The priority of soil conservation for meeting the objectives of many of the Millennium Development Goals (MEA, 2015) is high on the international development agenda. Therefore, novel approaches, with an emphasis on combining DSM and geodiversity assessment, for popularising digital soil inventories are needed to address the emerging concerns about the functioning of soils systems in the provision of services required by modern societies within present planetary constraints. A key consideration of this research is highlighting a simplistic, objective and replicable numerical approach for deriving a *GDIx* product at a landscape level, with rudimentary technical expectations placed on the end-user, at minimal expense to representing the reality of each input element (partial index) diversity. Landscape character assessment for sustainable land use decision making is becoming more holistic and broader in scope to understand better the connections between biodiversity and geodiversity within changing resource gradients.

One of the objectives of this study was to assess the relationship, both spatially and statistically, between geodiversity distribution within the UTDM to several strategic public-domain landscape resource coverages currently in active practice: *National Land Cover*, *National Agric-Land Capability*, *Regional Bio-Resource Units* and *National Vegetation Protection Status*. These relationships highlighted the opportunity benefits of including areas of high geodiversity into systematic conservation planning to address a range of current policy, strategy and landscape-action sustainability issues for regional landscape planning. Applying the *GDIx* map and the various land resource maps presents an original contribution to the region. The results were informative for showing that the geodiversity is statistically associated in the UTDM region, albeit with a small association effect, with land capability, regional bio-resource units, national biome types, and biome conservation status. Areas of medium and high *GDIx* were shown to be positively associated with areas of medium to high land capability and predominantly localised to Grassland and Savanah Biomes. An encouraging finding is a relationship between medium and high *GDIx* values with vulnerable regions and only moderately threatened land resource conservation. For that matter, no areas of high diversity were associated with land resources classified as endangered. These results support the my views that applying a *GDIx* in land resource planning has excellent potential for recognising and accepting the dual benefits for the conceptual, empirical and spatial landscape modelling concerning overall landscape diversity.

The *GDIx* method was tested across an area of approximately 11 500 km^2. Since the methodology relied on the utility of high-resolution (temporal and spatial) open-source covariates to derive the partial indices of geodiversity at a region-level, there should not be any limitations for considering its application to other geographical settings in South Africa. There is no superior level of geodiversity quantification and representation. While the methodology presented in this study provides a pragmatic approach for geodiversity assessment, further adaptions based on personal user preference and research objectives such as sampling grid resolution, method of geodiversity assessment and quantification, spatial data analysis, the scale of analysis and application of the final product provides an opportunity for further exploration. Importantly, it must be emphasised that the present study's geodiversity results merely represent the diversity (richness and abundance) concerning the observed partial indices for the region where no other studies on geodiversity have been undertaken from this perspective. Therefore, the concept of geodiversity remains somewhat “flexible” in that future studies may indeed consider a new suite of partial indices for geodiversity quantification. The influence of this flexibility is reflected in the results for the partial-index diversity decay calculation based on including multiple sub-partial index scores for each abiotic component to the final *GDIx* values.

The last decade has seen a significant increase in the quality of landscape indices through the application of geoinformatics research aimed directly at assessing geodiversity. Over this period, land-surface analysis and classification have seen rapid improvements in the rate and quality of geomorphometric computational approaches, a key factor for the utility of digital geomorphic mapping. In truth, the substantive contribution of geographical information systems and Remote Sensing has provided a platform for users to digitally access environmental information readily at different spatial scales and levels of sophistication. This has contributed significantly to providing new information about the mechanisms that underlie the current patterns of geodiversity. RS further provides a segway to endless opportunities to merge knowledge from different domains (e.g., software, mathematics, engineering, geomorphology) to study landscapes more interdisciplinary. GIS systems provide a solid platform to synoptically explore the parameters driving geodiversity and promote a systematic digital approach to accurately mapping landforms and geomorphic systems with different climate, geological, and topographic settings. Specifically, advances in RS of land surface terrain features and geospatial modelling systems that support numerical modelling of surface biophysical processes have reformed traditional geomorphic analyses and mapping. Pioneering techniques, such as these, may benefit future geodiversity research since the focus is on integrating quantitative

topographic information, on a cell-by-cell (gridded format) basis, with some form of geospatial platform with an ever-increasing reliance on geostatistical approaches and artificial intelligence to describe the dynamic landscape and topographic change, patterns, and complexity. These digital systems, data and tools, often referred to as geomorphic decision support systems (GDSS), are increasingly favoured for their contributions to diagnostic assessments and modelling of the terrain to achieve enhanced interpretations of scale, patterns, processes and diversity of landscape features and systems.

Internationally, there is growing awareness and acceptance of geodiversity in territorial management, particularly concerning the preservation of natural diversity to promote resilient landscapes. The present study is expected to intellectually incentivise future studies on geodiversity in South Africa as persistent environmental pressures will soon demand maps that express this concept to become more frequent as decision-support tools. If prioritised as a significant element for adaptive management, the *GDIx* assessment presented in this study can inform many kinds of planning and decision-making supporting sustainable development. It implicitly considers the diversity contributions of seven critical abiotic components related to geology, soils, climate, hydrology, atmospheric, topography and geomorphology to overall landscape geodiversity importance. One of the primary outputs is a simple cartographic representation of ranked diversity importance of the quality of geodiversity for the entire UTDM that is easily understandable by a varied audience. Therefore, the *GDIx* map(s) should be considered a tool for integrated natural resource management, monitoring and reporting at a regional level. Given the structural ties between geodiversity and biodiversity, this will facilitate leveraging usable land and biodiversity management recommendations and action programs in an integrated approach to environmental management and geoconservation.

CHAPTER 5

5 DIGITAL SOIL MAPPING WITH REFINED TERRAIN CLASSIFICATION PROTOCOLS AS A METHOD FOR ENHANCING SOIL ECOTOPE MAPPING WITHIN THE KWAZULU-NATAL BIORESOURCE CLASSIFICATION SYSTEM[5]

"Although some ecotope studies have been done, the ecotope concept has probably received much less attention than it should have."[4]

[4] Laker, M., 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

[5]This chapter is submitted to Geoderma Regional: Atkinson, J.T, de Clercq, W.P., Rozanov, A.B., 2021. Refining Digital Soil Mapping Protocols to Improve Soil Ecotope mapping in the KwaZulu-Natal Bioresource Classification System. *In Review.*

[5]The work was also presented at the annual Combined Congress of the Soil Science Society of South Africa, the South African Weed Science Society, the South African Society of Crop Production and the Southern African Society for Horticultural Sciences (SASHS), 24-26 January 2020, Virtual/ Online Event, South Africa.

5.1 BACKGROUND

The preservation, development and sustainable use of agricultural land have been highlighted as critical for economic growth and poverty reduction in South Africa. These interventions are no longer political ideologies but now directly underpin the essence of sustainable livelihoods necessary to support South Africa's growing population, expected to grow to 82 million by 2035 (Goldblatt, 2010). As with other countries globally, South Africa is expected to significantly amplify food production to meet the expanding population's needs from the same or even fewer existing natural resources. Furthermore, considerable attention has been placed on developing critical national policy documents to aggressively address apartheid-era legacies of poverty and land inequality. The desiderata of these new transformations and land assessment models emphasise responsibility to, among other things, conservation agriculture, agroecological land governance and agro-economic sustainability by landowners. Due primarily to limitations in optimum climate-soil combinations, the South African landscape is characterised by a scarcity of high potential agricultural land. Indeed, South Africa has a rich and diverse agricultural sector ranging from an intensive, large-scale commercial agricultural sector to low intensity, small-scale, and subsistence farming sector competing for approximately 82% (100 million ha) of its 127 million ha land area (DEA, 2016). However, by international standards, i.e., North America, Asia and Europe, the country's agricultural potential is low (DEA, 2016). Only 14% of this agricultural land receives sufficient rainfall for arable crop production, with a scant 3% considered genuinely fertile and only 1% reliably under permanent crops. The remaining 67% includes the natural grasslands considered suitable for grazing and livestock farming. Still, while precise estimates are not available for the country, the General Household Survey of 2009 and the Census for commercial farms, 2007 (STATSSA, 2009) estimate that almost 20.7% of South African households were engaged in some form of agricultural production, with KwaZulu-Natal alone containing 25.4% of these households. The high reliance on agriculture in the Province further highlights its contribution to regional growth and development outcomes. The sector is responsible for approximately 5.5% of the Province's overall gross value added (GVA) production. However, it is further responsible for almost 30% of National agriculture output (KZNPPC, 2013).

Nevertheless, there has been a steady decline in the number and the average cultivated area of farming units in KwaZulu-Natal over the last fifty years. Jewitt *et al.* (2015) note that in 1960, the average cultivated area per person in South Africa was estimated to be 0.55 ha. This figure decreased to 0.3 ha by 1993, and in KwaZulu-Natal in 2011, commercial and subsistence agriculture equated to 0.14 ha per person, representing a significant decline over time. As South Africa's agricultural land continues to decline irrevocably due to urbanisation and development pressure, this will ultimately lead to changes in land use and non-agricultural developments on high and medium value agricultural land, thereby aggravating an already high national hunger index. As of 2017, this hunger index stands at 51.6%, with another 28.2% at risk of hunger, posing a direct threat to national (and household) food security (DAFF, 2017). Incidentally, it is estimated that, by 2030, an additional 120 million ha of land is required to support the country's growth in food requirements (DAFF, 2017). The caveat to reaching this milestone is that agriculturally productive provinces, such as KwaZulu-Natal, have a limited supply of high-value agricultural land. Interventions are needed to protect and rehabilitate agricultural land through coordinated planning to protect zoned agricultural land. Therefore, it is essential to establish a practical arrangement between national and provincial authorities regulating the sub-division of agricultural land. In order to ensure coordinated, intergovernmental relations, there needs to be national standards and norms for the approval of subdivisions and use change applications. Norms and standards will apply to government at all three levels and ensure that the same factors will be considered, reducing the possibility of conflicting decisions.

Additionally, this will ensure administrative justice and compliance with the law (DAFF, 2015). In KwaZulu-Natal, it has been accepted that detailed natural resource information based on scientifically accurate and relevant criteria is required to develop spatial layers that planners, developers, local government, and other stakeholders can use to guide future development (KZNPPC, 2013). Against this backdrop, the access to readily interpretable soil and crop information is increasingly being prioritised by provincial planning commissions as critical inputs to decision support systems (DSS) for sustainable land management.

5.2 INTRODUCTION

A significant operational decision support resource already developed to provide a reconnaissance appraisal of the natural resources for both environmental impact assessments and the agricultural potential of KwaZulu-Natal is the classification of the Province into a collection of Bioresource Units (BRUs) (Camp, 1995). In total, KwaZulu-Natal is segmented into 590 BRUs, each of which is sufficiently homogeneous in environmental factors (climate, soil association, vegetation type and terrain form), such that uniform land use practices, production techniques and levels, can be defined with a reasonable degree of accuracy (Camp, 1995). The national Land Type survey (LTS) (ARC, 2003) provides the primary data for describing essential soil-landscape information, a descriptive guide for delineating BRU boundaries and associated BRU inventories. Each BRU can further be defined by a range of conceptual soil ecotopes that represent the potential of the soil for production (MacVicar, 1974). Specifically, each ecotope is defined by a unique set of soil properties (form, texture and depth) and soil surface characteristics (rockiness). This means that a significant difference will be found in the natural resources of one ecotope and another. There will be no significant advantage in the further subdivision of land. This soil ecotope concept was first described by Schoeman & Macvicar (1978), who emphasised the need for evaluating land by placing an agricultural potential on it. Incidentally, any estimates of agricultural potential must consider the factors which determine yield, namely management, climate, soil type and slope. It, therefore, follows that land potential will vary according to climate, soil and slope, with continuous management. In this sense, an ecotope provides a convenient option for any user to readily interpret and differentiate between different productive potentials of soils for a given surveyed piece of land. Regrettably, due to technological limitations, the different ecotopes were not spatially defined for KwaZulu-Natal when conceived initially and are instead expressed as percentages of the total area of the BRU and presented as inventories part of the Bioresource Classification Report Writer Program (BRW) (version 9) (Camp, 1995). The programme's primary aim is to serve as a land use evaluation system that can describe the natural resources of KwaZulu-Natal and achieve a sound matching of all forms of land use with the sustainable use of natural resources. The information provided by the BRU program is regarded as the first appraisal of land potential available to planners in the agricultural and ecological fields.

It is vital to know the essential characteristics of soils in order to produce food and ensure sustainable livelihoods. Understanding their properties, distribution across the landscape, potential and limitations, and the management strategies necessary to overcome these limitations (Reddy & Singh, 2018). Therefore, to improve the productive capacity of soils, there is a need to develop soil-based land use plans. Soil surveys provide this much-needed information by investigating the distribution of soils in a particular area, determining essential characteristics, delineating map units and describing the results in a readily interpretable local legend (Van Zijl & Botha, 2016). However, more recently, the heightened demand for finer resolution, locally relevant spatial soil data has been fuelled from both the agricultural and government sectors (Van Zijl & Botha, 2016). The presentation, predictions, and possibilities of the BRU program, therefore, requires revision. In the last several decades, advances in remote sensing (RS) and geographic information technology (GIT) have catalysed the

evolution of digital mapping systems such as digital soil mapping (DSM), predictive soil mapping (PSM) and digital soil assessment (DSA). It is widely accepted that these formalised approaches are the benchmark for generating numerical or statistical soil resource data by predicting soil distribution based on theoretical soil landscape models.

Furthermore, these models characterise soils as a function of ancillary environmental variables, including climate, organisms, topography, parent material, time and spatial location, recognised as the primary factors governing soil formation (Jenny, 1941; Grunwald, 2010; McBratney *et al.*, 2011). DSM and geographical information systems (GIS) may provide a timely and cost-effective cartographic solution for producing and visualising the soil information necessary to categorise and demarcate the diversity and fragility of ecotope regions in KwaZulu-Natal. Lagacherie & McBratney (2006) point out a clear need for precise quantitative relationships between soils and critical environmental factors to facilitate soil data collection and soil modelling worldwide. Such relationships form the basis of DSM techniques, which are widely considered to represent the future of soil surveys (de Carvalho Junior *et al.*, 2014).

A further consideration for improving the current utility of the BRU classification is to provide improved soil-landscape covariate descriptions through amended terrain morphological unit (TMU) classification. The BRU's are presently categorised according to the LTS five TMU classes: crest, scarp, midslope, footslope and valley bottom. These TMUs may be too granular to describe many soil-landscape covariates presently available through medium-high resolution earth observation and digital elevation models (DEM). Recent studies by Atkinson *et al.* (2020) have shown that better descriptions of landform features, using improved classification approaches such as geomorphons (Jasiewicz & Stepinski, 2013), have the potential to describe soil-landscape relationships in better detail. In particular, the representation of legacy soil-landscape covariates within geomorphons may contribute to the concept of present landscape pattern-process character assessment, i.e., the distinct, recognisable and consistent pattern of elements in the landscape that differentiates one landscape from another. In addition, these observed soil-geomorphology relationships can be used as a mapping tool to model ecotope characteristics and generate a model capable of relating the pedogenetic factors with the possible distribution of soils over geomorphon units within each BRU. Enhanced procedures for mapping soil-landscape relationships make it possible to spatially model and better represent ecotope regions whilst providing an associated representation of model prediction error. The quantitative prediction and representation of soil properties and soil-landscape relationships within the BRU are necessary for characterising the spatial variability of the soil ecotope inventories and understanding the coevolution of soils with TMUs.

Furthermore, introducing quantitative methods into the BRU DSS framework may facilitate expanding the original conceptual soil scientist's mental (qualitative) model of soil distribution in the landscape. Using quantitative modelling approaches, this tacit knowledge can be made explicit on maps, resulting in improvements in existing maps. Moreover, these relationships can be understood and applied by a broader range of end-users (Silva *et al.*, 2016b). In its current format, the BRU baseline recommendations may be limited in providing timely responses to rapid contextual changes and unable to fully explore the capacity of the land to support various land uses. Notably, the inclusion of predictive soil mapping methods in the BRU program involves creating and populating spatially explicit information coupled with spatial and non-spatial inferences systems (Grunwald, 2009) may position the program as a dynamic decisions software tool rather than simply an information source. For land managers and policymakers, dynamic software tools can facilitate adaptive farm management by leading them through an optimal decision path, recording data efficiently, analysing it and generating a series of evidence-based recommendations for site-specific land management

(Rossi *et al.*, 2014). This would facilitate the necessary contemporary linkages between judicious spatial land use planning and land resource inventories in KwaZulu-Natal.

This study was carried out to spatially define soil and crop ecotope regions for KwaZulu-Natal within the BRU Program using readily accessible GIS protocols. In this chapter, I do not seek to argue that the BRU program is by any means obsolete. Instead, I highlight that by taking advantage of new developments in informatics, i.e., GIS, DEMs, DSM and predictive resource-related assessment, it is expected that the inclusion of a quantitative modelling approach to ecotope mapping may provide new opportunities to assess and classify agricultural land suitability concerning specified kinds of land use and service simultaneously. Specifically, I test the usefulness of the Random Forest (RF) (Breiman *et al.*, 1984) classification and regression ensemble to characterise preliminary soil ecotope mapping units using a large source of reliable soil point data across a 73 127 ha BRU within the uThukela District Municipality. The RF algorithm, fully implemented in the ArcGIS Pro environment, is an attractive modelling technique for predicting soil classes in unmapped areas and updating soil ecotope maps in KwaZulu-Natal. It has high prediction performance, is robust to noise, has low bias and variance and is not sensitive to overfitting (Díaz-Uriarte & De Andres, 2006). The intention is to better define the relationship between soil ecotope and TMU by applying the *geomorphon* approach to stratify the landscape into ten well-defined landforms. Here I aim to qualitatively assess the relationships between geomorphon units, derived from 3-arc the second Shuttle Radar Topography Mission DEM, to better describe the spatial relationship between ecotope units and TMUs within the selected selected areas BRU which currently is limited to five landform descriptions.

5.3 MATERIALS AND METHODS

5.3.1 Study area

This study was carried out in BRU TUc6 (Colenso), located in the central region of the uThukela District Municipal (UTDM), which is located between the Kingdom of Lesotho and the western boundary of the coastal Province of KwaZulu-Natal, South Africa (**Figure 5.1**). BRU TUc6 encompasses an area of 73 127 ha between longitudes *29°34'34.92"E* and *29°52'58.32"E* and latitudes *28°41'54.32"S* and *29°02'23.52"S.* Each BRU is identified using a code based on rainfall and altitude. The upper case letters "R to Z" denote the rainfall range, and the lower case letters "a to f" denote the altitude range. BRU TUc6 indicates that the rainfall (TU) is from 650 to 750 mm p.a., the altitude (c) range is from 900 to 1400 metres above sea level. This is the 6th occurrence of the TUc code in KwaZulu-Natal. Varying climatic conditions dominate the District, with the region characterised by a combination of two climate classifications: temperate, dry winter, warm summer rainfall (Cwb) and arid, steppe, hot climates (BSh) (Conradie, 2013). Mean annual precipitation (MAP) derived from 50 years of historical climate data (Schulze & Maharaj, 2007) indicate rainfall to be 120 mm for the southern hemisphere summer rainfall period with a deviation of 69 mm to 18 mm during the winter season with a deviation of 27 mm (Smith, 2006). Typical mean annual temperature (MAT) range from approximately 10.6°C to 22.3°C, with mean maximum and minimum temperatures ranging to 27.7°C for December and June in parts, highlighting severe temperature deviations between summer and winter periods. In terms of its climate capability classification (C5), TU6c is considered to have moderate to severe limitations with moderately restrictive growing seasons due to low temperatures, frost and or moisture stress (Camp, 1995).

Elliot & Escott (2015) characterise the landscape as an assortment of gently undulating hills through to a rolling and partly broken landscape across a notable altitudinal gradient extending through the District (932 m to 1 521 m) with predominantly moderate slopes varying between 5 to 12% gradient. The region is mainly dominated by sandstone, shale, mudstone, dolerite, quartzite, dolomite, granite, diabase, and basalt saprolitic material. Moreover, altitudinal variations highly influence geological exposure, especially for the well-layered Karoo Supergroup (Elliot & Escott, 2015). The geological and geomorphological diversities are reflected in the pedodiversity of the area in terms of pedon-depth, topsoil depth, drainage, fertility and soil texture. 57.8% of the soils are considered shallow, while 38.8% are duplex soils (increased clay illuviation with depth), and 77.3% are considered to have moderate to poor drainage. Typical soil groups include, in varying spatial assemblages and without limitation: *Histosols, Gleysols, Vertisols, Umbrisols, Fluvisols, Luvisols, Lixisols, Planosols, Plinthosols, Ferralsols, Acrisols, Arenosols, Leptosols, Cambisols* and *Stagnosols*. In terms of arability, while between 21 to 26% of the BRU soils are considered suitable for either perennial or annual cropping, only about 4.7% of these soils are classified as high potential soils, i.e., humic soils typically associated with old land surfaces, young alluvial soils or well and moderately drained soils with effective rooting depths no less than 800 mm with no mechanical limitations to ploughing and clay fraction of sandy soils limited to 15% where slope gradients do not exceed 3% with conservation interventions necessary where slope gradient(s) exceed 10% and subsoil clay fraction exceeds 15% (Camp, 1995). The BRU edaphic profile is further manifested by its vegetation pattern, dominated by grassland and wooded grassland of the KwaZulu-Natal Highland Thornveld vegetation type. Finally, BRU TU6c contains five perennial and two non-perennial rivers that collectively account for 5% (3 645 ha) of the BRU area.

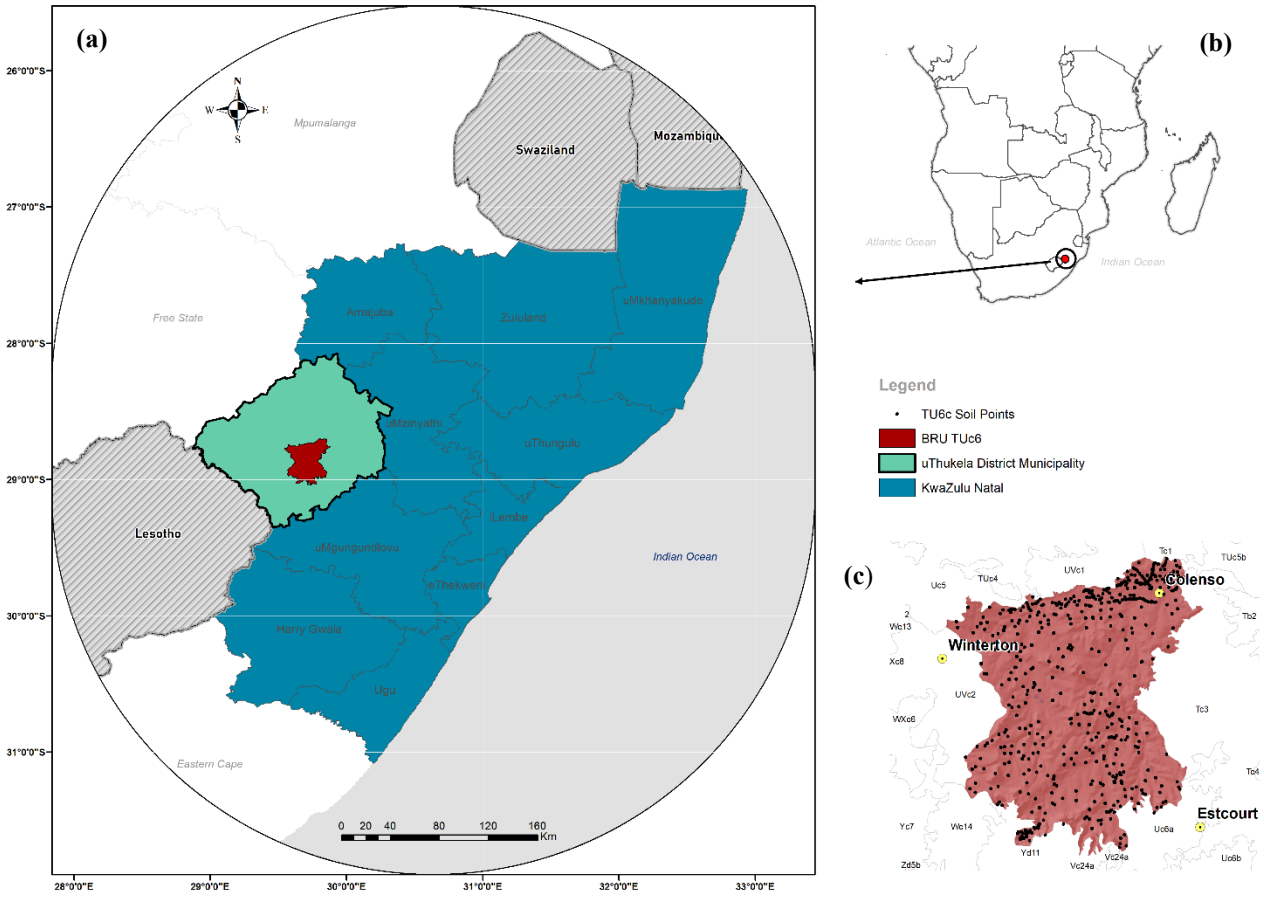


Figure 5.1: KwaZulu-Natal - (a) Location of study site, BRU TU6c situated in uThukela District Municipality (b) Continental location of BRU TU6c in Southern Africa, (c) soil point locations within BRU TU6c used to train, test and validate soil type and soil ecotope soil maps.

5.3.2 The Bioresource Report

The Bioresource Report Writer (BRW) program (Version 9), maintained by the KwaZulu-Natal Department of Agriculture and Rural Development (KZNDARD, 2015), is a desktop application with a comprehensive repository of natural resource information intended to classify and interpret various bioresources for KwaZulu-Natal delineated at a spatial scale of 1:50 000. The BRU is a demarcated area of land throughout which there are recurring patterns of soils, vegetation, climate and, to a lesser degree, terrain form (Camp, 1995). The BRW provides the platform for uniform recommendations on land use and natural resource management for regional and farm development planning (Pratt *et al.*, 1977). Besides the BRU report(s), the BRW has several unique features and capabilities not contained in any other ecosystem service support framework in current circulation in KwaZulu-Natal. These include average yield maps for over 140 well-known crops based on climate and soil ecotope, maps of soil and veld degradation, crop yield variability based on different global warming scenario's, and an application that allows users to vary their yield output predictions based on a management factor (**Figure 5.2**).

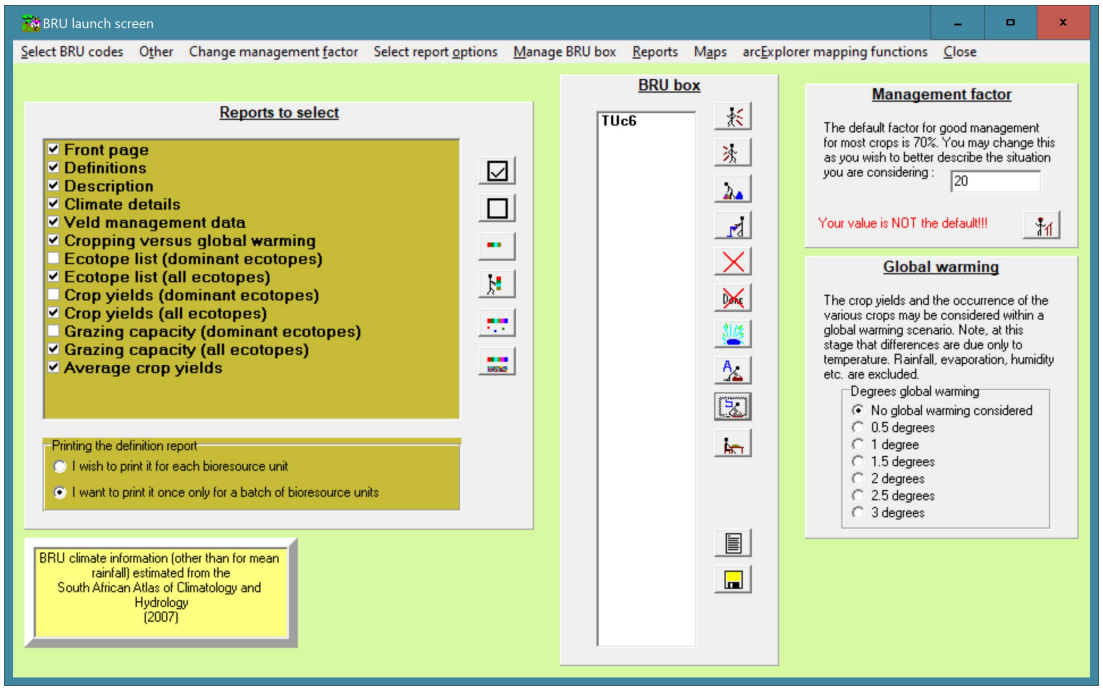


Figure 5.2: BRW application landing page. From this page, users have full access to define an array of BRU products based on user needs and preferences.

Despite this data-rich vault of information, the BRW application and related attribute inventories have not seen a significant update in sixteen years. While many immediate improvements may be gained by simply improving the desktop application and programming infrastructure, this study will instead highlight possible improvements in DEM applications to improve how the BRW displays the general description of terrain and soil-landscape characteristics for respective BRU's. While it may seem elementary to undertake such a revision, consider that for BRU TUc6, which covers 73 127 ha, the BRW provides a single description of altitude, terrain, and slope for the entire region (**Figure 5.3**), which is far too granular for site-specific descriptions.

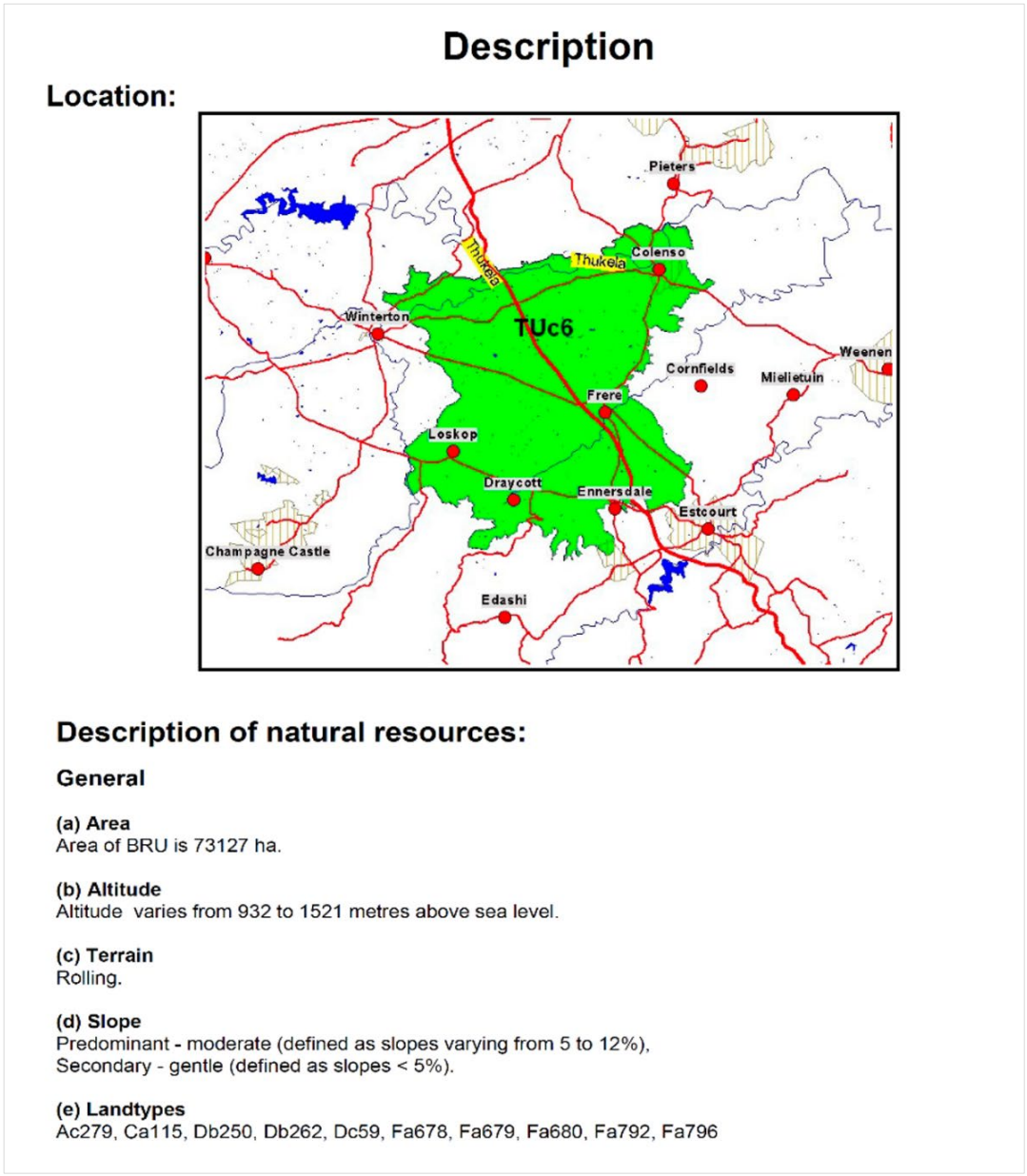


Figure 5.3: Each BRU report provides a description and overview of general site and natural resource properties. These characteristics are understood to be uniform within the entire 73 127 ha of the BRU.

5.3.3 SRTM DEM

There is an obvious benefit in providing better BRU general landscape descriptions through efficient applications of terrain datasets. In this regard, the Shuttle Radar Topographic Mission v4.1 (hereafter SRTM) project, managed by the Consortium for Spatial Information (CSI) of the Consultative Group for International Agricultural Research (CGIAR), and distributed freely by USGS, provides global topographic data at 5 deg x 5 deg tiles covering approximately 80% of the Earths terrain at 1- and 3 arc-second resolution (30 m and 90 m along the equator, respectively).

The SRTM has a reported vertical RMSE of 16 m and horizontal circular error of 20 m at a 90% confidence interval (Farr & Kobrick, 2000; Rabus *et al.*, 2003; Rodriguez *et al.*, 2005; Farr *et al.*, 2007). The SRTM data available as 3-arc second (approximately 90 m resolution) DEMs is an attractive option as an elevation source to integrate with the BRW for several reasons. First, the product remains a crucial spatial resource for many widely used terrain applications. So there is continual data improvement efforts and rigorous quality control of the SRTM products (Jarvis et al., 2008a), resulting in a highly trusted data product. Second, unlike earlier versions of the SRTM DEM (Version 2) that were known to have large areas of data voids mainly occurring in topographically steep terrain or where water or heavy shadow prevented the quantification of elevation (areas not well observed by the SRTM radar) (Sharma *et al.*, 2010), the current version is increasing better post-processed offering a seamless void-filled (pits and peaks) continuous top-of-canopy elevation surface.

Additionally, when higher resolution national or sub-national auxiliary DEMs are available, a point coverage is produced of the elevation values at the centre of each cell of the auxiliary DEM within void areas. These are used to fill large voids in the 3-arc second SRTM DEM. Finally, Atkinson *et al.* (2020) were able to show that SRTM DEM data provide a well-balanced (accuracy vs processing requirements) landscape-scale terrain product for deriving a range of DSM covariate layers for regional-scale soil-landscape applications. Their study further highlighted that depending on the terrain surface's particular soil covariate property and variability, SRTM DEM surfaces could outperform higher resolution DEM surfaces such as the ASTER Global DEM and even LiDAR surfaces generalised to lower resolution(s). For BRU TUc6, the SRTM tile numbers used for this study were SRTM3S29E029 and SRTM3S30E029 (http://srtm.csi.cgiar.org).

5.3.4 Soil and crop ecotope

The BRW program provides a range of natural resource inventory information related to ecotope. This includes dominant veld ecotope, with grazing capacity per ecotope and average grazing capacity in ha per animal unit, dominant soil ecotopes and crop yields per ecotope. This study focuses primarily on predicting the spatial variability of dominant soil ecotopes (**Figure 5.4b**), which can be related to crop yield for each dominant ecotope if necessary (**Figure 5.4c**). The soil ecotopes were developed using combinations of soil form, soil texture, effective rooting depth of the soil, slope gradient and surface rockiness (**Figure 5.4a**) (Camp, 1995). The ecotopes listed in the BRU inventories are based on information on the National Land Types, i.e., the disaggregation of land types into cropping potential land units and are considered the probable soil patterns of those units. In the BRU inventory, crops suitable for the BRU and each ecotope, and the production level (tons/ha/annum) in the case of crop ecotopes, are supplied at a stated level of management. Crop ecotopes are based on specific soil characteristics, and each of these has a symbol, and a combination of the symbols provides a code for the crop ecotopes. In addition, it is possible to investigate the potential for a wide variety of crops for which crop production models exist and for any other crop providing its growth and site requirements, particularly climate and soil, are known (MacVicar *et al.*, 1974; Schoeman & Macvicar, 1978; Camp, 1995). Based on the Taxonomic Soil Classification System for South Africa (SCWG, 1991), a crop ecotope located in TUc6 and defined as B.1.2.a.0 would be a well-drained soil with a topsoil clay percentage exceeding 35%, with an effective rooting depth between 500-800 mm located on a 0 to 3% slop with no mechanical limitations due to rocks.

Similarly, any ecotope described as B.1.2.a.0 but found in a different BRU would be expected to have the same soil factors but differ in altitude and rainfall between the respective BRUs. Moreover, other climatic factors such as heat units and evaporation are likely to be different so that the potential crop production will be

different for these two ecotopes. Incidentally, the ecotopes, both for cropping and veld, are not spatially defined but expressed as percentages of the total area of the BRU (**Figure 5.4b**).

They can only be positively identified and mapped once a soil survey has been done, incorporating local expert knowledge. A critical consideration outlined by Le Roux *et al.* (2016) is that while the ecotope concept has a sound scientific basis, they are primarily abstract since each ecotope does not have a defined location, i.e., coordinates. It is, therefore, probably wiser in the interim to consider them as 'hypothetical' ecotopes.

(a)

	Soils
A	Humic Soils (>450 mm deep) (No,Kp,Ia,Ma)
B	Well and moderately drained soils (Hu,Cv,Gf,Fw(10-22),Sd,Ct,Sp,Bv,Oa)
C	Alluvial soils (Vf,Du)
D	Mottled and moderately drained soils (Av,Pn,Gc)
E	Mottled and poorly drained soils (We,Lo,Wa,Cf,Kd(10-15)
F	Black (Margalitic) soils (Bo,Ik,My,Mw,Tk,Ar)
G	Black (Margalitic) poorly drained soils (Wo,Rg)
H	Young soils (Ms,Gs)
I	Other poorly drained soils (Ka,Kd(16-19),Fw(>30) and Wa,Lo in bottomlands)
J	Duplex soils (Sw,Va,Ss,Es,Se)
K	Organic soils and Wetlands
Y	Man made
Z	Podzol

	Rockiness
0	No mechanical limitations
1	Many stones but ploughable
2	Large stones and boulders, unploughable
3	Very shallow soil on rock or lack of soil

	Clay (%)
1	> 35%
2	15 - 35%
3	< 15%

	Depth (mm)
1	>800
2	500-800
3	300-500
4	200-300

	Slope (%)
a	0 - 3
b	4 - 4 - 12
c	13 - 15
d	16 - 20
e	21 - 40
x	> 40

(b)

Dominant soil ecotopes

(Area of ecotope must be at least 100 ha and percentage of total area at least 2%)

Ecotope	Area(ha)	% of BRU	Clay content %	Depth range (mm)	Slope range %	Fb rating	Terrain units	Soil series
B.1.2.a.0	1978.5	2.7	>= 36	501 - 800	1 - 3	5.1	3, 4, 5	Av27, Av37, Cv17, Cv27, Gf12, Hu17, Hu27, Hu37, Oa37, Oa47, Pn27, Sd11, Sd21, Sd22, Sd31
B.1.2.b.0	1729.1	2.4	>= 36	501 - 800	4 - 12	5.0	3, , 5	Av27, Av37, Cv17, Cv27, Gf12, Hu17, Hu27, Hu28, Hu37, Oa37, Oa47, Pn27, Sd11, Sd21, Sd22, Sd31

(c)

Crop yields (Dominant soil ecotopes)

Management level is 20%
The potential yield is expressed in tons/ha/annum
Ecotopes selected according to the following criterior -
Area of ecotope must be at least 100 ha and percentage of total area at least 2%

	Cabbage transplanted in January	Cabbage transplanted in February	Cabbage transplanted in March	Cabbage transplanted in April	Cabbage transplanted in May	Cabbage transplanted in June
B.1.2	15.9	17.5	17.7	18.4	18.2	12.5
B.2.1	17.7	19.5	19.7	20.5	20.2	13.9
B.2.2	16.8	18.5	18.7	19.5	19.2	13.2

	Cabbage transplanted in July	Cabbage transplanted in August	Cabbage transplanted in September	Canola	Carrot - hybrid sow September	Carrot - open pollinated sow September
B.1.2	16.3	21.0	23.5	0.8	16.6	11.5
B.2.1	18.1	23.4	26.1	0.9	23.7	16.5
B.2.2	17.2	22.2	24.8	0.8	21.3	14.8

Figure 5.4: (a) Soil ecotope properties. Ecotope soils are based on the South African Taxonomic System (1991-2019) for soil classification. (b) Description of selected dominant soil ecotopes and linkages to LTS terrain unit description and soil series description. (c) Description of selected crop yields based on dominant soil ecotopes. The crop yield(s) t/ha/yr, is based on a user-defined management factor. This example is defined as 20%.

5.3.5 Generating primary terrain attributes

Developed by Jasiewicz & Stepinski (2013), the geomorphon approach for terrain unit classification is designed to be computationally efficient, flexible to scale, and adept to DEM slope position. Using a pattern recognition algorithm based on a 3 x 3 local neighbourhood search radius from a central pixel focal point, the model can identify the ten most common landform units within the landscape, namely, flat, peak, ridge, shoulder, spur, slope, footslope, hollow, valley and pits. Several recent studies have highlighted the application of geomorphons for TMU mapping under local conditions. Flynn *et al.* (2019a) combined geomorphons with a technique known as "Disaggregating and Harmonising of Soil Map Units Through Resampled Classification Trees" (DSMART)(Odgers *et al.*, 2014) to disaggregate the LTS into a farm-scale soil map at a resolution suitable for farming applications. Atkinson *et al.* (2020) demonstrated how DEM spatial variations and geomorphon parameterisation influence the geomorphic terrain characterisation for soil-landscape feature extraction from three different derived DEM models calculated at two distinctive and operationally relevant, pixel resolutions, 30 m and 90 m within the uThukela region.

For creating the 2D geomorphon model in this study, the SRTM DEM was input to GRASS GIS (v 7.4.1)(GRASS Development Team, 2016) with the workflow using the *r.geomorphon* extension and the System for Automated Geoscientific Analyses SAGA v2.3.2 (Conrad *et al.*, 2015), respectively. A central feature in the generation of geomorphic surfaces is the *a posteriori* optimisation of three primary model parameters: maximum search radius (lookup distance), flatness threshold (t-degrees) and Skip radius (cells). A further consideration is the search radius (L-cells), representing the maximum distance for each pixel's line-of-sight (*LOS*) calculations. By adjusting these parameters, users can effectively skip minor terrain variations to capture more massive landforms and uniformly optimise the final geomorphon surface outputs related to the resolution of the DEM. In this study, I have adopted a similar approach to Luo & Liu (2018) and Atkinson *et al.* (2020) by iteratively selecting model parameters that offered a reasonable balance between capturing landform elements' accuracy and model computational cost. After several training iterations of the *r.geomorphon* model, with each model iteration characterised by a set of specific parameter combinations, I concluded that for the UTDM region, the following optimisation parameters best represented the geomorphon surfaces across the 90 m SRTM DEM: outer search radius =2 700 m, inner search radius = 5 40 m and flatness threshold = 1.2^{o}.

Slope gradient, represented as percentage values for this study, was produced at 90 m resolution from the SRTM void-filled 3-arc second DEM using the *Slope* tool in ArcGIS (ArcPro Version 2.8.0)(ESRI, 2021) *Spatial Analyst* suite of surface tools. Aggregated slope gradient values were derived for geomorphon features, which would further be used to segment the derived spatial soil ecotope boundaries to provide a better BRU site description for inclusion in the BRW. The intention is to provide a series of descriptive statistics, with the possible inclusion of terrain resource maps that focus on detailing the occurrence and frequency of slope gradient and geomorphon features within each BRU unit. It is expected that the inclusion of detailed slope and TMU properties in the BRU will undoubtedly provide a far more comprehensive description of terrain, slope and land types regardless of BRU extent.

5.3.6 Modelling soil units with soil profile data

For this study, three primary sources of soil information were used to analyse and validate soil ecotope properties within BRU TUc6 (**Figure 5.5**): the Land Type soil profile database, the Provincial soil database of agroecological assessments obtained from the KwaZulu-Natal Department of Agriculture and Rural Development (KZNDARD), and finally the soil points assimilated from the Soils of the Tugela Basin study (TBS). Each dataset is hereafter briefly described. The soil profile database (SPD) accompanying the LTS of South Africa, with original field mapping at a spatial scale of 1:50 000 and eventual production mapping by the Agricultural Research Council's Institute for Soil, Climate and Water at 1:250 000 (ARC, 1991; ARC, 2003) was used for this study. The SPD contains around 2 500 modal soil profiles and a further 10 000 soil series identification samples to support field soil diagnosis providing quantitative data about a range of soil properties across the greater part of South Africa (Paterson *et al.*, 2015). Typical soil information includes sample location, usually with several samples per location at various depths, information on soil organic carbon content, as well as fractions of the textural components of clay, sand and silt and often a broadly classified land use (Schulze & Schütte, 2020). The more than 6 000 Land Types in the LTS, from which the BRU spatial units are heavily designed on, are well represented by the SPD data (approximately one observation per 300 ha) with the areas and percentages of each soil series given in the documentation on each Land Type and extensively interrogated.

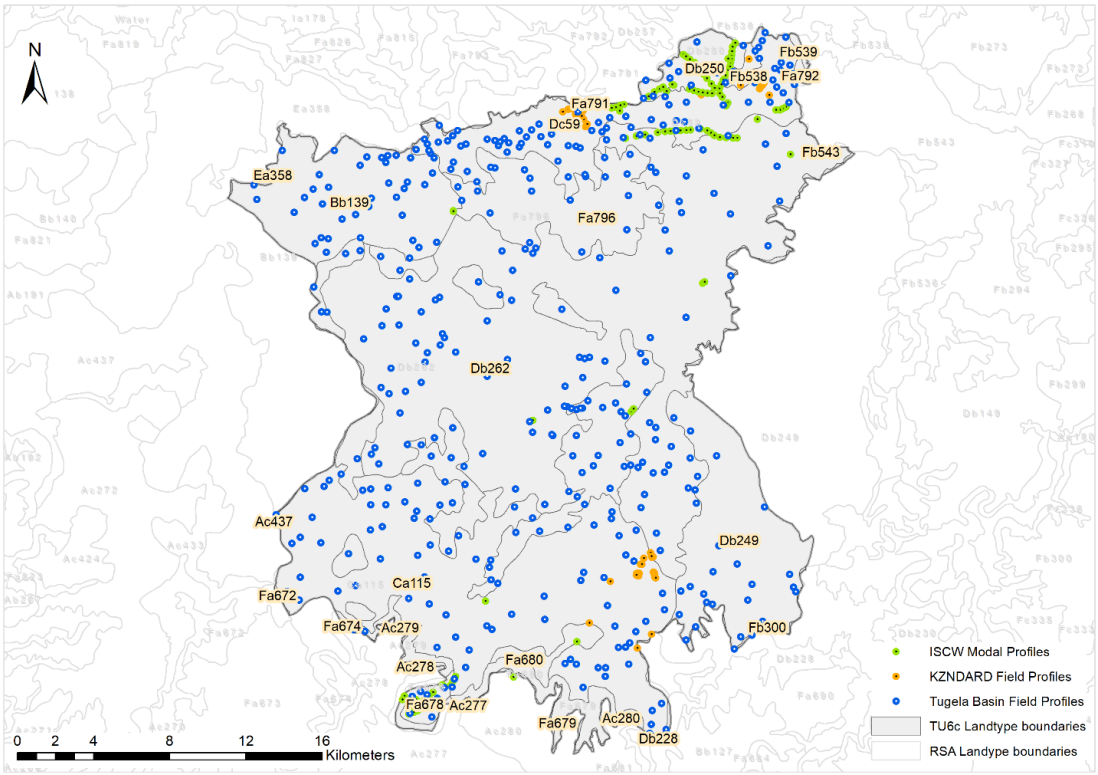


Figure 5.5: Detailed overview of soil sample locations in BRU TUc6 intersected with Land Type boundaries. Soil sample coverage is well represented by the combined contribution of ISCW, KwaZulu-Natal DARD and Tugela Basin soil data. A total of 25 Land Type Units are recognised in BRU TUc6.

Van den Bergh *et al.* (2009) used the land type point database to develop spatial methodologies for semi-detailed soil mapping of the KwaZulu-Natal Province. Wiese *et al.* (2016) leveraged the SPD with field observational data to describe soil carbon using vertical distribution functions and kriging. Stalz (2007) mapped salt-affected soils in the Orange River irrigation scheme by mapping salt-affected plants using remote sensing. A thorough review of the Land Type mapping procedures can be found in Paterson *et al.* (2015) and Van Zijl & Botha (2016).

The second source of soil point information is derived from the KZNDRD's Natural Resource Sub-Directorates (NRS) field assessment(s) as part of their efforts to characterize the natural resources, optimize agricultural potential and enhance productivity through sustainable natural resource management in KwaZulu-Natal (Mitchell, 2020). The NRS is mandated to identify, survey and research current land use, soil types, veld condition and land degradation in various disciplines (soil, veld, climate and hydrology) in the province on a scientific basis. Moreover, the Unit is tasked with maintaining the BRU Programme by compiling additional resource data, improving data accuracy, and determining agricultural potential and sustainability for optimised land use planning. Therefore, soil data is primarily surveyed at farm level with soil profile attributes reported based on expert knowledge using diagnostic and taxonomic classifications. Considering the scale of data collection, i.e., typically 1:10 000, these points are highly clustered but provide a helpful solution as potential independent model validation points for the final RF ecotope predictions.

The final soil dataset used is provided by the semi-detailed survey of the TBS (Van der Eyk *et al.*, 1969). Atkinson *et al.* (2020) provide a useful summary of the genesis of the TBS, including critical applications and limitations of the legacy soil dataset. In the context of this study, the spatial soil data still offers a suitable contribution considering the degree of detail captured for a wide range of soil profile attributes. This is highlighted because more than 22 000 soil profiles were examined in detail across the 31 500 km^2 Tugela Basin with locational positions hand-recorded on hard-copy photo-mosaics and topographical maps during the study. More than 250 modal profiles and 819 samples were collected and analysed to determine particle size analysis, soil reaction (pH, KCL), exchangeable cations, and soil organic carbon for each soil pit. While the survey's purpose was not to provide a blueprint for detailed land use planning for the region, it remains relevant because it is still well aligned to the LTS protocol(s). This allows the soils of the Tugela Basin to be interoperable at a regional rather than national scale while revealing the broad zonal arrangements in the soil series distribution, with each spatial unit potentially comprising as many as several associated series (Van der Eyk et al., 1969). **Figure 5.6** provides a valuable overview of the adopted workflow detailing the soil attributes extracted and processed for each soil dataset.

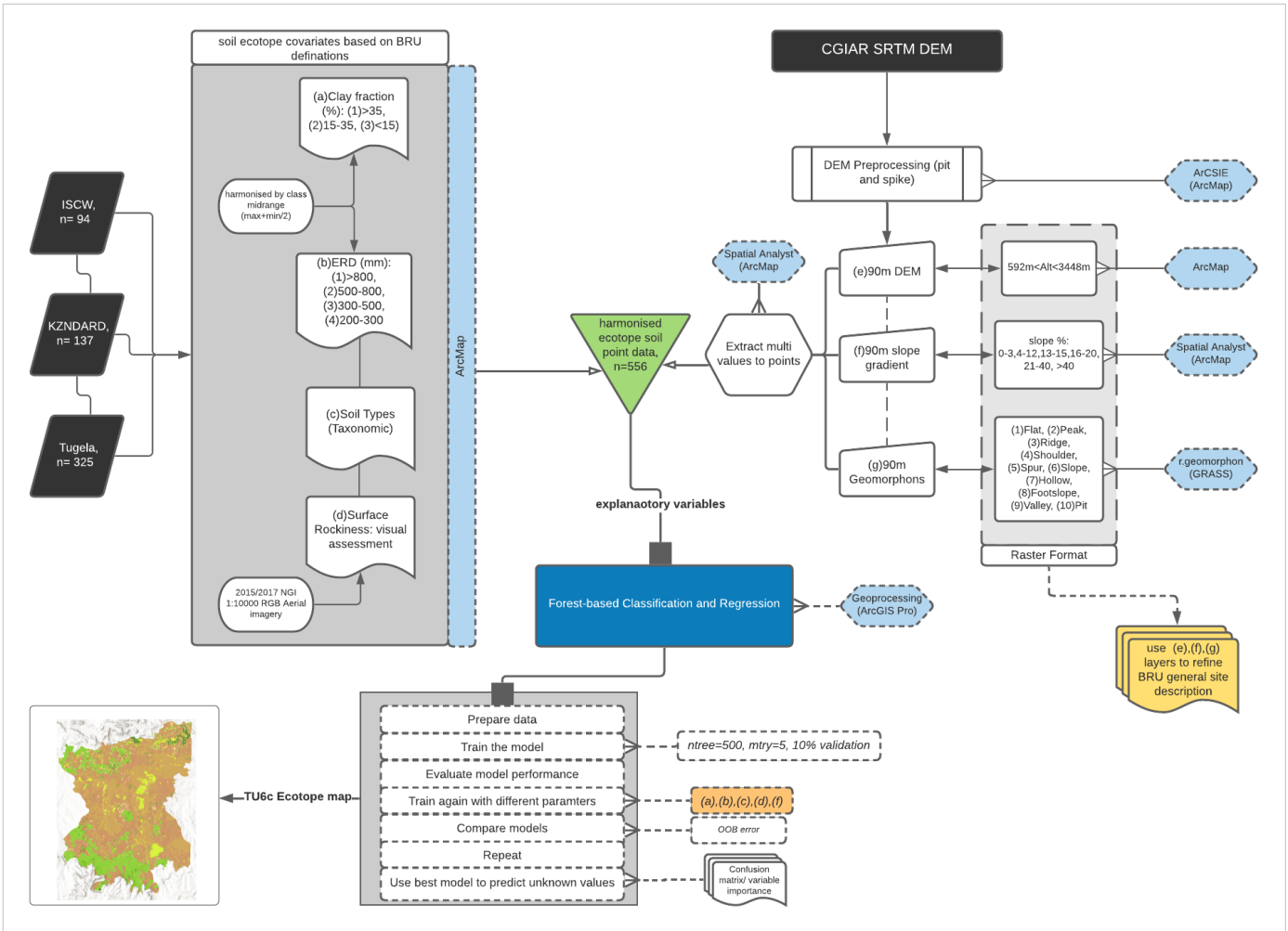


Figure 5.6: ArcGIS Pro Random Forest conceptual workflow model for digitally mapping soil ecotope and soil type in BRU TUc6.

5.3.7 Application of Random Forest classification for modelling soil units

Classification of the soil type and soil ecotopes was performed with random forest (RF) (Breiman, 2001) implemented in ArcGIS Pro Forest-based Classification and Regression GeoAnalytic Tools suite of applications. The RF classifier is conceptually similar to a decision tree. Specifically, it is a non-parametric ensemble method wherein many different classification trees are aggregated to produce a more stable and accurate classification than a single decision tree (Cutler & Edwards Jr, 2007). Each tree is built on a bootstrap sample (e.g. 500) of the given data. In a typical bootstrap sample, approximately 63% of the original observations occur at least once. Observations in the original data set that do not occur in a bootstrap sample are called out-of-bag (OOB) observations. To construct the ensemble, the different trees are combined using bagging (bootstrap aggregating). The resulting “forest” is a “random” forest because, at each split, only a random subset of the candidate predictors is considered for the binary partition (Häring & Dietz, 2012). The trees are fully grown, and each tree is used to predict the OOB observations.

The predicted class of observation, i.e., ecotope, is calculated by a majority vote of the OOB predictions for that observation, with ties split randomly. This de-correlates the trees, improves the variance reduction and finally leads to more accurate predictions (Bühlmann & Yu, 2002; Strobl & Malley, 2009). Accuracies and error rates are computed for each observation using the out-of-bag predictions and averaged over all observations. Because the OOB observations are not included in the fitting of the ensemble, the OOB estimates can be considered valuable cross-validated accuracy estimates (Cutler et al., 2007). RF has similar advantages to decision trees that for this study provide several benefits to modelling soil ecotopes within BRUs: the ability to model high dimensional non-linear relationships with only a few user-defined parameters, handling of categorical and continuous predictors, insensitive to missing data (Heung & Bulmer, 2014) and the prediction of OOB data avoids overfitting (Lagomarsino & Tofani, 2017). RF's additional characteristics include its ability to provide variable importance measures and relative robustness concerning noise in training data (Hua & Xiong, 2005).

Implementing the RF model in ArcPro is a relatively recent development and offers significant interoperability with existing geospatial and geostatistical applications. A significant advantage of using the ArcPro RF model, compared to other non-spatial statistical platforms, is that the tool can be used in two operation modes. First, the *Train, a model, to assess model performance* option can be used to evaluate the performance of different models to explore different explanatory variables and model optimisation settings. After a suitable model has been identified, users may use the *Fit a model and predict to features* or *Predict to raster values* option to calculate various outputs, i.e., output trained features, output predict features, output variable importance table, and statistical summary report results. Notably, the ArcPro RF models do not extrapolate; they can only classify or predict a value that the model was trained on. Therefore, models must be trained using training features and explanatory variables within the range of target features and variables, i.e., soil and ecotope properties. The tool will fail if categories exist in the prediction explanatory variables but are absent in the training features (ESRI, 2021). To optimize RF application in ArcPro, it was determined that the two default parameters for the RF model: (1) the number of predictors to be used in each tree building process (*mtry*) and (2) the number of trees to be built in the forest (*ntree*) selected by the program were suitable for meeting the research objectives and did not require further optimisation.

5.4 RESULTS AND DISCUSSION

5.4.1 Updated terrain morphological units and slope gradient classification

The 90 m geomorphon classifications of terrain morphology resulted in all ten geomorphon features well represented across BRU TUc6. Furthermore, the landscape segmentation into distinctive TMUs, with associated area coverage recorded per unit, represents a significant improvement to the existing general terrain description outlined in the BRU Report (**Figure 5.3**). The geomorphon classification results have also been summarised to represent the total percentage coverage of each TMU. Interestingly, while slope features appear to occupy a sizeable single-unit coverage of 34%, it would appear that BRU TUc6 is comprised of a series of ridges, spurs and valley bottoms that collectively account for 43% of the terrain area. Incidentally, this would suggest that the general terrain morphology is a series of ridges and valley bottoms and, in essence, as indicated in the BRU Report, best described as "rolling" in nature. Therefore, the geomorphons validate the BRU descriptions and, by introducing a spatial overview of the TMUs within the BRU, can provide strategic and regional precision representation(s) of terrain morphology. The BRU report does not provide reconnaissance terrain feature coverage (extent) or regionality (location) and, therefore, limits the ability of users to optimise their land use planning strategically. If terrain feature coverage map(s) were included in the BRU Report, it would allow users to notice areas with typically limited lower agricultural potential, for instance, summits, ridges, hollows, and pits. However, more importantly, users can verify the extent of these features. The same principle applies to the spatial representation of slope gradient, which is not well defined at a regional level within the BRW program. **Figure 5.7** highlights that while a series of undulating terrain properties characterise the region, the predominant slope gradient for BRU TUc6 is classified as less than 15%. This is well below the legislated 20% slope gradient for the cultivation of non-perennial crops (DAFF, 1983). As a first approximation, by simply assessing the slope gradient variability, land users can begin to draw estimations of land capability for the region. A slope gradient of 15% or less could be considered arable, i.e., land capable of being cultivated and suitable for the production of crops, but with slight boundaries and low erosion risks, or severe boundaries, with low arable potential and high erosion risks (Camp, 1995; KZNDARD, 2015).

Furthermore, regions where the slope gradient does not exceed 8% unless under certain crops such as sugarcane or pastures would be suitable for irrigation depending on the irrigability of the soils (Smith, 2006; Schulze *et al.*, 2007). Concomitantly, the regions identified as having slope gradient exceeding 15%, while non-arable, may still facilitate some form of perennial veld and pasture cultivation and support wildlife through game farming. Undoubtedly, the ability to spatially align land-use practices with environmental limitations provides a far superior approach to sustainable land and resource management, particularly within the BRU. Combing the geomorphon classification with the slope gradient results provides additional detail regarding the frequency of TMUs occurrence within each slope class (**Figure 5.8**). The results highlight that all ten geomorphon features occur within slope classes 0 to 3%, occupying just over 62% of the total BRU with approximately 35% BRU area comprised of geomorphons within 4 to 12% slope gradient. Put differently, approximately 97% of the BRU area contains all geomorphon features below 12% slope gradient confirming the undulating terrain characteristics described in the BRU report. Therefore, the remaining 3% of the BRU area is occupied by geomorphons exceeding a 12% slope gradient and concentrated predominantly along the southernmost border of the BRU. If analyses such as these are included in the BRU report, end-users will have a richer perspective of regional terrain morphology to facilitate better decision-support planning.

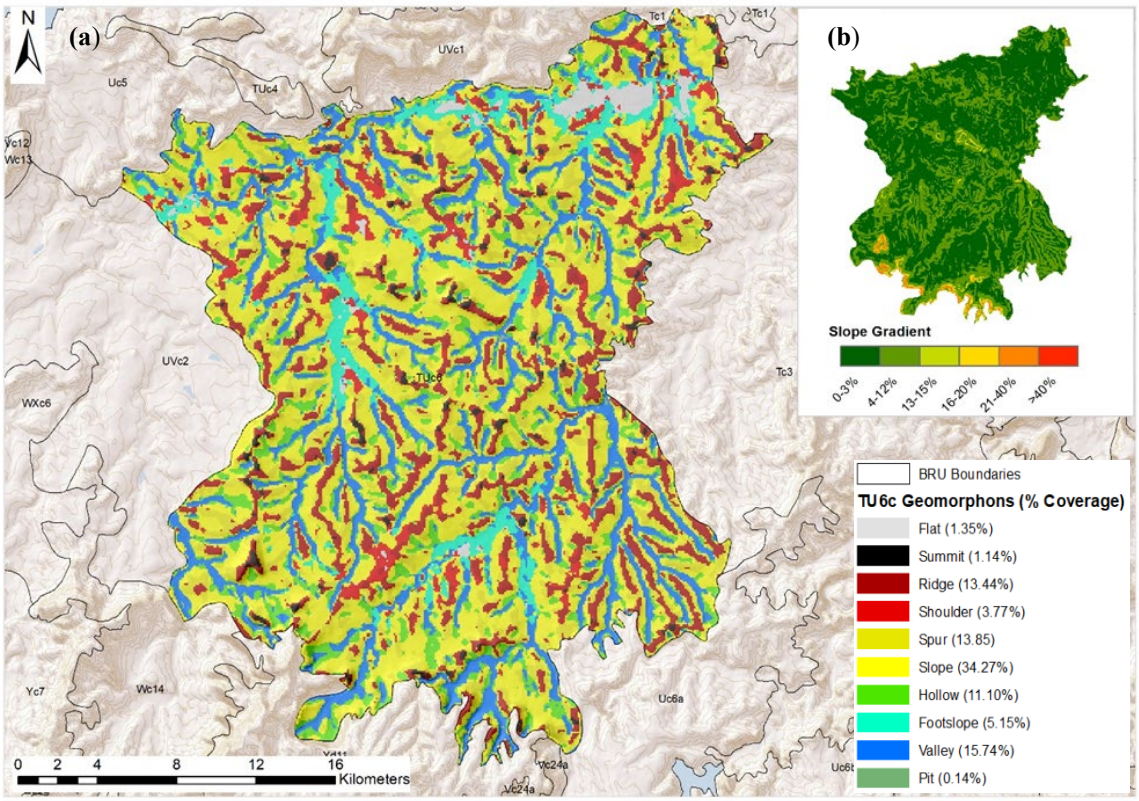


Figure 5.7: (a)Map of BRU TUc6 characterised by ten modelled geomorphon features against backdrop of LTS boundaries and (b) inset map of slope gradient percentage.

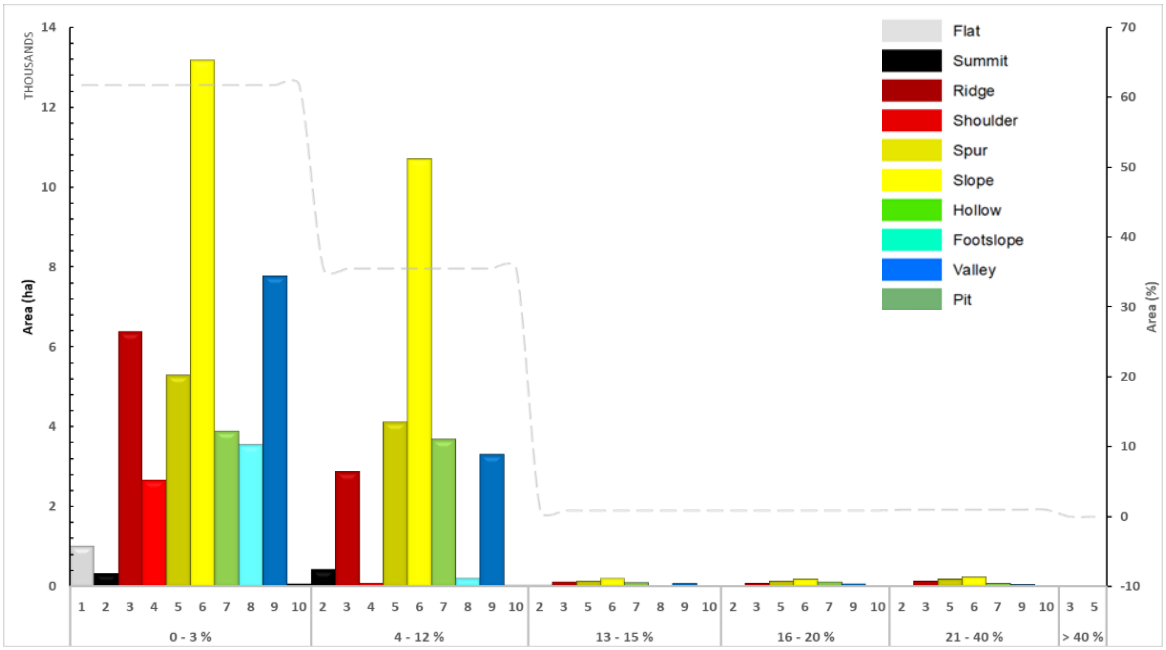


Figure 5.8: Histogram of geomorphon features showing the morphon feature extent in ha (y_1-axis) segmented by slope class category in per cent (x-axis) with cumulative morphon feature coverage per slope category in per cent (y_2-axis).

5.4.2 Mapping soil types and soil ecotopes using forest-based classification and regression

5.4.2.1 RF model characteristics

The forest-based classification model (RF) was extrapolated to 95% (69 660 ha) of the TUc6 area. The regions where no predictions were possible (3 467 ha) were limited to the areas where no soil samples were available and coincided with the peripheral portions of the BRU. RF models do not extrapolate; they can only classify or predict a value on which the model was trained. Therefore, models must be trained using training features and explanatory variables within the target features and variable's spatial location. The RF model characteristics are presented in **Table 5.1** and contain the basic model parameters for both the soil ecotope and soil type prediction models. A helpful approach to implementing the RF model in ArcPro is selecting *Advanced Foret Options*, which then applies a data-driven approach to derive optimal forest characteristics and improve model performance. This explains why both models have the same number of trees, i.e., 100 trees, but their respective tree depth range and mean tree depth values differ. The tree depth range shows the minimum tree depth found in the forest and maximum tree depth found in the forest. If the maximum depth parameter is set to 100, but the range and mean depth of the data suggest that a much smaller depth is used most of the time, then decreasing the maximum depth parameter can enhance the model's performance since it decreases the likelihood of overfitting the model to the data (ESRI, 2021). While 100 trees were selected as default values for both models, the soil type and soil ecotope predictions achieved optimal model performance with 45 and 69 trees, respectively. Each tree in the forest is assigned a unique number of randomly selected variables reported in the *Number of Randomly Sampled Variables* option. The default number is determined by combining the number of features and the number of variables available. For regression, it is one-third of the total number of explanatory variables.

Table 5.1: RF model characteristics for soil type and soil ecotope

	Soil Type	Soil Ecotope
Number of Trees	100	100
Leaf Size	1	1
Tree Depth Range	34-61	59-58
Mean Tree Depth	45	69
% of Training Available per Tree	100	100
Number of Randomly Sampled Variables	2	2
% of Training Data Excluded for Validation	10	10

5.4.2.2 RF model performance

The RF model predicted all seven soil types described in the soil database to map soil type (**Figure 5.10**) and all twenty-seven ecotope classes (**Figure 5.11**). The model performance is measured by the OOB error, calculated as the mean squared error (MSE) constructed on the model's ability to predict soil type classes based on the observed predictor values in the training dataset. Recall that Each tree of the RF is s constructed using bootstrap sampling. A different bootstrap sample is taken for each tree, which is approximately 2/3 of the available observations. OOB observations are those not included in the bootstrap sample. The proportion of misclassified samples in these samples (OOB error) can be used to measure generalisation error, otherwise referred to as model misclassification rate (Peters *et al.*, 2007). The RF model in ArcPro calculates OOB errors based on the percentage of incorrect classifications for each category among trees that did not see a subset of the trees in the forest. Since all input variables were available as 90 m resolution raster formats, the resulting model was then used to predict soil type spatially and soil ecotopes at 90 m for BRU TUc6.

The MSE of the classifications for soil type predictions (**Table 5.2**) and soil ecotope (**Table 5.3**) can be interpreted as the overall proportion of incorrect OOB classifications among all modelled categories. Firstly, our results highlight a clear difference in ensemble *ntree* performance. The 100-tree iteration show lower model OOB error and, therefore, better classification performance. This is evident by the lower MSE values for the 100 tree models, which translates into an OOB accuracy classification accuracy (1 – OOB error) of 76.2% and 69.7% for the soil type and ecotope predictions respectively. The continuity of model trends between the soil type and soil ecotope model prediction is a further promising result. That is, soil types identified as being either well or poorly modelled showed similar trends in ecotope prediction performance, albeit within different orders of magnitude. For instance, soil type and ecotope for categories D, G, H and J showed lower prediction performance in their respective models compared to the remaining classes. The corollary is that the remaining categories show better model predictions in both soil type and ecotope predictions. Generally, soil data with higher, balanced sampling frequencies tend to have higher prediction accuracy. Higher sample frequencies allow more OOB samples to be passed down the RF model tree and classified. As tree numbers increase in the forest, the OOB error can better estimate the classification error. Although an inverse relationship between predicted soil class sampling frequency and OOB error often exists (Brungard, 2009; Reza Pahlavan Rad *et al.*, 2014), a positive relationship with relatively low OOB error with low sampling frequencies were observed in this research. This is not uncommon in unbalanced datasets, as in this study, resulting in upwardly bias OOB estimates. There is extensive discussion in the literature regarding the influence of *mtry* optimisation on model performance. Cutler *et al.* (2007) reported that different values of *mtry* had no impact on the accurate classification rates of their model. Other performance metrics (sensitivity, specificity, kappa, and ROC AUC) remained stable while Brungard (2009) showed that overall OOB errors were remarkably constant using 1, 4, 8, 12, 16, 20 and 22 *mtry* values.

On the other hand, Strobl *et al.* (2008) reported that *mtry* strongly influenced predictor variable importance estimates. In fact, Heung *et al.* (2014) reported that lower OOB error rates could be expected when using fewer training points to build an RF model. Their study evaluated the optimisation of *mtry* values ranging from 1 to 27 and assessed the OOB error rates from 50 replicates for each *mtry* value. Interestingly, they reported that smaller *mtry* values retained more 'randomness' in RF's randomised variable selection process.

In this study, the selected ArcPro default setting of the primary tuning parameter, *mtry* value, which is the square root of the total number of predictors (M), is a good choice since a total of only five predictors, i.e., clay fraction, soil depth, surface rockiness, slope gradient and altitude, were available for splitting at each tree

node. Indeed, the inclusion of additional predictor variables to the RF model in future iterations may necessitate further optimising the *mtry* parameters. Breiman (2001) recommends that several covariate subsets be tried, and the size that returns the lowest overall OOB error is chosen. Incidentally, due to the conflicting evidence reported in the literature, the I recommend selecting the default value(s) within the ArcPro RF application with optimisations applied *ex-ante* based on what value might be suitable for the RF ensemble model. Second, comparing the observed OOB error with the frequency of soil type and ecotope sampling at each level could indicate that these series displayed strong relationships with predictor covariates and thus low OOB error. Hence, the OOB error could be reduced by increasing sample size and identifying solid relationships between individual soil properties and covariates used to predict soil types (Heung *et al.*, 2014). The ArcPro RF model explained between 69.4% and 99.3% of soil type variability and 41.8 to 99.1% of soil ecotope variability for various classes. Based on the extent of the survey area, the sampling intervals of the three different soil databases (~ 132 points/ ha), and the spatial resolution of the soil prediction covariates (90 m), the results are more than satisfactory. In mapping World Reference Base (WRB) Reference Soil Groups, Barthold *et al.* (2013) achieved an error rate of 51.6% OOB. Even with 672 soil observations, Stum *et al.* (2010) found an OOB error of 55.2% in predicting soil series. Häring *et al.* (2012) saw an OOB error range from 0.05 to 0.55 with a median of 0.31 for 9 924 soil profiles.

Table 5.2: Model out of Bag errors for soil type prediction with 50 and 100 tree ensemble

Number of trees	Frequency	50	100
MSE		26.01	23.83
B	41	20.67	16.14
D	80	27.11	24.02
E	30	2.64	1.34
F	21	1.47	0.75
G	63	27.37	22.79
H	161	27.23	26.15
J	160	32.35	30.65

Finally, Brungard (2009) recommends exploring the utility of other variations of RF models, including balanced random forests and balanced and weighted random forests. These models may be better suited where one or more soil types or classes are much more common across a landscape and where observation numbers are highly imbalanced (one or several classes) is much larger than the others). In balanced and balanced and weighted random forests, the objective is to lower error rates for small classes by equalizing class errors (Breiman, 2001). The balanced errors are achieved at the expense of a higher class error and a higher overall OOB error. These two models are not yet offered in the ArcPro Forest-based regression and classification toolset; however, readers are referred to Chen *et al.* (2004) and Brungard (2009) for a detailed review of their respective applications.

Table 5.3: Model out of Bag errors for soil ecotope prediction with 50 and 100 tree ensemble

Number of trees	50	100
MSE	33.60	30.30
B110	1.17	0.60
B111	12.41	6.33
B112	1.17	0.60
B122	2.28	1.16
B210	6.53	3.33
D110	22.47	17.92
D111	49.46	44.45
D112	35.29	32.92
E240	10.78	10.40
E322	3.92	2.00
E340	0.28	0.14
F120	13.92	7.10
F212	0.39	0.20
G120	9.36	4.77
G121	29.63	20.94
G122	39.58	34.47
H122	16.99	8.66
H223	2.74	1.40
H240	1.56	0.80
H343	40.28	39.06
J120	14.37	11.44
J121	4.13	2.11
J122	25.49	20.07
J220	42.81	35.88
J221	55.35	52.51
J222	62.17	58.02
J223	0.65	0.33

5.4.2.3 Spatial prediction of the most probable soil types and ecotopes

Independently, both sets of digital soil maps represent a significant contribution to the existing BRU Report ecotope module. Including a BRU soil type map into the BRW may improve regional and farm-level planning, particularly in rural areas where new agronomic development initiatives are prioritised. A synoptic representation of general soil types and associated properties will allow the user to manage and coordinate their land use more effectively by prioritising agricultural outputs in regions characterised by productive soil types and supportive of medium to high agronomic potential. Presently, the BRU report relies on the LTS database to determine generalised soil properties with detailed soil and landscape characteristics derived *ex-ante* from field assessments and then used to derive soil ecotopes. Even with the rising costs of conducting detailed site and soil surveys, combined with the time-consuming reliance on expert judgement to develop site-specific soil maps, portions of the Province are well mapped. Since many of these detailed soil and resource maps are not easily accessible in spatial formats or interoperable with other sub-national geospatial and RS data, strategically managing these resources, top-down does not promote integrated natural resource management (Paterson *et al.*, 2015; Van Zijl & Botha, 2016; van Zijl, 2019). Users would benefit directly from having access to summarised soil coverage data detailed for each BRU as a first-tier evaluation. Including

summary data presented in **Table 5.4** would give users a valuable perspective of each soil type, their extent and alignment with the WRB descriptions enabling better alignment of strategic land use objectives and expectations with separate soil limitations.

Table 5.4: Mapped soil type units for TUc6 comparing modelled extent for each class with Land Type coverage reported in the BRW.

Soil Type	WRB Reference Group	Modelled Extent (ha)	Modelled Coverage (%)	Land Type Coverage (%)
B - Well and moderately drained soils	Ferralsols; Cambisols; Arenosols; Acrisols; Lixisols; Nitisols	2 114.4	3	21.5
D - Mottled and moderately drained soils	Plinthosols; Ferralsols; Arenosols; Acrisols; Lixisols; Cambisol	7 468.8	11	0.6
E - Mottled and poorly drained soils	Stagnosols; Plinthosols; Arenosols; Acrisols; Lixisols	4 156.8	6	14.2
F - Black (Margalitic) soils	Chernozerms; Luvisols; Leptosols; Fluvisols; Phaeozems	4 882.4	7	5.1
G - Black (Margalitic) poorly drained soils	Gleyisols; Vertisols; Umbrisols	8 604.0	12	0.6
H - Young soils	Leptosols; Acrisols; Lixisols; Cambisols	16 822.4	24	14.2
J - Duplex soils	Solonetz; Planosols; Luvisols; Lixisols	25 611.2	37	38.8
Total		**69 660**	**100**	**95**

The final spatial predictions of soil type and soil ecotope for TUc6 are presented in s **Figures 5.10** and **5.11**, and associated classification diagnostics are presented in **Table 5.5** and **Table 5.6**, respectively. The results show that TUc6 is dominated by duplex soils followed by young, shallow soils, accounting for 61% of the modelled BRU area. These soil types, including those which may be poorly drained, require significantly higher management and conservation considerations to maintain or even maximise agricultural output(s). Furthermore, the results suggest that higher agronomic priority soils (B, D and F) collectively only occupy approximately 21% of the BRU. It has been mentioned earlier that due to the computational limitations of the RF model, it was not possible to model soil type for the full extent of TUc6. Therefore direct comparisons with the results in the BRW are not possible. However, a significant outcome of the spatial prediction is that the modelled soil type coverage is well aligned to the current soil pattern reported for the BRU and based on the LTS soil properties. A significant advantage of running the RF model in ArcPro is is that these summary model performance results can easily be embedded directly into the BRW for each spatial prediction map. These results show that based on a randomly excluded training and validation subset, the RF model performance ranged from 76% to 91%, depending on the modelled soil type and 92% to 98% for modelled soil ecotopes. For validation purposes, 10 percent of the training data is excluded by default within the ArcPro RF model (ESRI, 2021). The model is then used to predict the values for the test data, and the prediction accuracy of the model is measured in comparison to the observed values using data that was not used during the training process.

Table 5.4 highlights the comparison between modelled soil type pattern and reported soil type pattern for TUc6. The highest association appears most evident for duplex soils, where modelled coverage was estimated to be 37% compared to 38.8% reported in the BRW. Other notable associations include Black (Margalitic) soils estimated to occupy 7% of the BRU against a reported 5.1%. A favourable result must be that all soil types and soil ecotopes reported overall accuracy that far exceeded expectations compared to similar regional DSM applications. Van Zijl *et al.* (2013) produced soil association maps for an area near Newcastle in KwaZulu-Natal with 30 to 67% accuracy. Flynn *et al.* (2019b) evaluated the utility of the DSMART algorithm (Odgers *et al.*, 2014) to map five broad soil association classes at Cathedral Peak in eastern KwaZulu-Natal and obtained accuracies ranging from 35 to 93%. In fact, DSM results reported in this chapter can improve the accepted prediction accuracies of 65% expected from traditional soil maps quoted by Marsman & de Gruijter (1986).

A discernible observation in both prediction maps, which is more evident in the soil ecotope surface, is the apparent "feature-noise" in the north-east regions of the TUc6 compared to the more generalised, smooth feature boundaries in the central parts of the BRU. Additionally, these same areas were discovered to have lower classification accuracy, < 5%, than the surrounding areas (**Figures 5.10b & 5.11b**). The primary sources of sensitivity of the RF model in these regions can be attributed to three possible causes: first, there are definite regions within TUc6 that display systematic patterns of medium to high spatial variability related to each model covariate. Systematic variation is a gradual or marked change in soil properties as a function of landforms, geomorphic elements, soil-forming factors or soil management (Jenny, 1941). Here spatial variability is understood to constitute pedogenic state factors contributing to different soil-forming environments and consequent spatial variability of soil morphologies (Wilding *et al.*, 1994; Florinsky *et al.*, 2002). Variability in the spatial distribution of soil input data can strongly influence the quality of results from a logical, empirical, and physical model of soil and landscape processes (Burrough, 1993). Lin *et al.* (2005) interpret soil variability as influenced by a combination of soil-forming factors acting through space and time. Specifically, I identify five space-time factors as functions of soil variability, namely spatial extent or area size, spatial resolution or map scale, spatial location or physiographic region, specific soil property or process, and finally, a temporal factor.

The exact expression of these factor-relationships is complex and beyond the scope of this research. However, it is accepted that as spatial extent, spatial resolution, or time scale increase, the magnitude of soil variability will increase. **Figure 5.9** displays the spatial variability, represented by standard deviation, of the five model covariate layers for TUc6. Dark regions signify higher variability within the data to an upper limit of 1 standard deviation. There is a clear association between regions of high variability and the RF model covariates, i.e., clay fraction and soil depth. Evidently, the highest clay and soil depth variability regions coincided with the regions that showed the lowest purity in spatial predictions of soil ecotope and soil type.

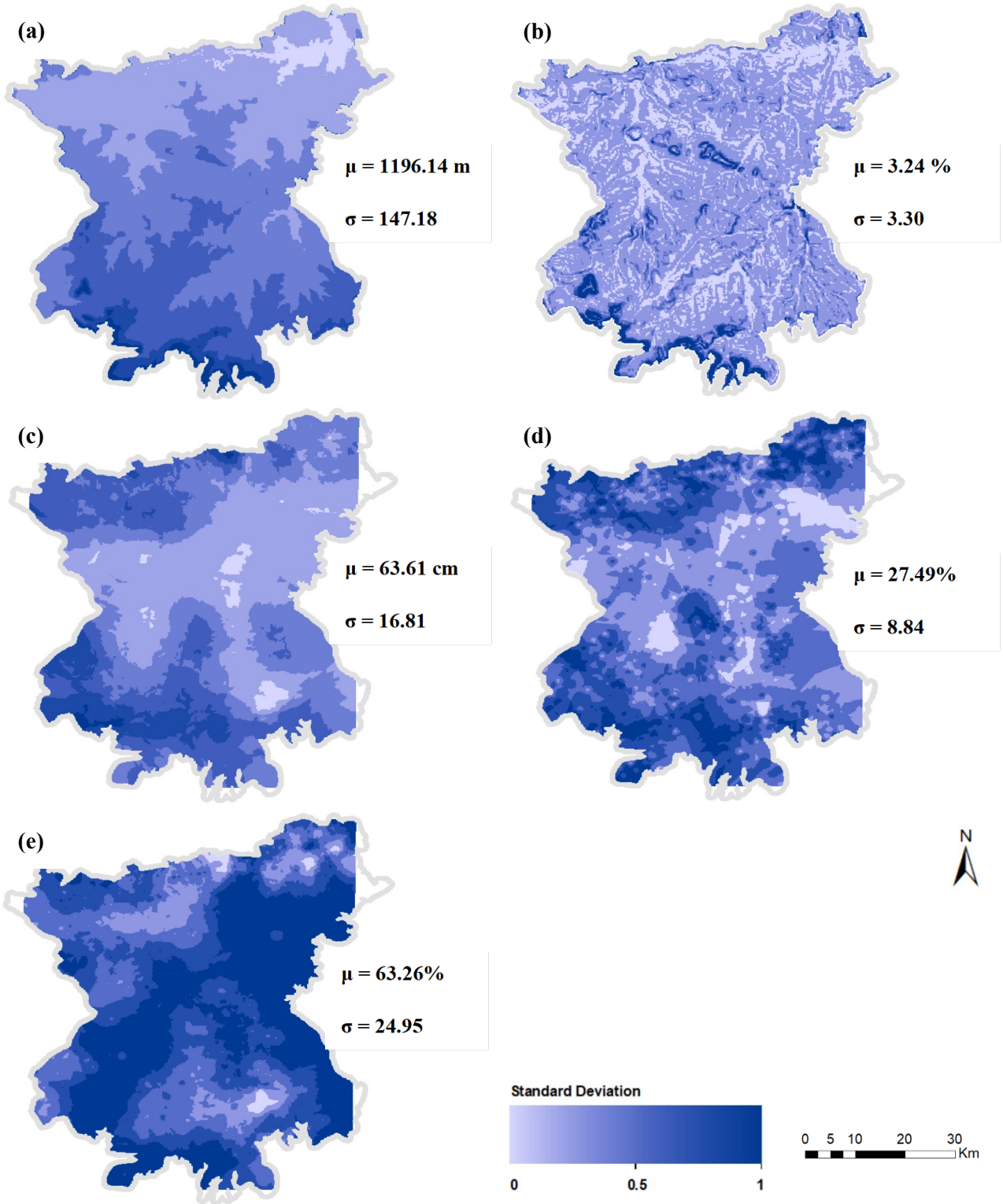


Figure 5.9: 90 m resolution covariate variability maps for TUc6 with population means (μ) and standard deviation of data (σ) displayed for (a) elevation, (b) slope gradient, (c) depth, (d) clay fraction, (e) surface rockiness.

When rationalising the sub-optimal prediction results in some BRU areas, the RF model's sensitivity to class imbalance is a second factor worth considering. An excellent example of this is the soil ecotope spatial prediction which identifies 27 soil ecotopes. The issue of class imbalance has already been described earlier concerning its influence on OOB error prediction and is well described in the literature (Mitchell, 2011; Janitza & Hornung, 2018). There is no doubt that classification trees are greatly affected by unbalanced samples because they tend to classify new observations into the class in which the majority of training observations originated. In our analysis, the ArcPro RF would have randomly drawn subsamples (or bootstrap samples) of observations from the original balanced sample. Typically, these subsamples can comprise as much as 63.2% of the observations contained in the original sample (Janitza & Hornung, 2018). By contrast, the resulting subsamples do not generally include precisely the same number of observations from each class. That is, the subsamples may be unbalanced or extremely unbalanced if many more observations are drawn by chance from one class than another.

Consequently, the level of class imbalance in the subsample directly relates to *n*, the original sample size. For a large *n*, the chances of a more substantial class imbalance in the subsample are small, while for a small *n*, the odds are much higher (Chen *et al.*, 2004). Janitza & Hornung (2018) point out that large class imbalance in subsamples is especially a problem for smaller samples. The imbalance in class representation causes trees to tend towards predicting classes that were more frequently represented in the in-bag sample, or equivalently, those that were less often represented in the out-of-bag sample, resulting in higher OOB error rates. The significant advantage of using ArcPro to run the RF model is that the user is alerted to any class imbalances through an error message stating that at least one of the categories in the *Variable to Predict* parameter contains relatively few values which can reduce the model's accuracy. The influence of this class imbalance and sample size on OOB error on the purity of the prediction of soil ecotope is presented in **Figure 5.11b.**

Finally, the modelled soil classes' taxonomic detail is also considered a primer for poor spatial prediction outputs. High impurity in the predicted soil ecotope map can be related to a region's high soil or taxonomic diversity. In this case, results resemble the same error pattern as the predicted soil types, but the extent of classification error is larger. This is because there are fewer observations in each ecotope class when the number of taxonomic classes increases. Jafari & Ayoubi (2013) observed similar results, who also observed that increasing taxonomic detail reduced prediction accuracy.

Further, to assess soil map uncertainty in the Netherlands, Kempen *et al.* (2009) used the Shannon index of entropy (H′) (Longuet-Higgins, 1971). It is accepted that heterogeneous areas display high entropy and low purity, and their prediction is accompanied by high uncertainty. The soil ecotope prediction map in TUc6, which contains many single-pixel ecotope class features, indicates this phenomenon.

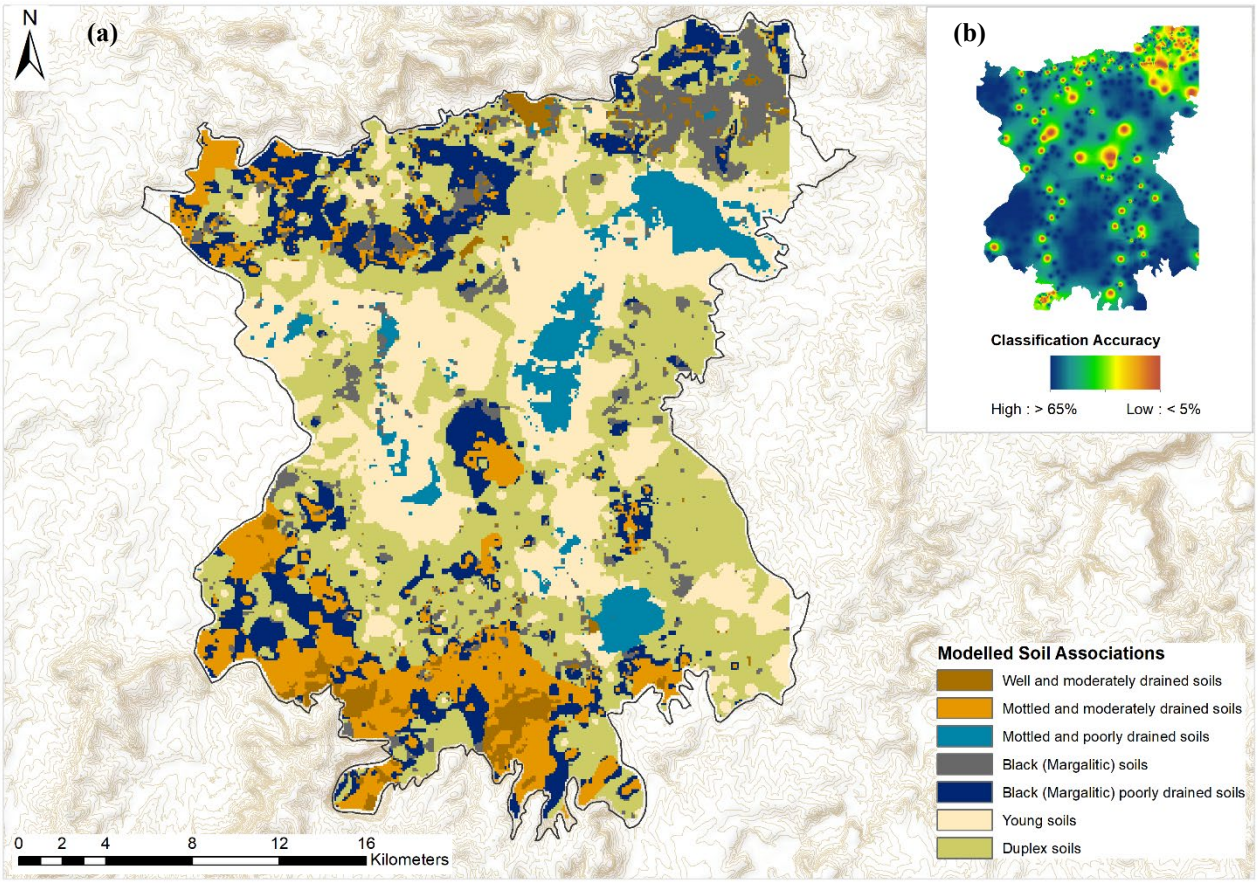


Figure 5.10: (a) Soil class-map predicted by the random forest model at the level of reference soil types for TUc6. (b) Mapped classification accuracy of predicted soil type based on 10% independent validation dataset out-of-bag error.

Despite selected regions displaying less than favourable prediction accuracies, The superior capability of RF is its prediction performance (Prasad *et al.*, 2006; Barthold *et al.*, 2013). The classification diagnostics for soil type (**Table 5.5**) and soil ecotope (**Table 5.6**) support this assessment with high prediction accuracy in model and regionalisation performance. As part of the Forest-based classification and regression model outputs, the RF provides classification diagnostic results for the training and validation data sets. Diagnostics are calculated using a confusion matrix, which keeps track of instances in which the category of interest is correctly classified and incorrectly classified and instances in which other categories are incorrectly classified.

Table 5.5: Training and validation classification diagnostics for soil type spatial prediction using RF model

	Training				Validation			
Soil Type	**F1-Score**	**MCC**	**Sensitivity**	**Accuracy**	**F1-Score**	**MCC**	**Sensitivity**	**Accuracy**
B	0.82	0.81	0.91	***0.97***	0.00	-0.11	0.00	***0.80***
D	0.83	0.80	0.79	***0.95***	0.59	0.51	0.62	***0.87***
E	0.76	0.77	1.00	***0.97***	0.55	0.58	1.00	***0.91***
F	0.55	0.60	1.00	***0.94***	0.18	0.18	0.50	***0.84***
G	0.75	0.72	0.82	***0.94***	0.31	0.22	0.33	***0.84***
H	0.83	0.77	0.75	***0.91***	0.50	0.42	0.38	***0.78***
J	0.79	0.72	0.71	***0.89***	0.43	0.36	0.31	***0.76***

Table 5.6: Training and validation classification diagnostics for soil ecotope spatial prediction using RF model

	Training				Validation			
Category	**F1-Score**	**MCC**	**Sensitivity**	**Accuracy**	**F1-Score**	**MCC**	**Sensitivity**	**Accuracy**
B110	0.67	0.70	1.00	***0.99***	0.40	0.49	1.00	***0.95***
B111	0.90	0.90	1.00	***1.00***	0.00	-0.02	0.00	***0.97***
B112	0.71	0.74	1.00	***0.99***	1.00	1.00	1.00	***0.97***
B122	0.92	0.92	1.00	***1.00***	-	-	-	**-**
B210	0.78	0.80	1.00	***0.99***	0.00	-0.02	0.00	***0.97***
D110	0.89	0.89	0.92	***0.99***	0.00	-0.04	0.00	***0.92***
D111	0.78	0.79	0.65	***0.98***	-	-	-	**-**
D112	0.80	0.80	0.72	***0.98***	0.57	0.55	0.67	***0.95***
E240	0.86	0.85	0.90	***0.99***	0.40	0.38	0.50	***0.95***
E322	0.33	0.43	1.00	***0.93***	0.33	0.43	1.00	***0.94***
E340	0.70	0.73	1.00	***0.99***	1.00	1.00	1.00	***0.98***
F120	0.83	0.84	1.00	***0.99***	0.00	-0.02	0.00	***0.95***
F212	0.59	0.64	1.00	***0.98***	-	-	-	**-**
G120	0.78	0.80	1.00	***0.99***	0.29	0.39	1.00	***0.92***
G121	0.94	0.94	0.89	***1.00***	0.40	0.38	0.33	***0.95***
G122	0.78	0.78	0.67	***0.98***	0.00	-0.04	0.00	***0.92***
H122	1.00	1.00	1.00	***1.00***	0.00	-0.02	0.00	***0.97***
H223	0.59	0.64	1.00	***0.98***	0.00	-0.03	0.00	***0.94***
H240	0.83	0.84	1.00	***1.00***	-	-	-	**-**
H343	0.73	0.67	0.62	***0.88***	0.67	0.66	0.50	***0.89***
J120	0.95	0.95	1.00	***1.00***	0.00	-0.02	0.00	***0.97***
J121	0.90	0.90	1.00	***1.00***	0.29	0.28	0.50	***0.92***
J122	0.86	0.85	0.86	***0.99***	0.50	0.48	0.50	***0.97***
J220	0.76	0.76	0.72	***0.98***	0.00	-0.03	0.00	***0.94***
J221	0.63	0.66	0.46	***0.96***	0.00	-0.05	0.00	***0.90***
J222	0.58	0.57	0.47	***0.94***	0.25	0.25	0.17	***0.90***
J223	0.52	0.59	1.00	***0.98***	0.00	-0.03	0.00	***0.92***

These results are easily interpretable by the end-user: The sensitivity is expressed as the percentage of times a feature associated with an observed category was correctly predicted (ESRI, 2021). For example, in **Table 5.5**, if the aim is to predict soil type and soil type E has a sensitivity of 1.00, then all instances where soil type E is being predicted will correctly be labelled as soil type E. In contrast, if soil type D were incorrectly marked as E, it would not be included in its (E) sensitivity number. This would instead be reflected in soil type D sensitivity, which would mean one of the D features was not correctly predicted. In addition, training datasets are also provided with diagnostics to compare predicted values to observed values. Using these diagnostics lets users determine how well their model(s) matches the training data. ArcPro's RF model includes two additional metrics in the classification diagnostics: the F1-score and the Matthews Correlation Coefficient (MCC)(Matthews, 1975; Baldi *et al.*, 2000). Both are computed from the model(s) classification confusion matrices. Both are considered standard performance metrics for machine learning with a natural extension to a multiclass case represented by this study (Gorodkin, 2004). In short, the F1 score (2 * (Precision * Recall) / (Precision + Recall)) is a measure of a model's accuracy on a dataset. It incorporates both classification Recall and Precision to provide a more balanced view of model performance, and it is defined as the harmonic mean of the model's precision and recall with values approaching 1.00, indicating ideal performance (Padilla *et al.*, 2020).

There is value in the F1-score in situations where false positives are more costly than false negatives, or there is a considerable imbalance in class frequency. Alternatively, the MCC is a more reliable statistical metric. It produces a high score when the prediction obtains good results in all of the four confusion matrix categories (true positives, false negatives, true negatives, and false positives)(Chicco & Jurman, 2020). Similar to a correlation coefficient, the range of values of MCC lie between -1 to +1. A model with a score of +1 is a perfect model, and -1 is a poor model. It is one of the main advantages of MCC because it makes the method easy to interpret. Because the MCC is perfectly symmetric, no class is more important than the other; switching the positive and negative will still produce the same result (Baldi *et al.*, 2000; Chicco & Jurman, 2020). In particular, the summary performance results for the spatial prediction of soil type and soil ecotope can be a valuable addition to the BRW, as no such quantitative measurements are currently available for any of the respective soil ecotope descriptions in the BRUs. More importantly, since the soil prediction maps are generated at finer spatial resolutions (90 m ≈ 0.81 ha), users can investigate the benefit of higher precision land management approaches applicable to sub-BRU level, i.e., managing agricultural resources at farm level rather than the entire 73 127 ha region.

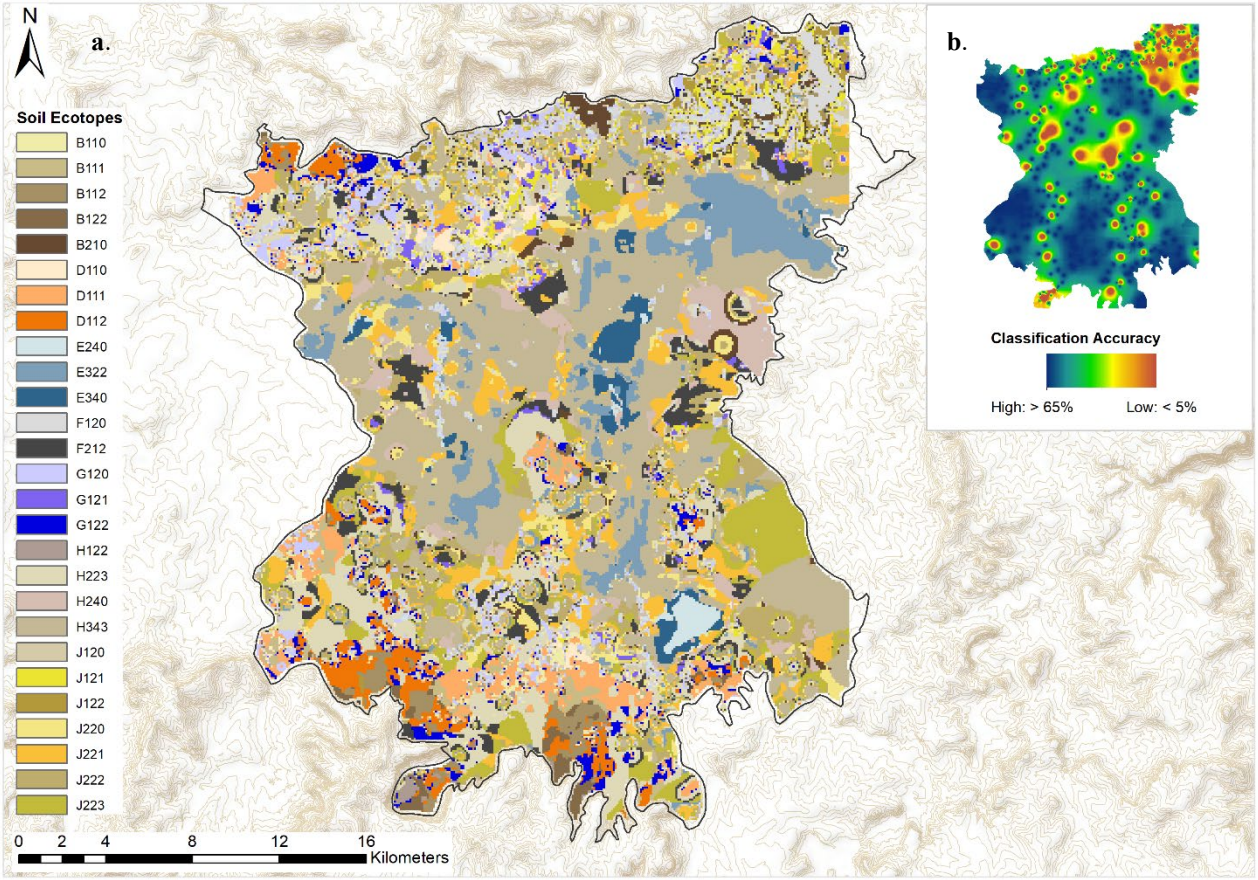


Figure 5.11: Soil-class map predicted by the random forest model at the reference soil ecotope class level for TUc6. (b) Mapped classification accuracy of predicted soil ecotope based on 10% independent validation dataset out-of-bag error.

5.4.2.4 Variable importance

Using the variable ranking provided by the ArcPro Forest-based classification and regression algorithm, it was possible to determine the contributions of each explanatory variable (covariate) to the soil predictions. Variable importance aims to create a more parsimonious model that details variables that are meaningful to model performance. Importance is calculated using Gini coefficients, which can be thought of as the number of times a variable is responsible for a split. The impact of that split is divided by the number of trees, i.e., splits are each decision within a decision tree (ESRI, 2021). It is important to note that variable importance is a diagnostic that explains which variables are driving the model's results and is not a measure of model accuracy or how well the model is predicting (Strobl *et al.*, 2008). ArcPro calculates variable importance as the sum of the Gini coefficients from all the trees for each variable listed and provides a helpful breakdown of the percentage of the total sum of Gini coefficients which can be readily displayed as an optional bar chart displaying the importance of each variable used in the model (**Figure 5.12**). Each bar shows the distribution of the variable importance values across all the validation runs. The distribution of the variable importance is an indicator of the stability of the trained forest model. If the importance of a variable is changing broadly across the validation runs, indicated by a long box in the chart, this may indicate an unstable random forest model. An unstable model can often be improved by adjusting the RF *ntree* parameter (see §*5.3.7)* to capture more complex relationships in the soil data (Grömping, 2009). For clarity, I have shown the variable importance output for both soil type and soil ecotope together. In both cases, clay fraction is the highest-ranked variable, describing roughly 25 to 27% model importance. This is followed by surface rockiness, which accounts for more than 23% and soil depth at 20% of the model importance when mapping soil ecotope. These results are not surprising when considering the covariate variability results, which corroborate that clay fraction, surface rockiness and depth are highly spatially variable within TUc6.

Although the overall importance of the slope gradient was low, the results must be considered with the topographic expression of the study site. **Figure 5.9** underlines the broad spatial variability of the 90 m slope gradient coverage with only select regions of high slope variability. However, slope gradient is considered one of the pivotal soil-forming state factors (Jenny, 1941; McBratney *et al.*, 2003; Florinsky, 2016). In regions of high topographic expression, we would expect to observe more significant spatial variability and higher variable importance. In both instances, i.e., spatial variability and variable importance, the slope gradient representation are far more descriptive than presently provided by the BRW. Users still have to contend with slope gradient representations that are too generalised and representative for the entire BRU – limiting sub-regional site-specific management. Therefore, the BRW would benefit from including such metrics in the report system since it still provides no spatial context of any soil covariate distributions or quantitative accuracy measure of any sort for evaluating soil type, soil ecotope, or associated crop-yield estimations (expert-based or model-driven) relating to decision-support performance.

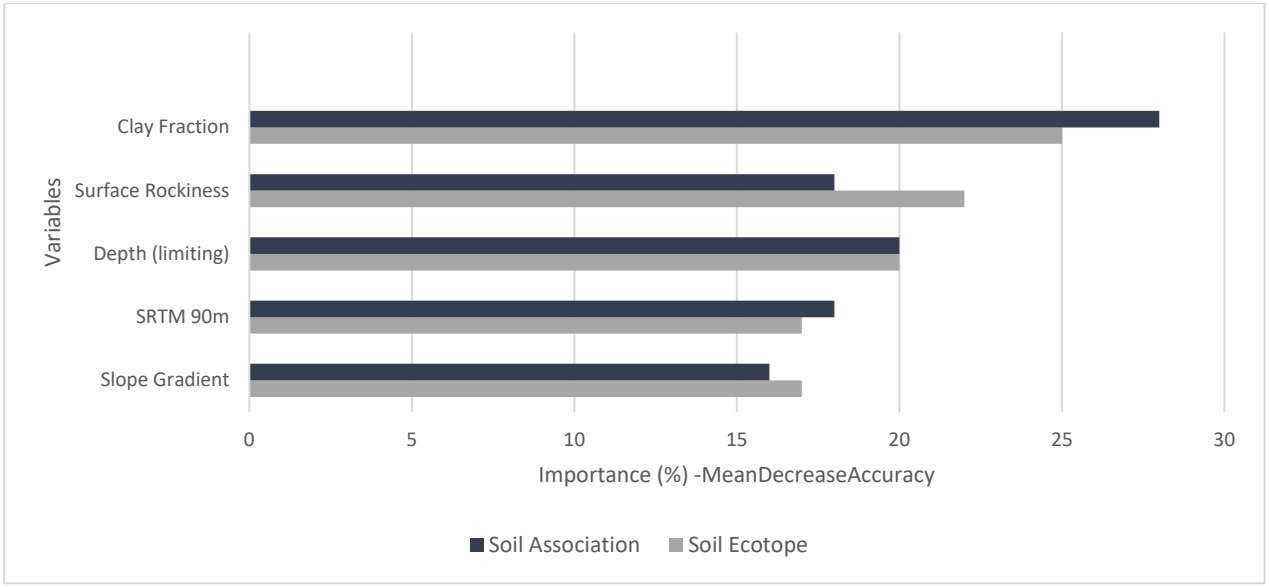


Figure 5.12: Variable importance plot for RF model covariates based on a mean decrease in prediction accuracy.

5.5 CONCLUSION

The BRW has not seen a major revision for over two decades. Still, the natural resource information it contains provides land managers, policymakers and farmers with invaluable access to regional and farm level qualitative estimations of agricultural productivity. There is a need to preserve this information while simultaneously providing modern measures of land management recommendation at multiple scales to the end-user. This work was the first attempt in KwaZulu-Natal to produce a full extent soil type and soil ecotope map for a non-specific BRU using a Forest-based classification and regression tree approach (RF). This chapter has demonstrated that using the ArcPro RF model to predict soil type and soil ecotope properties using a set of environmental covariates, including detailed soil calibration locations, can be a gateway approach for updating the natural resource information delivered by the BRW. I have further demonstrated that our quantitative GIS soil-landscape model leveraged the existing relationship, defined in the BRU, between soil properties and five environmental variables: slope gradient, elevation, pedon depth, surface rockiness and in-field clay fraction. This approach is deliberate to ensure seamless interoperability with the existing BRW data repository, facilitate harmonised interpretation of results between the current BRU recommendations and the new modelled results and impose minimal capacity-building or upskilling effort to the end-user. A leading consideration regarding the continued relevancy of the BRW is that the program is unable to readily meet the demands of a noticeable change in agronomic land-use patterns, both in size and ownership, in KwaZulu-Natal. In the past, the program's relevance was driven by the need for users to characterise their resource availability better, concerning limitations and land potential across a broad region since most commercial landholdings were large in extent (Jewitt *et al.*, 2015). However, the growing emphasis on equitable access and equitable opportunity to land has resulted in these large tracts of land not only diminishing in size but now including changes from agriculture to non-agricultural land use activities.

Moreover, there is an expectation for these smaller, subdivided landholdings to produce favourable agricultural yields consistently. The ability to re-imagine the BRW at a farm scale would significantly increase its application as a regional decision support tool. This study developed high-resolution soil type and soil ecotope maps (90 m ≈ 0.81 ha), which can be used for many environmental, vegetation type, agronomic, or land-management studies. Since the soil prediction maps are generated at finer spatial resolutions, users can take the opportunity to test the effectiveness of higher precision land management approaches for large farms, as opposed to managing agricultural resources over the entire 73 127 ha region.

The utility of RF successfully resulted in maps of soil type and soil ecotope for BRU TUc6. No other reports or studies have attempted to map the soil resources for an entire BRU using GIS or achieve the degree of accuracy reported in this study. Overall prediction accuracy varied from 76% to 91% when modelling all seven soil types and 92% to 98% for the 27-modelled soil ecotopes. Model OOB rates were positively correlated with class frequency and negatively correlated with ecotope taxonomic detail. The outlined protocol for modelling soil and ecotope classes includes the spatial overview of covariate variability within the feature domain, which can be included as a value-added product in the BRW and the variable importance of these covariates on model performance. Regarding variable importance, for TUc6, it was shown that clay fraction is the highest-ranked variable in soil type and soil ecotope prediction, describing roughly between 25-27% of model importance. The next most significant factor affecting soil ecotope modelling is surface rockiness, with 23% importance, and soil depth, with 20% importance.

This chapter has also outlined how model and prediction performance can be improved since all results were obtained through default parameter settings optimised based on the data distribution and RF model diagnostics. Indeed, improved model predictions may be gained through necessary program-setting adjustments. However, these need not be at the expense of determining if the model's default settings achieve the desired objective. The main advantage of the ArcPro RF model is its ability to determine model optimisation(s) based directly on the input data type, distributions and particular prediction model selected. While this may seem to limit the ability to specific model customisations by more experienced users, it does allow for a more egalitarian approach to model design by a broader end-user demographic. A second possible case for improving the model outputs relates directly to the environmental covariate and soil data quality (spatial and attribute). I have shown that class imbalance in regards to RF model importance is an important consideration. However, a key consideration worth acknowledging is the significant role that spatial resolution of covariates plays in soil prediction (Smith *et al.*, 2006; Cavazzi *et al.*, 2013; Atkinson *et al.*, 2020).

Further improvements must be justified against the costs of obtaining higher or lower resolution data and the impacts on processing and computation time. This study has further shown that the inclusion of a more superior representation of slope gradient and TMU, through introducing the concept of geomorphons into BRU's, not only provides a far richer interpretation of the general BRU site descriptions but underpins our understanding of soil-landscape and geomorphological relationship within the BRU. The lack of detailed terrain description is a fundamental limitation of the existing BRW, which can be immediately improved with the accessibility of high-resolution SRTM 90 m DEM. I not only report on the improvements that such DEMs offer on slope gradient and TMU descriptions but present the benefit of supporting these qualitative soil-landscape descriptions with spatial maps as tools for synoptic landscape evaluation.

CHAPTER 6

6 CONCLUSION: THE OUTLOOK FOR REMOTE SENSING AND GEOGRAPHICAL INFORMATION SYSTEMS FOR IMPROVED SOIL MAPPING IN KWAZULU-NATAL

"without fieldwork, the exceptions to the current understanding of soil landscapes cannot be discovered. Therefore, increases in mapping efficiency due to advances in remote sensing and spatial pedologic modelling should not be mistaken for a reason to invest fewer resources to fieldwork."[5]

[5] Brevik, E.C., Calzolari, C., Miller, B.A., Pereira, P., Kabala, C., Baumgarten, A., Jordán, A., 2016. **Soil mapping, classification, and pedologic modeling: History and future directions**. *Geoderma* 264, 256-274.

6.1 SYNTHESIS

In South Africa, traditional soil mapping and landscape assessment approaches have served land development needs well. The gradual progression of soil and land assessment focused on classification and inventory to include quantifying spatial soil and environmental relationships and formalising expert knowledge on soil-landscape relationships would not be possible without GIS and RS facilities. Undeniably, advances in computational algorithms, numerical and predictive analysis and geo-visualisation combined with readily accessible multi-resolution remotely sensed terrain data have revolutionised digital geomorphic analysis, particularly in geopedology – the integration of geoforms and soils. Indeed, the assessment of large (un)known areas and the ability to evaluate uncertainty as well as new ways to visualise soil-landscape relationships have meant that digital soil methods have become essential additions to the soil surveyors toolset. However, the multi-disciplinary application of geospatial technologies to address numerous conceptual and formative contemporary soil-landscape endeavours is not without its challenges. Foremost, approaches in digital soil mapping and digital geomorphological mapping are still defined from multiple perspectives. Despite global efforts to establish harmonised assessment protocols, these approaches still lack precise criteria for formalising regional frameworks of soil-landscape digital analyses. Notwithstanding new capabilities, many inquiries persist regarding advantages and limitations of topography and landform representation, landform taxonomy, terrain pattern-process characterisation, uncertainty estimation and issues involving the scale-dependence of environmental properties combined with the mapping objectives at various scales. A key consideration to resolving many of these uncertainties lies in the ability of users to address the advantages and limitations of DEM surfaces since their quality inherently define the representation, and thereby the accuracy of application, of soil-landscape relationships. In recent years, it appears that researchers are becoming interested in formalising pedological expertise within the realm of rapidly advancing geo-computation tools.

The ease of automation, swift replication and acceptable representation of an expanded set of morphometric classes using geomorphons have shown to be a suitable alternative to the basic 5-class Land Type discrete relief models for South Africa. This study aimed to develop a sustainable and practical framework for integrating better descriptions of terrain unit characterisation into digital terrain analysis for soil assessment and mapping using advanced geomorphometric and remote sensing approaches. By emphasising current issues of concern in DSM and terrain analysis, this study has attempted to broaden and strengthen the soil science community's understanding of how to optimise better and incorporate proven geospatial services and tools that have become increasingly mainstream in terrain mapping under regional circumstances. As part of this study, soil-landform units are assessed, quantified, modelled, and visualised using high and medium resolution DEMs to classify local landform elements and relate these elements to specific soil geopedological associations. A key feature of the study presents aspects of DEM quality on soil-landscape covariates selection, explains ways to assess DEM error, and presents the critical role of DGM in addressing present-day sustainable development issues regarding KwaZulu-Natal. Starting on the south coast of the Province, the study successfully evaluated how spatial resolutions and search window generalisation influence the characterisation of topographic and terrain data generalisation and the potential impacts of geopedological inputs to DSM. The study found that sensor selection and DEM resolution significantly influence the utility of the application. When representing scale-dependent terrain processes to be modelled, the generalisation approach to be used must be carefully considered.

The scale-interpolation dichotomy implies a degree of inaccuracy in identifying topographic features across the site, with optimal results achieved when the neighbourhood size and DEM pixel resolution correspond to the degree of topographic variability within the landscape. For a coastal region of KwaZulu-Natal, this study

has confirmed that terrain attributes respond differently to resolution changes, particularly when switching DEM resolutions between generalisation methods. Indeed, this study has confirmed much of what has been established internationally that DEM resolutions and terrain variable extraction for soil-landscape modelling should align with the landscape-scale processes being modelled. Based on the overall findings, this study recommends that users select DEM surface models primarily influenced by research and specific to a landscape. The selection of DEMs should be based on their performance either at a regional or local level.

Shifting the focus to the central interior region of KwaZulu-Natal, this study was able to descriptively introduce the concept of local ternary patterns, otherwise known as geomorphons. The methods presented herein provide an appealing outlook of geomorphon characterisation across varying DEM resolutions to highlight how different DEM source and spatial resolutions influence the representation of soil-landscape patterns and process phenomena. Among the research results, a test was conducted of the notional scale-independent property of the geomorphon approach to various soil parameters and terrain parameters with varying DEM resolutions. The findings have demonstrated that DEM resolution influences the extraction, generalisation and representation of digitally derived soil-terrain attributes and the extent of terrain units. Interestingly, geomorphons may be more appropriate for setting landscape and regional conditions than direct application at the farming level. Notably, this study has revealed that DEM selection must consider the purpose for which the application is being developed. When DEMs are not fit for purpose, digital representations and analyses of geomorphic properties may be severely compromised. As a result, practitioners of digital soil mapping should be aware of the implications of inherently weak data for soil-landscape and digital soil mapping analysis when incorporating elevation datasets, especially in South Africa.

Furthermore, the research outputs have demonstrated how the integration of soil-landscape covariates within geomorphons may contribute to the concept of present landscape pattern-process character assessment, as defined by the distinctive, recognisable, and consistent set of elements in the landscape that distinguish one landscape from another. Although this study does not definitively describe the characteristics of each geomorphon for complete landscape character assessment, there is potential benefit in the synoptic delineation of a geomorphic pattern signature to define agricultural land potential prioritisation and categorisation for a diverse range of land use planning functions.

Further unpacking the geopedological characterisation of the KwaZulu-Natal interior, albeit, from a landscape development perspective, this study successfully introduced a simplistic, objective and replicable numerical approach to assess the landscape geodiversity for the entire uThukela District Municipality. Although the importance of geodiversity in landscape assessment is often ignored, this study highlights how hydrographic lithostratigraphic, pedological, climatic, topographic, atmospheric and geomorphological diversity can be clearly connected. An innovative approach to quantifying landscape diversity using geospatial grid-based methods is presented herein, making KwaZulu-Natal one of the first to apply it. Another key highlight is the overarching endogeneity of pedological and topographic influences on geodiversity quantification. There is a need to recognize the value of novel approaches to addressing emerging concerns about the functioning of soil systems in the provision of services required by modern societies.

Consequently, landscape character assessment for sustainable land use decision-making is becoming more holistic and more inclusive in scope to better understand biodiversity and geodiversity, leading to a combination of DSM and geodiversity assessment to popularize digital soil inventories under changing resource gradients. Globally, there is increasing awareness and acceptance of the importance of geodiversity in territorial management, particularly in promoting resilient landscapes through the preservation of natural diversity. It is expected that this study will motivate future studies on geodiversity in South Africa, as persistent

environmental pressures will soon require maps expressing this concept to become more common as decision-support tools. Agriculture remains a major driver of food security and economic development in South Africa. Nevertheless, South Africa faces a declining supply of high potential agricultural land, which is crucial for the sector's long-term sustainability. Socio-economic pressures have led to the irreversible loss of medium to high potential agricultural land by converting land to uses other than agriculture. This has resulted in a reduction of land available for farm production and increased alienation of agricultural land. Due to limited agricultural land in South Africa and land in general, land managers, policymakers, natural resource managers, farmers, and investors face a severe challenge of identifying and prioritising the most appropriate and suitable land use options. Thus, decisions related to land use and land management impact the quantity and quality of agricultural resources available to society and the agro-ecological goods and services they provide. There is no doubt that sustainable intensification is necessary to address the challenges of population growth and food security. However, land use should be optimally aligned to maximise the potential of the land in a manner that does not deplete or compromise agricultural resources. In this challenge, using a suite of natural resource, climate and terrain information provides an opportunity to define, characterize, and manage agricultural resources and explore or develop pragmatic avenues of agricultural land use prioritisation in KwaZulu-Natal. The Bioresource Unit Classification (BRU) is the primary source of long-term average yield data for various crops and guides agricultural development for KwaZulu-Natal. However, many of these crop yield estimates are outdated and require a revisit. Therefore, remote sensing technologies can significantly expand the investigative reach of resource managers and decision-makers in the Province.

Concluding the research in the Central Drakensberg region of the Province, the study detailed the protocol for spatially defining soil ecotope and soil type inventories part of the Bioresource Classification Report Writer Program (BRW). The research findings demonstrate how forest-based classification methods and DSM can effectively predict soil type and soil ecotope properties using a set of improved geopedological covariates. In particular, the approach(s) described in this study allows seamless interoperability with the existing BRW decision support platform and data repository, simplifies the interpretation of results between the current BRU recommendations and the modelled results, and requires minimal capacity-building effort on the part of users. Currently, the BRW can only assess the natural resources and agricultural production potential at a regional level given its terrain morphological linkages to the National Land Type Survey. Thus, there is limited opportunity to further disaggregate the large BRUs without first supplementing the coarse-scale resource data with detailed *in situ* field surveys. Therefore, accurately predicting and spatially representing smaller, homogenous agricultural management units, such as soil ecotopes, should enable better farm-level management of medium and high-value agricultural resources within the landscape as a first approximation. The added value to the end-user of using DSM approaches to features like ecotopes or soil type is that large areas can be systematically mapped repeatedly with measures of uncertainty included in the final mapping results. Thus, any additional field surveying would be driven by further precision-management value-added benefit or attempts at validating model predictions and supporting the often *defacto* expert-driven approach to resource assessment.

6.2 DIRECTIONS FOR FURTHER RESEARCH

The recognition of the importance of DEM selection for multi-scale soil-landscape applications is certainly not new. In fact, the numerous innovations in the digital representation of variables describing the topographic surface have been a major commodity of the geospatial revolution in the last two decades. Likewise, the formative relationship of topography as a critical functional characteristic of soil formation in any landscape means that DEMs will remain a featured priority in all present and future soil-landscape assessments. Indeed, while progress is regularly being made in a wide range of DEM, DSM and DGM domains, the continual exploration of advanced GIS-based applications of place-based geomorphic landscape assessment remains high on the research agenda. In actual fact, as long as advances in these respective domains are continually being pursued, there will always be a need to address new concepts and knowledge integration in the context of analytical and geospatial capabilities under varying environmental settings. This is further emphasised by the confluence of digital geopedological technology as standardised, functional solutions to address a range of sustainable land management needs. Moreover, these opportunities will always be framed within the context of the evolving approaches that are either expert-based rule sets *vs* completely automated classifications or a combination of both to meet specific objectives. Similarly, as new advances introduce new opportunities for exploration, application and knowledge transfer, they introduce new problems that will require holistic understanding and detailed attention.

Notwithstanding the significant benefits of data modelling and GIS science in addressing real-world challenges, it would be highly optimistic to assume that fieldwork description can entirely be separated from the DSM framework. So while new avenues of geospatial optimisation are necessary, these initiatives should not be at the expense of thorough fieldwork undertakings. On the contrary, digital and conventional soil assessment approaches should aim to support rather than entirely supplement the other within the various soil mapping stages. The turnkey framework presented in this study offers a pragmatic approach for balancing the requisite data preparation requirements across RS datasets and GIS platforms to maximize the operational effectiveness of DSM products within a regional setting but scalable up to a national level if necessary. A notable feature of the framework is that it deliberately imposes minimal barriers to uptake by the end-user. As a result, it should facilitate a wide range of applications, ranging from research to field-based assessments, by experts and novices alike. Moreover, all DEMs used to develop this research are available through open access or at competitive prices, and all DGM, DSM, and cartography of the final products were processed using either open-source GIS platforms or software in a format easily accessible to open-source platforms. The potential for further refinement and adaptions of the methods detailed in this research to benefit the local strategic landscape management objectives and much broader soil-landscape desiderata is well-positioned.

Three focus areas of the current research outputs that would benefit the most from particular research attention are the multi-resolution soil-landscape characterisation of the Bergville region, the grid-based approach to geodiversity quantification of the uThukela District Municipality and lastly, the soil ecotope and soil type mapping for BRU TU6c. While the study suggests that the inclusion of legacy soil-landscape covariates within geomorphons is useful in assessing current landscape pattern-process complexity, users should consider the measurement scale for the terrain parameters under observation, i.e., discrete versus continuous. It follows that this concept also applies to field-based over measured estimates of soil properties since field-based estimates are almost always discrete or categorical in nature. However, as this study has established, simply improving the precision of soil measurements or terrain resolution does not directly translate to better yielding predictive soil-landscape relationships. Therefore, future research opportunities may need to reconcile the treatment of computational and geographic scales for both DGM pattern-process applications and the cartographic

representation of results and consider the currency and applicability of older legacy datasets in the present-day land use changes and concerns.

The assessment of geodiversity in this study is based on the grid-based assessment of edaphic partial indices. However, the environmental covariates examined are neither conclusive nor exhaustive. Of course, this opens up new possibilities through further consideration of alternative or complementary covariates but also limits replicability and potential understanding of geodiversity within the same region. Interestingly, this is a well-known limitation of geodiversity measurements globally and the subject of ongoing research. Furthermore, beyond this study, there have not been any other attempts to quantify geodiversity for KwaZulu-Natal outside of the conventional geoconservation and geoheritage focal themes. Therefore, while this study represents a landmark contribution of geodiversity assessment to the region, no other studies corroborate the results. This study should then gain further exposure as a benchmark for similar succeeding geodiversity research or ground-truthing efforts in the region.

Finally, the application of improved DEMs to derive geomorphons for improved terrain unit descriptions combined with digital mapping of soil ecotopes using superior machine-learning methods as a newly proposed protocol for updating the resource information presented by the BRW a novel intervention from the research. To the my knowledge (at the time of print), no other regional study has attempted to provide a DSM solution for spatially and synoptically defining soil ecotopes for an entire BRU. The next evolution of this protocol would require a thorough field assessment to validate soil ecotope predicted boundaries and properties and further refine DSM model parameters to yield better and more efficient results. Future works may consider further (re)defining soil type and soil ecotope maps for other BRUs and possibly spatially evaluating soil properties with crop yield. This would be immensely beneficial as an entirely vertically integrated solution to sustainable agricultural management. Additional benefits may include predicting soil properties at multiple depths across BRU soil types, possibly according to GlobalSoilMap project specification since South Africa is a member country of the partnership. The inclusion of statistical and independent model accuracy metrics for soil type and soil ecotope mapping will provide much-needed measures of quantitative resource evaluation. Aligning the BRW to such a data schema will facilitate harmonisation with internal standards GlobalSoilMap project specifications which have become a priority in DSM in recent years.

REFERENCES

A-Xing, Z., Smith, M., Rongxun, W., Jing, G. (2008). *The impact of neighbourhood size on terrain derivatives and digital soil mapping*. In Q. Zhou, B. Lees, G. Tang (Eds.), *Advances in digital terrain analysis* (pp. 333-348): Springer, Berlin, Heidelberg.

Adediran, A.O., Parcharidis, I., Poscolieri, M., Pavlopoulos, K. (2004). Computer-assisted discrimination of morphological units on north-central Crete (Greece) by applying multivariate statistics to local relief gradients. *Geomorphology, 58*(1-4), 357-370.

Aguilar, F.J., Agüera, F., Aguilar, M.A., Carvajal, F. (2005). Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy. *Journal of Photogrammetric Engineering and Remote Engineering, 71*(7), 805-816.

Ahrens, R.J. (2008). *Digital soil mapping with limited data*: Springer Science & Business Media.

Ai, T. (2007). The drainage network extraction from contour lines for contour line generalization. *ISPRS Journal of Photogrammetry and Remote Sensing, 62*(2), 93-103.

Ai, T., Li, J. (2010). A DEM generalization by minor valley branch detection and grid filling. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(2), 198-207.

Alba Alonso, S., Ibáñez Martí, J.J., Lobo Aleu, A., Zucarello, V. (1998). Pedodiversity and global soil patterns at coarse scales *Geoderma, 83*, 171-192.

Albani, M., Klinkenberg, B., Andison, D., Kimmins, J. (2004). The choice of window size in approximating topographic surfaces from digital elevation models. *International Journal of Geographical Information Science, 18*(6), 577-593.

Allen, J., Somerfield, P., Gilbert, F. (2007). Quantifying uncertainty in high-resolution coupled hydrodynamic-ecosystem models. *Journal of Marine Systems, 64*(1-4), 3-14.

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., Jetz, W. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific data, 5*, 1-15.

Amrhein, C.G. (1995). Searching for the elusive aggregation effect: evidence from statistical simulations. *Environment and Planning, 27*(1), 105-119.

Andersen, H., Reutebuch, S.E., McGaughey, R.J. (2005). Accuracy of an IFSAR-derived digital terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing, 31*(4), 283-288.

Anderson, A.B., Wang, G., Gertner, G. (2006). Local variability based sampling for mapping a soil erosion cover factor by co-simulation with Landsat TM images. *International Journal of Remote Sensing, 27*(12), 2423-2447.

Andrle, R., Abrahams, A.D. (1989). Fractal techniques and the surface roughness of talus slopes. *Earth surface processes and landforms, 14*(3), 197-209.

Araujo, A.M., Pereira, D.Í. (2018). A new methodological contribution for the geodiversity assessment: applicability to Ceará State (Brazil). *Geoheritage, 10*(4), 591-605.

ARC. (1991). *A procedure for describing soil profiles*.Agricultural Research Council- Institute for Soil Climate and Water, Report No. GB/A/91/67. Agricultural Research Council: ISCW, Pretoria, South Africa.

ARC. (2003). *1:250 000 scale Land Type Survey of South Africa*.Agricultural Research Council, Agricultural Research Council: ISCW, Pretoria. South Africa.

Argyriou, A.V., Sarris, A., Teeuw, R.M. (2016). Using geoinformatics and geomorphometrics to quantify the geodiversity of Crete, Greece. *International Journal of Applied Earth Observation and Geoinformation, 51*, 47-59.

Arrouays, D., Vion, I., Kicin, J.L. (1995). Spatial analysis and modeling of topsoil carbon storage in temperate forest humic loamy soils of France. *Soil Science, 159*(3), 191-198.

Arrouays, D., McKenzie, N., Hempel, J., de Forges, A.R., McBratney, A.B. (2014). *GlobalSoilMap: basis of the global spatial soil information system*: CRC press.

Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A.C., Morgan, C.L., Roudier, P., Poggio, L., Mulder, V.L. (2020). Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional, 20*, e00255.

Atkinson, J., de Clercq, W., Rozanov, A. (2020). Multi-resolution soil-landscape characterisation in KwaZulu Natal: Using geomorphons to classify local soilscapes for improved digital geomorphological modelling. *Geoderma Regional, 22*, 1-18.

Atkinson, J.T., Rozanov, A.B., De Clercq, W.P. (2017). Evaluating the effects of generalisation approaches and DEM resolution on the extraction of terrain indices in KwaZulu Natal, South Africa. *South African Journal of Geomatics, 6*(2), 245-261.

Bailey, J.J., Boyd, D.S., Hjort, J., Lavers, C.P., Field, R. (2017). Modelling native and alien vascular plant species richness: At which scales is geodiversity most relevant? *Global Ecology and Biogeography, 26*(7), 763-776.

Bainbridge, W.R., Marneweck, G.C. (1996). *Natal Drakensberg Park: Information sheet for the site designated to the List of Wetlands of International Importance in terms of the Convention on Wetlands of International Importance especially as Waterfowl Habitat.* South African Wetlands Conservation Programme.Department of Environmental Affairs and Tourism, 24/21/3/3/3/14. Pretoria.

Baker, V.R. (1986). Fluvial landforms. *Geomorphology from space*, 255-315.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A., Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics, 16*(5), 412-424.

Barber, C.P., Shortridge, A. (2005). Lidar elevation data for surface hydrologic modeling: Resolution and representation issues. *Cartography and Geographic Information Science, 32*(4), 401-410.

Barichievy, K., Botha, C., (2017). *Verifying land assessment techniques using high resolution yield data*, Combined Congress: Soil, Horticulture and Grassland, Western Cape, South Africa.

Barrett, C.B. (2010). Measuring food insecurity. *Science, 327*(5967), 825-828.

Barthold, F., Wiesmeier, M., Breuer, L., Frede, H.-G., Wu, J., Blank, F. (2013). Land use and climate control the spatial distribution of soil types in the grasslands of Inner Mongolia. *Journal of arid environments, 88*, 194-205.

Baugh, C.A., Bates, P.D., Schumann, G., Trigg, M.A. (2013). SRTM vegetation removal and hydrodynamic modeling accuracy. *Water Resources Research, 49*(9), 5276-5289.

Baumgardner, M. (1999). *Soil Databases*. In M. Summer (Ed.), *Handbook of soil science*: CRC press.

Baxter, S., Oliver, M.A. (2005). The spatial prediction of soil mineral N and potentially available N using elevation. *Geoderma, 128*(3-4), 325-339.

Beasom, S.L., Wiggers, E.P., Giardino, J.R. (1983). A technique for assessing land surface ruggedness. *Journal of Wildlife Management*

Behrens, T., Förster, H., Scholten, T., Steinrücken, U., Spies, E.D., Goldschmitt, M. (2005). Digital soil mapping using artificial neural networks. *Journal of plant nutrition and soil science, 168*(1), 21-33.

Beier, P., Brost, B. (2010). Use of land facets to plan for climate change: conserving the arenas, not the actors. *Conservation Biology, 24*(3), 701-710.

Bell, J.C., Cunningham, R.L., Havens, M.W. (1992). Calibration and validation of a soil-landscape model for predicting soil drainage class. *Soil Science Society of America Journal, 56*(6), 1860-1866.

Benito-Calvo, A., Pérez-González, A., Magri, O., Meza, P. (2009). Assessing regional geodiversity: the Iberian Peninsula. *Earth surface processes and landforms, 34*(10), 1433-1445.

Bennett, J.M., Mattsson, L. (1999). *Introduction to Surface Roughness and Scattering* Washington, DC: Optical Society of America.

Betard, F., Peulvast, J.P. (2019). Geodiversity hotspots: Concept, method and cartographic application for geoconservation purposes at a regional scale. *Environmental Management, 63*(6), 822-834.

Beven, K. (2001). On fire and rain (or predicting the effects of change). *Hydrological processes, 15*(7), 1397-1399.

Beven, K.J., Kirkby, M.J. (1979). A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Journal, 24*(1), 43-69.

Bishop, J., Shroder Jr., M., Colby, J. (2003). Remote Sensing and Geomorphometry for studying relief production in high mountains. *Geomorphology, 55*, 345–361.

Bishop, M.P., James, L.A., Shroder Jr, J.F., Walsh, S.J. (2012). Geospatial technologies and digital geomorphological mapping: Concepts, issues and research. *Geomorphology, 137*(1), 5-26.

Bishop, T., McBratney, A., Laslett, G. (1999). Modelling soil attribute depth functions with equal-area quadratic smoothing splines. *Geoderma, 91*(1-2), 27-45.

Blum, W.E. (2016). *Role of soils for satisfying global demands for food, water, and bioenergy*. In *Environmental resource management and the nexus approach* (pp. 143-177): Springer.

Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A., Selige, T. (2001). Soil regionalisation by means of terrain analysis and process parameterisation. *Soil classification, 7*, 213.

Boixadera, J., Antúnez, M., Claret, R.M.P. (2016). Soil–landscape relationships in the Empordà basin (Catalonia, NE Iberian Peninsula). *Spanish Journal of Soil Science, 6*(3).

Bonham-Carter, G.F. (2014). *Geographic information systems for geoscientists: modelling with GIS* (Vol. Vol. 13): Elsevier.

Borrelli, P., Panagos, P., Langhammer, J., Apostol, B., Schütt, B. (2016). Assessment of the cover changes and the soil loss potential in European forestland: first approach to derive indicators to capture the ecological impacts on soil-related forest ecosystems. *Ecological Indicators, 60*, 1208-1220.

Botha, G.A., Singh, R. (2012). *Geology, geohydrology and Development Potential Zonation of the uThukela District Municipality: Specialist contribution towards the Environmental Management Framework*.Council for Geoscience, Pietermaritzburg, KwaZulu-Natal.

Bouma, J. (1989). *Using soil survey data for quantitative land evaluation*. In *Advances in soil science* (pp. 177-213): Springer.

Bouma, J. (2020). Contributing pedological expertise towards achieving the United Nations sustainable development goals. *Geoderma, 375*, 114508.

Bouma, J., Batjes, N., Groot, J. (1998). Exploring land quality effects on world food supply. *Geoderma, 86*(1-2), 43-59.

Brady, N.C., Weil, R.R. (2008). *The nature and properties of soils* (Vol. 13). Upper Saddle River, NJ: Prentice Hall

Brand, R.F., Collins, N., du Preez, P.J. (2015). A phytosociology survey and vegetation description of inselbergs in the uKhahlamba-Drakensberg Park World Heritage Site, South Africa. *koedoe, 57*(1), 1-12.

Brand, R.F., Scott-Shaw, C.R., O'Connor, T.G. (2019). The alpine flora on inselberg summits in the Maloti—Drakensberg Park, KwaZulu-Natal, South Africa. *Bothalia-African Biodiversity & Conservation, 49*(1), 1-15.

Breiman, L. (2001). Random Forests. *Machine learning, 45*(1), 5-32.

Breiman, L., Friedman, J., Stone, C., Olshen, R. (1984). *Classification and regression trees* (2nd ed.). CA; Wadsworth: Pacific Grove.

Brenner, A.C., DiMarzio, J.P., Zwally, H.J. (2007). Precision and accuracy of satellite radar and laser altimeter data over the continental ice sheets. *IEEE Transactions on Geoscience and Remote Sensing, 45*(2), 321-331.

Bretar, F., Arab-Sedze, M., Champion, J., Pierrot-Deseilligny, M., Heggy, E., Jacquemoud, S. (2013). An advanced photogrammetric method to measure surface roughness: Application to volcanic terrains in the Piton de la Fournaise, Reunion Island. *Remote Sensing of Environment, 135*, 1-11.

Bridges, E. (1994). *Soil conservation: an international issue*. In D. O‟ Halloran, Green, C., Harley, M., Stanley, M. & Knill, J. (Ed.), *Geological and Landscape Conservation* (pp. 11-15). London: Geological Society.

Brilha, J. (2005). *Geological heritage and geoconservation: nature conservation in its geological aspect*. Braga, Portugal.

Brown, D.G., Lusch, D.P., Duda, K.A. (2001). Supervised classification of types of glaciated land- scapes using digital elevation data. *Geomorphology, 21*, 233–250.

Brubaker, K.M., Myers, W.L., Drohan, P.J., Miller, D.A., Boyer, E.W. (2013). The use of LiDAR terrain data in characterizing surface roughness and microtopography. *Applied and Environmental Soil Science, 2013*.

Brubaker, S., Jones, A., Lewis, D., Frank, K. (1993). Soil properties associated with landscape position. *Soil Science Society of America Journal, 57*(1), 235-239.

Brungard, C.W. (2009). *Alternative sampling and analysis methods for digital soil mapping in southwestern Utah*. (MSc Thesis), Utah State University, Logan, Utah.

Bubenzer, O., Bolten, A. (2008). The use of new elevation data (SRTM/ASTER) for the detection and morphometric quantification of Pleistocene mega dunes (draa) in the eastern Sahara and the southern Namib. *Geomorphology, 102*, 221-231.

Buchanan, B., Fleming, M., Schneider, R., Richards, B., Archibald, J., Qiu, Z., Walter, M.J.H., Sciences, E.S. (2014). Evaluating topographic wetness indices across central New York agricultural landscapes. *18*(8), 3279-3299.

Bünemann, E.K., Bongiorno, G., Bai, Z., Creamer, R.E., De Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T.W., Mäder, P. (2018). Soil quality–A critical review. *Soil Biology and Biochemistry, 120*, 105-125.

Burek, C.V., Potter, J. (2006). *Local geodiversity action plans: setting the context for geological conservation.* English Nature Research Reports (0967-876X).English Nature, ENRR560. Northminster House, Peterborough.

Burrough, P.A. (1993). Soil variability: a late 20th century view. *Soils Fertility, 56*, 529-562.

Bushnell, T. (1943). Some Aspects of the Soil Catena Concept *Soil Science Society of America Journal, 7*(C), 466-476.

Cairncross, B. (2011). The National Heritage Resource Act (1999): can legislation protect South Africa's rare geoheritage resources? *Resources Policy, 36*(3), 204-213.

Callow, J., Van Niel, K., Boggs, G. (2007). How does modifying a DEM to reflect known hydrology affect subsequent terrain analysis? *Journal of Hydrology, 332*(1-2), 30-39.

Camp, K., G., T. (1995). *The Bioresource Units of KwaZulu-Natal.* Cedara Report N/A/95/32.Cedara College, KZN Department of Agriculture, Pietermaritzburg, South Africa.

Camp, K.G. (1999). *Guide to the use of the Bioresource Programme.*KZN Natural Resource Section: Technology Development and Training, Cedara Report No N/A/99/11. KZN Department of Agriculture, Cedara. KZNDARD

Carré, F., McBratney, A.B., Mayr, T., Montanarella, L. (2007). Digital soil assessments: Beyond DSM. *Geoderma, 142*(1-2), 69-79.

Carter, J.R. (1992). The effect of data precision on the calculation of slope and aspect using gridded DEMs. *Cartographica: The International Journal for Geographic Information and Geovisualization, 29*(1), 22-34.

Cavalli, M., Tarolli, P., Marchi, L., Dalla Fontana, G. (2008). The effectiveness of airborne LiDAR data in the recognition of channel-bed morphology. *Catena, 73*(3), 249-260.

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R. (2013). Are fine resolution digital elevation models always the best choice in digital soil mapping? *Geoderma, 195*, 111-121.

CGS. (2000). *Council for GeoScience: Cartographic Material: 2828 Harrismith, 2928 Drakensberg, 2930 Durban, 2830 Richards Bay* SACG [Digital map]. 1:250 000 Geological Series (South Africa Geological Survey).

Chabala, L.M., Mulolwa, A., Lungu, O. (2013). Landform classification for digital soil mapping in the Chongwe-Rufunsa area, Zambia. *Agro- Forest and Fisheries, 2*, 156-160.

Chang, K.-t., Tsai, B.-w. (1991). The effect of DEM resolution on slope and aspect mapping. *Cartography and geographic information systems, 18*(1), 69-77.

Chaplot, V., Walter, C., Curmi, P. (2000). Improving soil hydromorphy prediction according to DEM resolution and available pedological data. *Geoderma, 97*(3-4), 405-422.

Chen, C., Liaw, A., Breiman, L. (2004). Using random forest to learn imbalanced data. *University of California, Berkeley, 110*(1-12), 24.

Chen, Y., Wilson, J.P., Zhu, Q., Zhou, Q. (2012). Comparison of drainage-constrained methods for DEM generalization. *Computers & Geosciences, 48*, 41-49.

Chiang, S., Peterson, G. (1970). Soil catena concept for hydrologic interpretations. *Journal of soil and Water Conservation, 25*, 225-227.

Chicco, D., Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics, 21*(1), 1-13.

Cho, S.M., Lee, M. (2001). Sensitivity considerations when modeling hydrologic processes with Digital Elevation Models. *Journal of the American Water Resources Association, 37*(4), 931-934.

Chow, T.E., Hodgson, M.E. (2009). Effects of lidar post-spacing and DEM resolution to mean slope estimation. *International Journal of Geographical Information Science, 23*(10), 1277-1295.

Church, M. (2011). *Observations and experiments*. London: Sage.

Clark. (2020). TerrSet. Retrieved from https://clarklabs.org/terrset/

COGTA. (2017). *Provincial Cooperative Governance and Traditional Affairs: UThukela District Municipality District Growth and Development Plan.*Uthukela District Municipality, Provincial Cooperative Governance and Traditional Affairs, Uthukela, South Africa.

Comer, P.J., Pressey, R.L., Hunter Jr, M.L., Schloss, C.A., Buttrick, S.C., Heller, N.E., Tirpak, J.M., Faith, D.P., Cross, M.S., Shaffer, M.L. (2015). Incorporating geodiversity into conservation decisions. *Conservation Biology, 29*(3), 692-701.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1. 4. *Geoscientific Model Development, 8*(7), 1991-2007.

Conradie, D.C.U. (2013). *Appropriate Passive Design Approaches for the Various Climatic Regions in South Africa*. In L. Van Wyk (Ed.), *The Green Building Handbook, the Essential Guide* (Vol. 5, pp. 101-117). Cape Town, South Africa: Alive2Green.

Cooper, J.A.G. (1991). *Sedimentary models and geomorphological classification of river-mouths on a subtropical, wave-dominated coast, Natal, South Africa*. (PhD Thesis), the University of Natal, University of Natal, KwaZulu Natal, South Africa.

Cowen, D.J., Jensen, J.R., Hendrix, C., Hodgson, M., Schill, S.R., Macchiaverna, F. (2000). A GIS-assisted rail construction econometric model that incorporates LIDAR data. *Photogrammetric Engineering and Remote Sensing, 66*(11), 1323-1328.

Cowling, R.M., Richardson, D.M., Pierce, S.M. (2004). *Vegetation of Southern Africa*. Cambridge, UK: Cambridge University Press.

Cox, D., Oosthuizen, S., Dickens, C. (2015). *Moving from Integrated Water Resources Management (IWRM) to Integrated Natural Resources Management (INRM): A proposed framework for INRM at District Scale in South Africa***.**Institute of Natural Resources, TT 643/15. Water Research Commission. Pietermaritzburg, KwaZulu Natal, South Africa. I.o.N. Resources

Crofts, R., Gordon, J., Brilha, J., Gray, M., Gunn, J., Larwood, J., Santucci, V., Tormey, D., Worboys, G., (2020). *Guidelines for geoconservation in protected and conserved areas*. International Union for Conservation of Nature (IUCN).

Cutler, D.R., Edwards Jr, T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J. (2007). Random forests for classification in ecology. *Ecology, 88*(11), 2783-2792.

DAFF, (1983). *Conservation of Agricultural Resources Act no. 43 of 1983*, Pretoria: Government Printer. Department of Agriculture Forestry and Fisheries, Republic of South Africa.

DAFF. (2015). *Discussion document on the preservation and development of agricultural land***.**Department of Agriculture Forestry and Fisheries, Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

DAFF. (2017). *Draft Policy on the Preservation and Development of Agricultural Land***.**Department of Agriculture, Forestry and Fisheries: Natural Resources Directorate, Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa.

Dale, V.H., Kline, K.L. (2013). Issues in using landscape indicators to assess land changes. *Ecological Indicators, 28*, 91-99.

Datta, P.S., Schack-Kirchner, H. (2010). Erosion relevant topographical parameters derived from different DEMs—A comparative study from the Indian Lesser Himalayas. *Remote Sensing, 2*(8), 1941-1961.

Davis, F.W., Dozier, J. (1990). Information analysis of a spatial database for ecological land classification. *Photogrammetric Engineering and Remote Sensing, 56*(5), 605-613.

De Bruin, S., Stein, A. (1998). Soil-landscape modelling using fuzzy c-means clustering of attribute data derived from a digital elevation model (DEM). *Geoderma, 83*(1-2), 17-33.

De Bruin, S., Wielemaker, W., Molenaar, M. (1999). Formalisation of soil-landscape knowledge through interactive hierarchical disaggregation. *Geoderma, 91*(1-2), 151-172.

de Carvalho Junior, W., Lagacherie, P., da Silva Chagas, C., Calderano Filho, B., Bhering, S.B. (2014). A regional-scale assessment of digital mapping of soil attributes in a tropical hillslope environment. *Geoderma, 232*, 479-486.

De Smith, M.J., Goodchild, M.F., Longley, P. (2007). *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*: Troubador publishing ltd.

DEA. (2016). *South Africa Environment Outlook: A report on the state of the environment***.**Department of Environmental Affairs and Tourism, Department of Environmental Affairs and Tourism. Pretoria, South Africa.

Deacon, J. (2015). Great expectations: Archaeological sites and the national heritage resources act (act 25 of 1999). *The South African Archaeological Bulletin*, 220-223.

Debella-Gilo, M., Etzelmüller, B. (2009). Spatial prediction of soil classes using digital terrain analysis and multinomial logistic regression modeling integrated in GIS: Examples from Vestfold County, Norway. *Catena, 77*(1), 8-18.

Dent, D., Young, A. (1981). *Soil survey and land evaluation*: George Allen & Unwin.

Desmet, P.J.J. (1997). Effects of interpolation errors on the analysis of DEMs. *Earth surface processes and landforms: The Journal of the British Geomorphological Group, 22*(6), 563-580.

Díaz-Uriarte, R., De Andres, S.A. (2006). Gene selection and classification of microarray data using random forest. *BMC bioinformatics, 7*(1), 1-13.

Dietrich, W.E., Reiss, R., Hsu, M.L., Montgomery, D.R. (1995). A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. *Hydrological processes, 9*(3-4), 383-400.

Dikau, R., Brabb, E.E., Mark, R.M. (1991). *Landform classification of New Mexico by computer.* (2331-1258).US Dept. of the Interior, US Geological Survey,

Dixon, G. (1996). *Geoconservation: an international review and strategy for Tasmania.* Parks and Wildlife Service. Tasmania.

Djokic, D., (2011). *Creating a Hydrologically Conditioned DEM*, 2011 Esri International User Conference-Technical Workshops.

Dobos, E. (2006). *Digital soil mapping: as a support to production of functional maps*: Office for Official Publication of the European Communities.

Dobos, E., Montanarella, L. (2006). The development of a quantitative procedure for soilscape delineation using digital elevation data for Europe. *Developments in soil science, 31*, 107-605.

Dobos, E., Micheli, E., Baumgardner, M.F., Biehl, L., Helt, T. (2000). Use of combined digital elevation model and satellite radiometric data for regional soil mapping. *Geoderma, 97*(3-4), 367-391.

Dokuchaev, V., (1892). *Our Steppes–At One Time and Now*. Yevdokimoff Press, St. Petersburg.

dos Santos, F.M., Bacci, D.L.C., Saad, A.R., da Silva Ferreira, A.T. (2020). Geodiversity index weighted by multivariate statistical analysis. *Applied Geomatics*.

Doumit, J.A. (2018). Multiscale Landforms Classification Based on UAV Datasets. *Sustainability in Environment, 3*(2), 128-141.

Drăguţ, L.D., Blaschke, T. (2006). Automated classification of landform elements using object-based image analysis. *Geomorphology, 81*(3-4), 330-344.

Drăguţ, L.D., Blaschke, T. (2008). *Terrain segmentation and classification using SRTM data*. In *Advances in digital terrain analysis* (pp. 141-158): Springer.

Drăguţ, L.D., Eisank, C. (2012). Automated object-based classification of topography from SRTM data. *Geomorphology, 141*, 21-33.

Drăguţ, L.D., Csillik, O., Minár, J., Evans, I.S. (2013). Land-surface segmentation to delineate elementary forms from Digital Elevation Models. *Proceedings of the Geomorphometry*.

Driver, A., Sink, K.J., Nel, J.L., Holness, S., Van Niekerk, L., Daniels, F., Jonas, Z., Majiedt, P.A., Harris, L., Maze, K. (2012). *National Biodiversity Assessment 2011: An assessment of South Africa's biodiversity and ecosystems.* Department of Agriculture, Forestry and Fisheries South African National Biodiversity Institute. Pretoria, South Africa.

Dunn, M., Hickey, R. (1998). The effect of slope algorithms on slope estimates within a GIS. *Cartography, 27*(1), 9-15.

DWAF. (2020). *South African dams and river line shapefiles mapped at 1:500 000* [Shapefile]. Retrieved from: http://www.dwaf.gov.za/iwqs/gis_data/river/rivs500k.html

DWS. (2013). *Classification of Water Resources and Determination of the Comprehensive Reserve and Resource Quality Objectives in the Mvoti to Umzimkulu Water Management Area: Status Quo Assessment, Integrated Unit of Analysis Delineation and Biophysical Node.* Classification of Water Resources and Determination of the Comprehensive Reserve and Resource Quality Objectives in the Mvoti to Umzimkulu Water Management Area Report Number: RDM/WMA11/00/CON/CLA/0815. Department of Water and Sanitation. Pretoria, South Africa.

EKZNW. (2010). *2008 KZN Province Land Cover Mapping (from SPOT5 Satellite imagery circa 2008.* Data Users Report and MetaData Ezemvelo KwaZulu-Natal Wildlife. KwaZulu-Natal, South Africa. G.P. Ltd

Elhorst, J.P., Fischer, M.M., Getis, A. (2010). *Handbook of applied spatial analysis.* London, UK: Springer.

Elliot, F., Escott, B. (2015). *UThukela District Municipality: Biodiversity Sector Plan.* Provincial Biodiversity Sector Plan.Ezemvelo KZN Wildlife, Pietermaritzburg, KwaZulu Natal, South Africa.

EnviroGIS. (2016a). *Agricultural Suitability Zones and Land Capability Evaluation.* Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa. EnviroGIS

EnviroGIS. (2016b). *The Development of an Integrated National Geo-referenced Database Comprising of Land Capability, Land Suitability Agricultural Ecological Zones as well as Land Use Data and Information with and Applicable Scale of 1:50 000.*EnviroGIS Univen IGC Consortium, AG-861710. Department of Agriculture, Forestry and Fisheries. Pretoria, South Africa. EnviroGIS

Erasmus, D. (1998). *Forest Industry Soils Database (FSD) Co-operative.* Soil Survey Standards for Consultants.SAFCOL, SAFCOL, Pretoria, South Africa. S.A.F.C. Limited

Erikstad, L. (2013). Geoheritage and geodiversity management–the questions for tomorrow. *Proceedings of the Geologists' Association, 124*(4), 713-719.

Erskine, R.H., Green, T.R., Ramirez, J.A., MacDonald, L.H. (2007). Digital elevation accuracy and grid cell size: effects on estimated terrain attributes. *Soil Science Society of America Journal, 71*(4), 1371-1380.

ESRI, (2011). *How Aggregate works: ArcGIS Desktop Help 9.3*. In: ESRI (Ed.), http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=How%2BAggregate%2Bworks.

ESRI. (2021). *ArcGIS 10.5 & ArcPro 2.8.*Environmental Systems Research Institute Inc., New York Street, Redlands, CA, USA.

Evans, I.S. (1972). General geomorphometry, derivatives of altitude, and descriptive statistics. *Spatial analysis in geomorphology*, 17-90.

Evans, I.S. (1998). What do terrain statistics really mean? *Landform monitoring, modelling and analysis*, 119-138.

Evans, I.S. (2012). Geomorphometry and landform mapping: What is a landform? *Geomorphology, 137*(1), 94-106.

Evans, I.S., Hengl, T., Gorsevski, P. (2009). Applications in geomorphology. *Developments in soil science, 33*, 497-525.

eXtension. (2008). Remote Sensing Resampling Methods. Retrieved from at http://www.extension.org/pages/

Falorni, G., Teles, V., Vivoni, E.R., Bras, R.L., Amaratunga, K.S. (2005). Analysis and characterization of the vertical accuracy of digital elevation models from the Shuttle Radar Topography Mission. *Journal of Geophysical Research: Earth Surface, 110*(F2).

FAO. (2006). *Guidelines for soil description.* Rome: Soil Resources, Management and Conservation Service, Land and Water Development Division.Food and Agriculture Organisation of the United Nations, Rome, Italy.

Farr, T.G., Kobrick, M. (2000). Shuttle Radar Topography Mission produces a wealth of data. *Eos, Transactions American Geophysical Union, 81*(48), 583-585.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L. (2007). The shuttle radar topography mission. *Reviews of geophysics, 45*(2).

Fels, J., Zobel, R. (1995). *Landscape position and classified land type mapping for the statewide DRASTIC mapping project. North Carolina State University.* Technical report to the North Carolina Department of Environment, Health, and Natural Resources, Division of Environmental Management.North Carolina State University, Raleigh, North Carolina State University, Raleigh.

Ferrer-Valero, N. (2018). Measuring geomorphological diversity on coastal environments: A new approach to geodiversity. *Geomorphology, 318*, 217-229.

Fischer, G., van Velthuizen, H.T., Nachtergaele, F.O. (2000). *Global agro-ecological zones assessment: methodology and results.* IIASA Interim Report. IIASA, IIASA, Laxenburg, Austria.

Fisher, P.F. (1998). Improved modelling of elevation error with geostatistics. *GeoInformatica, 2*(3), 215-233.

Fisher, P.F., Tate, N.J. (2006). Causes and consequences of error in digital elevation models. *Progress in physical geography, 30*(4), 467-489.

Flanagan, D.C., Nearing, M.A. (1995). *USDA-Water Erosion Prediction Project: Hillslope profile and watershed model documentation.* National Soil Erosion Research Laboratory Report.USDA-Agricultural Research Service, USDA-Agricultural Research Service, Lafayette.

Florinsky, I. (2016). *Digital Terrain Analysis in Soil Science and Geology*. Russian Academy of Sciences, Pushchino, Russia: Academic Press.

Florinsky, I.V. (1998a). Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science, 12*(1), 47-62.

Florinsky, I.V. (1998b). Combined analysis of digital terrain models and remotely sensed data in landscape investigations. *Progress in physical geography, 22*(1), 33-60.

Florinsky, I.V., Kuryakova, G.A. (2000). Determination of grid size for digital terrain modelling in landscape investigations—exemplified by soil moisture distribution at a micro-scale. *International Journal of Geographical Information Science, 14*(8), 815-832.

Florinsky, I.V., Eilers, R.G., Manning, G., Fuller, L. (2002). Prediction of soil properties by digital terrain modelling. *Environmental Modelling & Software, 17*(3), 295-311.

Flynn, D.F.B., Gogol-Prokurat, M., Nogeire, T., Molinari, N., Richers, B.T., Lin, B.B., Simpson, N., Mayfield, M.M., DeClerck, F. (2009). Loss of functional diversity under land-use intensification across multiple taxa. *Ecology Letters, 12*(1), 22-33.

Flynn, T., Rozanov, A., de Clercq, W., Warr, B., Clarke, C. (2019a). Semi-automatic disaggregation of a national resource inventory into a farm-scale soil depth class map. *Geoderma*, 1136-1145.

Flynn, T., Van Zijl, G., Van Tol, J., Botha, C., Rozanov, A., Warr, B., Clarke, C. (2019b). Comparing algorithms to disaggregate complex soil polygons in contrasting environments. *Geoderma, 352*, 171-180.

Frankel, K.L., Dolan, J.F. (2007). Characterizing arid region alluvial fan surface roughness with airborne laser swath mapping digital topographic data. *Journal of Geophysical Research: Earth Surface, 112*(F2).

French, J. (2003). Airborne LiDAR in support of geomorphological and hydraulic modelling. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 28*(3), 321-335.

Galatowitsch, S., Frelich, L., Phillips-Mao, L. (2009). Regional climate change adaptation strategies for biodiversity conservation in a midcontinental region of North America. *Biological Conservation, 142*(10), 2012-2022.

Gallant, A.L., Brown, D.D., Hoffer, R.M., letters, r.s. (2005). Automated mapping of Hammond's landforms. *IEEE Geoscience and remote sensing letters, 2*(4), 384-388.

Gallant, J.C., Wilson, J.P. (1996). TAPES-G: a grid-based terrain analysis program for the environmental sciences. *Computers & Geosciences, 22*(7), 713-722.

Gallant, J.C., Hutchinson, M.F. (1997). Scale dependence in terrain analysis. *Mathematics and Computers in Simulation, 43*(3-6), 313-321.

Gao, J. (1997). Resolution and accuracy of terrain representation by grid DEMs at a micro-scale. *International Journal of Geographical Information Science, 11*(2), 199-212.

Garbrecht, L., Martz, J. (1994). Grid size dependency of parameters from digital elevation models. *Computers & Geosciences, 1* (20), 85-87.

Gardner, T., Sasowsky, K.C., Day, R.L. (1990). Automated extraction of geomorphometric properties from digital elevation data. *Zeitschrift fur Geomorphologie, Supplementband, 80*, 57-68.

Garrigues, S., Allard, D., Baret, F., Weiss, M. (2006). Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sensing of Environment, 103*(1), 81-96.

George, J., Suresh, K., Arya, K.R. (2018). Digital soil mapping in a Himalayan watershed using remote sensing and terrain parameters employing artificial neural network model. *Environmental Earth Science, 77* (5), 1-14.

Gerrard, A.J. (1981). *Soils and Landforms: An Integration of Geomorphology and Pedology*. London, UK: George Allen & Unwin (Publishers) Ltd.

Gesch, D.B., Larson, K. (1996). Techniques for development of global 1-kilometre digital elevation models. *Pecora Thirteen, Human Interactions with the Environment-Perspectives from Space, Sioux Falls, South Dakota*, 20-22.

Gesch, D.B., Maune, D., (2007). *Digital elevation model technologies and applications: the DEM users manual*, American Society for Photogrammetry and Remote Sensing, pp. 99-118.

Gesch, D.B., Oimoen, M., Zhang, Z., Meyer, D., Danielson, J., (2012). *Validation of the ASTER global digital elevation model version 2 over the conterminous United States*, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Congress, Melbourne, Australia, pp. B4.

Gessler, P.E., Moore, I., McKenzie, N.J., Ryan, P. (1995). Soil-landscape modelling and spatial prediction of soil attributes. *International Journal of Geographical Information Systems, 9*(4), 421-432.

Gessler, P.E., Chadwick, O., Chamran, F., Althouse, L., Holmes, K. (2000). Modelling soil-landscape and ecosystem properties using terrain attributes. *Soil Science Society of America Journal, 64*(6), 2046-2056.

Ghosh, S., Stepinski, T.F., Vilalta, R. (2009). Automatic annotation of planetary surfaces with geomorphic labels. *IEEE Transactions on Geoscience and Remote Sensing, 48*(1), 175-185.

Giasson, E., Clarke, R.T., Inda Junior, A.V., Merten, G.H., Tornquist, C.G. (2006). Digital soil mapping using multiple logistic regression on terrain parameters in southern Brazil. *Scientia Agricola, 63*(3), 262-268.

Gillin, C.P., McGuire, K. (2013). Digital terrain analysis to predict soil spatial patterns at the Hubbard Brook Experimental Forest. *Resources. Environment. Conservation*, 105.

Gillin, C.P., Bailey, S.W., McGuire, K.J., Prisley, S.P. (2015). Evaluation of LiDAR-derived DEMs through terrain analysis and field comparison. *Photogrammetric Engineering and Remote Sensing, 81*(5), 387-396.

Gingrich, P. (1992). *Introductory statistics for the social sciences*. Regina, Canada: Department of Sociology and Social Sciences, University of Regina.

Goldblatt, A. (2010). *Agriculture: Facts & Trends: South Africa.* Ceo WWF-RSA.World Wildlife Fund, Pretoria, South Africa.

Gonga-Saholiariliva, N., Gunnell, Y., Petit, C., Mering, C. (2011). Techniques for quantifying the accuracy of gridded elevation models and for mapping uncertainty in digital terrain analysis. *Progress in Physical Geography, 35*(6), 739-764.

Goodchild, M.F. (2011). Scale in GIS: An overview. *Geomorphology, 130*(1-2), 5-9.

Goodchild, M.F., Mark, D.M. (1987). The fractal nature of geographic phenomena. *Annals of the Association of American Geographers, 77*(2), 265-278.

Gordon, J.E., Barron, H.F. (2013). The role of geodiversity in delivering ecosystem services and benefits in Scotland. *Scottish Journal of Geology, 49*(1), 41-58.

Gordon, J.E., Barron, H.F., Hansom, J.D., Thomas, M.F. (2012). Engaging with geodiversity—why it matters. *Proceedings of the Geologists' Association, 123*(1), 1-6.

Gorodkin, J. (2004). Comparing two K-category assignments by a K-category correlation coefficient. *Computational biology and chemistry, 28*(5-6), 367-374.

Goudie, A.S. (2018). *The Human Impact on the Natural Environment: Past, Present and Future*. Malden, USA: Blackwell Publishing.

Goulden, T., Hopkinson, C., Jamieson, R., Sterling, S. (2016). Sensitivity of DEM, slope, aspect and watershed attributes to LiDAR measurement uncertainty. *Remote Sensing of Environment, 179*, 23-35.

Gousie, M.B., Franklin, W. (2005). Augmenting grid-based contours to improve thin-plate DEM generation. *Photogrammetric Engineering & Remote Sensing, 71*(1), 69-79.

Goyal, S., Seyfried, M., O'neill, P. (1998). Effect of digital elevation model resolution on topographic correction of airborne SAR. *International Journal of Remote Sensing, 19*(16), 3075-3096.

Grab, S., Knight, J. (2015). *Landscapes and landforms of South Africa*. Switzerland Springer International Publishing.

Grabs, T., Seibert, J., Bishop, K., Laudon, H. (2009). Modelling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology, 373*(1-2), 15-23.

Graham, B., Ashworth, G.J., Tunbridge, J.E. (2000). A Geography of Heritage: Power. *Culture and Economy*.

GRASS Development Team. (2016). Geographic Resources Analysis Support System (GRASS) Software. Retrieved from http://grass.osgeo.org.

Gray, M. (2004). *Geodiversity: valuing and conserving abiotic nature* (2nd ed.). West Sussex, UK: John Wiley & Sons.

Gray, M. (2008). Geodiversity: a new paradigm for valuing and conserving geoheritage. *Geoscience Canada*.

Gregory, K.J., Goudie, A.J.T. (2011). *Introduction to the discipline of geomorphology*. In *The SAGE Handbook of Geomorphology* (pp. 1-20). London: SAGE.

Greve, M.H., Kheir, R.B., Greve, M.B., Bøcher, P.K. (2012). Quantifying the ability of environmental parameters to predict soil texture fractions using regression-tree model with GIS and LIDAR data: The case study of Denmark. *Ecological Indicators, 18*, 1-10.

Greve, M.H., Greve, M.B., Kheir, R.B., Bøcher, P.K., Larsen, R., McCloy, K. (2010). *Comparing decision tree modelling and indicator Kriging for mapping the extent of organic soils in Denmark*. In *Digital Soil Mapping* (pp. 267-280): Springer.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T., DeAngelis, D.L. (2005). Pattern-oriented modelling of agent-based complex systems: lessons from ecology. *Science, 310*(5750), 987-991.

Grohmann, C.H., Sawakuchi, A.O. (2013). Influence of cell size on volume calculation using digital terrain models: A case of coastal dune fields. *Geomorphology, 180*, 130-136.

Grohmann, C.H., Smith, M.J., Riccomini, C. (2010). Multiscale analysis of topographic surface roughness in the Midland Valley, Scotland. *IEEE Transactions on Geoscience and Remote Sensing, 49*(4), 1200-1213.

Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician, 63*(4), 308-319.

Gruber, F.E., Baruck, J., Geitner, C. (2017). Algorithms vs surveyors: A comparison of automated landform delineations and surveyed topographic positions from soil mapping in an Alpine environment. *Geoderma, 308*, 9-25.

Gruber, F.E., Zieher, T., Rutzinger, M., Geitner, C., (2015). *Geomorphons and structure metrics for the characterization of geomorphological landscape regions in Austria*, EGU General Assembly Vienna, Austria, pp. 6873.

Grunwald, S. (2009). Multi-criteria characterization of recent digital soil mapping and modelling approaches. *Geoderma, 152*(3-4), 195-207.

Grunwald, S. (2010). *Current state of digital soil mapping and what is next*. In J.L. Boettinger, D.W. Howell, A.C. Moore, A.E. Hartemink, S. Kienast-Brown (Eds.), *Digital Soil Mapping. Bridging Research, Environmental Application, and Operation* (pp. 3-12). Springer, Netherlands: Springer.

Grunwald, S. (2016). *Environmental soil-landscape modeling: Geographic information technologies and pedometrics*: CRC Press.

Grunwald, S., Vasques, G.M., Rivero, R.G. (2015). Fusion of soil and remote sensing data to model soil properties. *Advances in Agronomy, 131*, 1-109.

GSA, (2010). *Geoscience Amendment Act, 16 of 2010*. In: C.F. Geoscience (Ed.), Gazette No. 33837. Council For Geoscience, Pretoria, South Africa Government Printers.

GTI. (2015). *2013 - 2014 South African National Land-Cover Dataset - Data User Report and Metadata.* South African National Land-Cover.South African Department of Environmental Affairs, Pretoria, South Africa. G. Image

GTI. (2019). *South African National Land-Cover 2018: Report and Accuracy Assessment.*South Africa Department of Environmental Affairs, E1434. Pretoria, South Africa. G. Image

Guber, A., Pachepsky, Y.A., Van Genuchten, M.T., Rawls, W., Simunek, J., Jacques, D., Nicholson, T., Cady, R. (2006). Field-scale water flow simulations using ensembles of pedotransfer functions for soil water retention. *Vadose Zone Journal, 5*(1), 234-247.

Guisan, A., Weiss, S.B., Weiss, A.D. (1999). GLM versus CCA spatial modelling of plant species distribution. *Plant Ecology, 143*(1), 107-122.

Guo, Q., Li, W., Yu, H., Alvarez, O. (2010). Effects of topographic variability and lidar sampling density on several DEM interpolation methods. *Photogrammetric Engineering and Remote Sensing, 76*(6), 701-712.

Guth, P.L. (1995). Slope and aspects calculations on gridded digital elevation models: examples from a geomorphometric toolbox for personal computers. *Zeitschrift für Geomorphologie. Supplementband*(101), 31-52.

Guth, P.L. (2003). *Eigenvector analysis of digital elevation models in a GIS: Geomorphometry and quality control*. In R.D. I. S. Evans, E. Tokunaga, H. Ohmori and M. Hirano (Ed.), *Concepts and Modelling in Geomorphology: International Perspectives* (pp. 199–220). Toyko: TERRAPUB.

Haas, J. (2010). *Soil moisture modelling using TWI and satellite imagery in the Stockholm region.* (MSc Thesis), Geoinformatik och Geodesi, Geoinformatik och Geodesi, Stockholm.

Hall, G. (1983). *Chapter 5: Pedology and Geomorphology*. In L.P. Wilding, N.E. Smeck, G. Hall (Eds.), *Pedogenesis and soil taxonomy: the soil orders*: Elsevier.

Hall, O., Falorni, G., Bras, R.L. (2005). Characterization and quantification of data voids in the shuttle radar topography mission data. *EEE geoscience and remote sensing letters, 2*(2), 177-181.

Hancock, G.R. (2005). The use of digital elevation models in the identification and characterization of catchments over different grid scales. *Hydrological Processes: An International Journal, 19*(9), 1727-1749.

Haneberg, W.C. (2006). Effects of digital elevation model errors on spatially distributed seismic slope stability calculations: an example from Seattle, Washington. *Environmental & Engineering Geoscience, 12*(3), 247-260.

Häring, T., Dietz, E., Osenstetter, S., Koschitzki, T., Schröder, B. (2012). Spatial disaggregation of complex soil map units: A decision-tree based approach in Bavarian forest soils. *Geoderma, 185*, 37-47.

Hartemink, A.E., Minasny, B. (2014). Towards digital soil morphometrics. *Geoderma, 230*, 305-317.

Hays, W. (1988). *Statistics, 4th ed*. New York: CBS College Publishing.

Helzer, C.J., Jelinski, D.E. (1999). The relative importance of patch area and perimeter–area ratio to grassland breeding birds. *Ecological Applications, 9*(4), 1448-1458.

Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences, 32*(9), 1283-1298.

Hengl, T., (2009). *A Practical Guide to Geostatistical Mapping*. In: http://spatial-analyst.net/book/About (Ed.), Accessed 2019.

Hengl, T., Rossiter, D.G. (2003). Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. *Soil Science Society of America Journal, 67*(6), 1810-1822.

Hengl, T., Evans, I. (2009). Mathematical and digital models of the land surface. *Developments in soil science, 33*, 31-63.

Hengl, T., Reuter, H. (2011). How accurate and usable is GDEM? A statistical assessment of GDEM using LiDAR data. *Geomorphometry, 2*, 45-48.

Hengl, T., Bajat, B., Blagojević, D., Reuter, H.I. (2008). Geostatistical modelling of topography using auxiliary maps. *Computers & Geosciences, 34*(12), 1886-1899.

Henning, D. (2018). *UGU District Municipality Environmental Management Framework***.** Strategic Environmental Management Plan.KZN-DEDTA, Sunninghill, JHB, RSA.

Henning, D., Phamphe, R., Munshi, S. (2014). *uThukela District Municipality Environmental Management Framework***.** Desired State Report. Nemai Consulting, Durban, Kwa-Zulu natal.

Herzfeld, U.C., Mayer, H., Feller, W., Mimler, M. (2000). Geostatistical analysis of glacier-roughness data. *Annals of Glaciology, 30*, 235-242.

Heung, B., Bulmer, C.E., Schmidt, M.G. (2014). Predictive soil parent material mapping at a regional-scale: a random forest approach. *Geoderma, 214*, 141-154.

Heuvelink, G.B. (1999). Propagation of error in spatial modelling with GIS. *Geographical information systems, 1*, 207-217.

Heuvelink, G.B., (2014). *Uncertainty quantification of GlobalSoilMap products*, GlobalSoilMap: basis of the global spatial soil information system. Proceedings of 1st GlobalSoilMap Conference, pp. 335-340.

Hickey, R. (2000). Slope angle and slope length solutions for GIS. *Cartography, 29*(1), 1-8.

Hjort, J., Luoto, M. (2010). Geodiversity of high-latitude landscapes in northern Finland. *Geomorphology, 115*(1-2), 109-116.

Hjort, J., Heikkinen, R.K., Luoto, M. (2012). Inclusion of explicit measures of geodiversity improve biodiversity models in a boreal landscape. *Biodiversity and Conservation, 21*(13), 3487-3506.

Hobson, R.D. (1972). *Chapter 8 - surface roughness in topography: a quantitative approach*. In R.J. Chorley (Ed.), *Spatial analysis in geomorphology*. New York, USA: Harper & Row New York.

Hodgson, M.E. (1995). What Gell Size Does the Computed Slope/Aspect Angle Represent? *Photogrammetric Engineering & Remote Sensing, 6*(5), 513-517.

Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G., Schuckman, K. (2005). An evaluation of lidar-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering & Remote Sensing, 71*(7), 817-823.

Hoffman, R., Krotkov, E., (1990). *Terrain roughness measurement from elevation maps*, Symposium on Visual Communications, Image Processing, and Intelligent Robotics Systems. International Society for Optics and Photonics, Philadelphia, PA, United States, pp. 104-114.

Hofton, M.A., Rocchio, L.E., Blair, J.B., Dubayah, R. (2002). Validation of vegetation canopy lidar sub-canopy topography measurements for a dense tropical forest. *Journal of geodynamics, 34*(3-4), 491-502.

Holmes, P.J., Grab, S.W., Knight, J. (2016). South African geomorphology: current status and new challenges. *South African Geographical Journal, 98*(3), 405-416.

Hook, P.B., Burke, I.C. (2000). Biogeochemistry in a shortgrass landscape: control by topography, soil texture, and microclimate. *Ecology, 81*(10), 2686-2703.

Horton, R.E. (1932). Drainage-basin characteristics. *Eos, transactions American geophysical union, 13*(1), 350-361.

Hu, G., Dai, W., Xiong, L., Tang, G., (2020). *Classification of Terrain Concave and Convex Landform Units by using TIN*, Geomorphometry, Perugia, Italy.

Huaxing, L. (2008). *Modelling terrain complexity*. In *Advances in digital terrain analysis* (pp. 159-176): Springer.

Hudson, B.D. (1990). Concepts of soil mapping and interpretation. *Soil Horizons, 31*(3), 63-72.

Hudson, B.D. (1992). The soil survey as a paradigm-based science. *Soil Science Society of America Journal, 56*, 836-841.

Hudson, P., Inbar, M. (2012). Land degradation and geodiversity: anthropogenic controls on environmental change. *Land Degradation & Development, 23*(4), 307-309.

Huggett, R. (1975). Soil landscape systems: a model of soil genesis. *Geoderma, 13*(1), 1-22.

Hutchinson, M. (2011). ANUDEM version 5.3, user guide. *Canberra: Fenner School of Environment and Society, Australian National University*.

Hutchinson, M.F., (1988). *Calculation of hydrologically sound digital elevation models*, Proceedings of the Third International Symposium on Spatial Data Handling. International Geographical Union Sydney.

Hutson, J.L. (1983). *Estimation of hydrological properties of South African soils*. (PhD Thesis), University of Natal, Pietermaritzburg, KwaZulu Natal, South Africa.

Ibáñez, J.J., Bockheim, J.G. (2013). *Soil Endemism and its Importance to Taxonomic Pedodiversity*. In *Pedodiversity* (pp. 202-217): CRC Press.

Ibáñez, J.J., Vargas, R.J., Vázquez-Hoehne, A. (2013). Pedodiversity state of the art and future challenges. *Pedodiversity, CRC Press (Taylor and Francis Group) Boca Ratón. California*, 1-28.

Ibáñez, J.J., De-Alba, S., Boixadera, J., King, D., Jones, R., Thomasson, A. (1995). The pedodiversity concept and its measurement: application to soil information systems. *European land information systems for agro-environmental monitoring*, 181-195.

Iwahashi, J. (1994). Development of landform classification using the digital elevation model. *Disaster Prevention Research Institute Annuals, 37*, 141-156.

Iwahashi, J., Pike, R.J. (2007). Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature. *Geomorphology, 86*(3-4), 409-440.

Iwahashi, J., Watanabe, S., Furuya, T. (2001). Landform analysis of slope movements using DEM in Higashikubiki area, Japan. *Computers & Geosciences, 27*(7), 851-865.

Iwahashi, J., Kamiya, I., Matsuoka, M., Yamazaki, D. (2018). Global terrain classification using 280 m DEMs: segmentation, clustering, and reclassification. *Progress in Earth and Planetary Science, 5*(1), 1-31.

Jackson, C., Mofutsanyana, L., Mlungwana, N. (2019). A Risk-based approach to heritage management in South Africa. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, XLII-2/W15*, 591-597.

Jahn, R., Blume, H., Asio, V., Spaargaren, O., Schad, P. (2006). *Guidelines for soil description*. Food and Agricultural Organisation of the United Nations, Rome: FAO.

James, L.A., Watson, D.G., Hansen, W.F. (2007). Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. *Catena, 71*(1), 132-144.

James, L.A., Walsh, S.J., Bishop, M.P. (2012). Geospatial technologies and geomorphological mapping. *Geomorphology*(137), 1-4.

Janitza, S., Hornung, R. (2018). On the overestimation of random forest's out-of-bag error. *PloS one, 13*(8), e0201904.

Jarihani, A.A., Callow, J.N., McVicar, T.R., Van Niel, T.G., Larsen, J. (2015). Satellite-derived Digital Elevation Model (DEM) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments. *Journal of Hydrology, 524*, 489-506.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E. (2008). *Hole-Filled Seamless SRTM Data V4***.** CGIAR-CSI SRTM 90m Database. International Centre for Tropical Agriculture (CIAT),

Jarvis, A., Rubiano, J.E., Nelson, A., Farrow, A., Mulligan, M. (2004). *Practical use of SRTM data in the tropics: Comparisons with digital elevation models generated cartographic data***.** International Center for Tropical Agricultura.Centro Internacional de Agricultura Tropical (CIAT), Apartado Aéreo, Cali, Colombia. CIAT

Jasiewicz, J., Stepinski, T., (2012). *Machine vision approach to auto-generation of high resolution, continental-scale geomorphometric map from DEM*, EGU General Assembly Conference Vienna, Austria, pp. 3373.

Jasiewicz, J., Stepinski, T.F. (2013). Geomorphons—a pattern recognition approach to classification and mapping of landforms. *Geomorphology, 182*, 147-156.

Jasiewicz, J., Netzel, P., Stepinski, T. (2014). Landscape similarity, retrieval, and machine mapping of physiographic units. *Geomorphology, 221*, 104-112.

Jasiewicz, J., Stach, A., Nowosad, J. (2015). *Terrain misclassification problem–analysis using pattern simulation approach*. In J. Jasiewicz, Z.b. Zwoliński, Mitasova H., H. T. (Eds.), *Geomorphometry for Geosciences*. Institute of Geoecology and Geoinformation, International Society for Geomorphometry, Pozman: Adam Mickiewicz University in Poznań.

JAXA. (2019). *Advanced Land Observing Satellite (ALOS) Project* J.A.E. Agency [Imagery]. Retrieved from: http://www.jaxa.jp/projects/sat/alos2/index_e.html

Jelinski, D.E., Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology, 11*(3), 129-140.

Jenness, J. (2007). *Some Thoughts on Analyzing Topographic Habitat Characteristics.* Remotely Wild.Jenness Enterprises, Flagstaff, AZ, USA.

Jenness, J. (2013). DEM surface tools for ArcGIS. Flagstaff, AZ, USA: Jenness Enterprises. Retrieved from http://www.jennessent.com/arcgis/surface_area.htm

Jenny, H. (1941). Factors of soil formation. McGraw-Hill, New York. *Factors of soil formation. McGraw-Hill, New York.*, -.

Jenson, S.K., Domingue, J.O. (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric Engineering and Remote Sensing, 54*(11), 1593-1600.

Jewitt, D., Goodman, P.S., Erasmus, B.F., O'Connor, T.G., Witkowski, E.T. (2015). Systematic land-cover change in KwaZulu-Natal, South Africa: Implications for biodiversity. *South African Journal of Science, 111*(9-10), 01-09.

Jones, E.J., McBratney, A.B. (2016). *What is digital soil morphometrics and where might it be going?* In *Digital Soil Morphometrics* (pp. 1-15): Springer.

Jones, K.H. (1998). A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers & Geosciences, 24*(4), 315-323.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, 290*(5802), 91.

Kafdar, K. (1994). The Collected Works of John W. Tukey, Volume VIII: Multiple Comparisons, 1948-1983. *Journal of the American Statistical Association, 89*(428), 1569-1569.

Kempen, B., Brus, D., Stoorvogel, J. (2011). Three-dimensional mapping of soil organic matter content using soil type-specific depth functions. *Geoderma, 162*(1-2), 107-123.

Kempen, B., Brus, D.J., Heuvelink, G.B., Stoorvogel, J.J. (2009). Updating the 1: 50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma, 151*(3-4), 311-326.

Kidd, D., Searle, R., Grundy, M., McBratney, A., Robinson, N., O'Brien, L., Zund, P., Arrouays, D., Thomas, M., Padarian, J. (2020). Operationalising digital soil mapping–Lessons from Australia. *Geoderma Regional*, e00335.

Kienzle, S. (2004). The effect of DEM raster resolution on first order, second order and compound terrain derivatives. *Transactions in GIS, 8*(1), 83-111.

Kim, D., Zheng, Y. (2011). Scale-dependent predictability of DEM-based landform attributes for soil spatial variability in a coastal dune system. *Geoderma, 164*(3-4), 181-194.

King, G., Acton, D., St. Arnaud, R. (1983). Soil-landscape analysis in relation to soil distribution and mapping at a site within the Weyburn Association. *Canadian Journal of soil science, 63*(4), 657-670.

King, L.C. (1942). *South African scenery. A textbook of Geomorphology*.

King, L.C. (1967). *South African Scenery: A Textbook of Geomorphology* (Vol. 3rd ed). Edinburgh, Scotland: Edinburgh. Oliver and Boyd, Ltd., Tweeddale Court.

Kinsey-Henderson, A.E., (2007). *The influence of the type of digital elevation models on catchment-scale slope calculations and erosion modelling*, International Congress on Modelling and Simulation. Citeseer, pp. 724-730.

Kirkby, M., Chorley, R. (1967). Throughflow, overland flow and erosion. *Hydrological Sciences Journal, 12*(3), 5-21.

Knight, J., Grab, S., Esterhuysen, A.B. (2015). *Geoheritage and geotourism in South Africa*. In S. Grab, J. Knight (Eds.), *Landscapes and Landforms of South Africa* (pp. 165-173). Switzerland: Springer International Publishing.

Kolasa, J., Rollo, C.D. (1991). *Introduction: the heterogeneity of heterogeneity: a glossary*. In J. Kolasa (Ed.), *Ecological heterogeneity* (pp. 1-23). New York Springer Verlag.

Kori, E., Onyango Odhiambo, B.D., Chikoore, H. (2019). A geomorphodiversity map of the Soutpansberg Range, South Africa. *Landform Analysis, 38*.

Korzeniowska, K., Korup, O., (2016). *Mapping gullies using terrain surface roughness*, Proceedings of the 19th AGILE International Conference on Geographic Information Science (AGILE 2016), Helsinki, Finland, pp. 14-17.

Kostrzewski, A. (2011). The role of relief geodiversity in geomorphology. *Geographia Polonica, 84*(Special Issue Part 2), 69-74.

Kozlowski, S. (2004). Geodiversity. The concept and scope of geodiversity. *Przeglad Geologiczn, 52*, 833-837.

Kraus, K., Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 53*(4), 193-203.

Krause, P., Boyle, D., Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in geosciences, 5*, 89-97.

Kumar, L. (2013). Effect of rounding off elevation values on the calculation of aspect and slope from a gridded digital elevation model. *Journal of Spatial Science, 58*(1), 91-100.

Kuo, W.L., Steenhuis, T.S., McCulloch, C.E., Mohler, C.L., Weinstein, D.A., DeGloria, S.D., Swaney, D.P. (1999). Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model. *Water Resources Research, 35*(11), 3419-3428.

Kuriakose, S.L., Devkota, S., Rossiter, D., Jetten, V. (2009). Prediction of soil depth using environmental variables in an anthropogenic landscape, a case study in the Western Ghats of Kerala, India. *Catena, 79*(1), 27-38.

KZNDARD, (2015). *Promoting Sustainable Use and Assessment of Natural Resources in KwaZulu-Natal*. In: M. KZNDARD (Ed.). KwaZulu_Natal Department of Agriculture and Rural Development: Natural Resource Section Cedara, KwaZulu-Natal.

KZNPPC. (2013). *Provincial Growth and Development Plan 2011-2030.* KZN Department of the Premier.KwaZulu Natal Provincial Planning Commission, Pietermaritzburg, KZN.

Lagacherie, P. (2008). *Digital soil mapping: a state of the art*. In A.E. Hartemink, A.B. McBratney, M.M. a-Santos (Eds.), *Digital soil mapping with limited data* (pp. 3-14). Boca Raton: Springer.

Lagacherie, P., McBratney, A. (2006). Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping. *Developments in soil science, 31*, 3-22.

Lagacherie, P., Legros, J., Burfough, P. (1995). A soil survey procedure using the knowledge of soil pattern established on a previously mapped reference area. *Geoderma, 65*(3-4), 283-301.

Lakhraj-Govender, R., Grab, S. (2019). Temperature trends for coastal and adjacent higher lying interior regions of KwaZulu-Natal, South Africa. *Theoretical and Applied Climatology, 137*(1-2), 373-381.

Lal, R., Bouma, J., Brevik, E., Dawson, L., Field, D.J., Glaser, B., Hatano, R., Hartemink, A., Kosaki, T., Lascelles, B. (2021). Soils and sustainable development goals of the United Nations (New York, USA): An IUSS perspective. *Geoderma Regional*, e00398.

Lang, M.W., Kim, V., McCarty, G.W., Li, X., Yeo, I.-Y., Huang, C., Du, L. (2020). Improved Detection of Inundation below the Forest Canopy using Normalized LiDAR Intensity Data. *Remote Sensing, 12*(4), 707.

Lark, R. (2012). Towards soil geostatistics. *Spatial Statistics, 1*, 92-99.

Lausch, A. (2015). Understanding and quantifying landscape structure - A review on relevant process characteristics, data models and landscape metrics. *Ecological Modelling, 295*, 31-41.

Lavin, J., (2013). *Palaeontological sensitivity map of South Africa*. In: SAHRA (Ed.). https://sahris.sahra.org.za/content/how-use-palaeontological-fossil-sensitivity-map/, Cape Town.

Lawler, J.J., Ackerly, D.D., Albano, C.M., Anderson, M.G., Dobrowski, S.Z., Gill, J.L., Heller, N.E., Pressey, R.L., Sanderson, E.W., Weiss, S.B. (2015). The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conservation Biology, 29*(3), 618-629.

Le Coz, M., Delclaux, F., Genthon, P., Favreau, G. (2009). Assessment of Digital Elevation Model (DEM) aggregation methods for hydrological modeling: Lake Chad basin, Africa. *Computers & Geosciences, 35*(8), 1661-1670.

Le Roux, P.A., Hensley, M., Van Rensburg, L.D., Botha, J.J. (2016). The cropping potential of South Africa: land evaluation results obtained during the last 50 years. *South African Journal of Plant and Soil, 33*(2), 83-88.

Lee, J. (1991). Analyses of visibility sites on topographic surfaces. *International Journal of Geographical Information System, 5*(4), 413-429.

Leenaers, H., Okx, J., Burrough, P. (1990). Employing elevation data for efficient mapping of soil pollution on floodplains. *Soil Use and Management, 6*(3), 105-114.

Legros, J.-P. (1996). *Cartographies des sols: de l'analyse spatiale à la gestion des territoires* (Vol. 10): PPUR presses polytechniques.

Leifman, G., Katz, S., Tal, A., Meir, R., (2003). *Signatures of 3D models for retrieval*, Proceedings of the 4th Israel-Korea Bi-National Conference on Geometric Modeling and Computer Graphics. Citeseer, pp. 159-163.

Letuka, K.W. (2016). *The potential impact of the Draft Preservation and Development of Agricultural Land Framework Bill on land use and planning laws*. (LLB Thesis), North-West University (South Africa), Potchefstroom Campus, North-West University.

Levin, S.A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. *Ecology, 73*(6), 1943-1967.

Levin, S.A. (1998). Ecosystems and the biosphere as complex adaptive systems. *Ecosystems, 1*(5), 431-436.

Li, H., Reynolds, J.F. (1993). A new contagion index to quantify spatial patterns of landscapes. *Landscape Ecology, 8*(3), 155-162.

Li, J., Wong, D.W. (2010). Effects of DEM sources on hydrologic applications. *Computers, Environment and Urban Systems, 34*(3), 251-261.

Li, X., Zhang, Y., Jin, X., He, Q., Zhang, X. (2017). Comparison of digital elevation models and relevant derived attributes. *Journal of Applied Remote Sensing, 11*(4), 046027.

Li, Z. (1988). On the measure of digital terrain model accuracy. *The photogrammetric record, 12*(72), 873-877.

Li, Z., Zhu, C., Gold, C. (2004). *Digital terrain modelling: principles and methodology*. Boca Raton: CRC Press, Taylor and Francis Group.

Li, Z., Trappe, W., Zhang, Y., Nath, B., (2005). *Robust statistical methods for securing wireless localization in sensor networks*, Proceedings of the 4th international symposium on Information processing in sensor networks. IEEE Press, pp. 12.

Liang, S. (2007). Recent developments in estimating land surface biogeophysical variables from optical remote sensing. *Progress in physical geography, 31*(5), 501-516.

Liang, X., Guo, J., Leung, L.R. (2004). Assessment of the effects of spatial resolutions on daily water flux simulations. *Journal of Hydrology, 298*(1-4), 287-310.

Libohova, Z., Winzeler, H.E., Lee, B., Schoeneberger, P.J., Datta, J., Owens, P.R. (2016). Geomorphons: Landform and property predictions in a glacial moraine in Indiana landscapes. *Catena, 142*, 66-76.

Lilburne, L., Eger, A., Mudge, P., Ausseil, A.G., Stevenson, B., Herzig, A., Beare, M. (2020). The Land Resource Circle: Supporting land-use decision making with an ecosystem-service-based framework of soil functions. *Geoderma, 363*, 114134.

Liljequist, D., Elfving, B., Skavberg Roaldsen, K. (2019). Intraclass correlation–A discussion and demonstration of basic features. *PloS one, 14*(7), e0219854.

Lin, F., Chen, X., Yao, H. (2017). Evaluating the use of Nash-Sutcliffe efficiency coefficient in goodness-of-fit measures for daily runoff simulation with SWAT. *Journal of Hydrologic Engineering, 22*(11), 05017023.

Lin, H. (2003). Hydropedology: Bridging disciplines, scales, and data. *Vadose Zone Journal, 2*(1), 1-11.

Lin, H., Wheeler, D., Bell, J., Wilding, L. (2005). Assessment of soil spatial variability at multiple scales. *Ecological Modelling, 182*(3-4), 271-290.

Lindsay, J.B., Creed, I.F. (2005). Removal of artefact depressions from digital elevation models: towards a minimum impact approach. *Hydrological Processes: An International Journal, 19*(16), 3113-3126.

Liu, X. (2008). Airborne LiDAR for DEM generation: some critical issues. *Progress in physical geography, 32*(1), 31-49.

Liu, X., Zhang, Z., Peterson, J., Chandra, S., (2007). *The effect of LiDAR data density on DEM accuracy*, Proceedings of the International Congress on Modelling and Simulation (MODSIM07). Modelling and Simulation Society of Australia and New Zealand Inc., pp. 1363-1369.

Longuet-Higgins, M. (1971). On the Shannon-Weaver index of diversity, in relation to the distribution of species in bird censuses. *Theoretical population biology, 2*(3), 271-289.

Lowenthal, D. (2005). Natural and cultural heritage. *International Journal of Heritage Studies, 11*(1), 81-92.

Luo, W., Liu, C.-C. (2018). Innovative landslide susceptibility mapping supported by geomorphon and geographical detector methods. *Landslides, 15*(3), 465-474.

Luo, Y., Liu, T., Wang, X., Duan, L., Zhang, S., Shi, J. (2012). Landform classification using soil data and remote sensing in northern Ordos Plateau of China. *Journal of Geographical Sciences, 22*(4), 681-698.

Luo, Y., Yang, S., Zhao, C., Liu, X., Liu, C., Wu, L., Zhao, H., Zhang, Y. (2014). The effect of environmental factors on spatial variability in land-use change in the high-sediment region of China's Loess Plateau. *Journal of Geographical Sciences, 24*(5), 802-814.

Macfarlane, D., Teixeira-Leite, A., Escott, B., Elliot, F., Thornhill, M., Richardson, J. (2014). *Ugu District Municipality: Biodiversity Sector Plan.* EKZNW Biodiversity Sector Plan Series.Ezemvelo KwaZulu Natal Wildlife, Pietermritzburg, KZN.

MacMillan, R., Shary, P. (2009). Landforms and landform elements in geomorphometry. *Developments in soil science, 33*, 227-254.

MacMillan, R., Pettapiece, W., Nolan, S., Goddard, T. (2000). A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems, 113*(1), 81-109.

MacMillan, R., Martin, T., Earle, T., McNabb, D. (2003). Automated analysis and classification of landforms using high-resolution digital elevation data: applications and issues. *Canadian Journal of Remote Sensing, 29*(5), 592-606.

MacMillan, R., Jones, R.K., McNabb, D.H., Systems, U. (2004). Defining a hierarchy of spatial entities for environmental analysis and modeling using digital elevation models (DEMs). *Computers, Environment and Urban Systems, 28*(3), 175-200.

MacVicar, C.N. (1974). Concerning the meaning of potential in agriculture. *South African Journal of Agriculture Extension, 3*, 1-4.

MacVicar, C.N. (1977). *Soil classification. A binomial system for South Africa*. Soil and Irrigation Research Institute, South Africa: Department of Agricultural Technical Services.

MacVicar, C.N., Scotney, D.M., Skinner, T.E., Niehaus, H.S., Loubser, J.H. (1974). A classification of land (climate, terrain form, soil) primarily for rainfed agriculture. *South African Journal of Agricultural Extention, 3*, 21-24.

MacVicar, C.N., de Villiers, J.M., Loxton R.F., Verster E., Lambrechts J.J.N., Merryweather F.R., Le Roux J., van Rooyen T.H., H.J., H. (1977). *Soil classification - A binominal system for South Africa*. Pretoria: Department of Agricultural Technical Services.

Malan, G.J. (2016). *Investigating the suitability of land type information for hydrological modelling in the mountain regions of Hessequa, South Africa*. (MSc Thesis), Stellenbosch University, Stellenbosch, South Africa.

Maling, D.H. (2016). *Measurements from maps: principles and methods of cartometry*: Butterworth-Heinemann.

Malinowska, E., Szumacher, I. (2013). Application of landscape metrics in the evaluation of geodiversity. *Miscellanea Geographica, 17*(4), 28-33.

Malo, D.D., Worcester, B., Cassel, D.K., Matzdorf, K. (1974). Soil-landscape relationships in a closed drainage system. *Soil Science Society of America Journal, 38*(5), 813-818.

Malone, B.P., Minasny, B., McBratney, A.B. (2017). *Using R for digital soil mapping*. Switzerland: Springer International Publishing

Malone, B.P., McBratney, A., Minasny, B., Laslett, G. (2009). Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma, 154*(1-2), 138-152.

Mander, Ü., Müller, F., Wrbka, T. (2005). Functional and structural landscape indicators: Upscaling and downscaling problems. *Ecological Indicators, 4*(5), 267-272.

Manosso, F.C., de Nóbrega, M.T. (2016). Calculation of geodiversity from landscape units of the Cadeado Range region in Paraná, Brazil. *Geoheritage, 8*(3), 189-199.

Marbut, C.F. (1937). *History of Soil Survey Ideas*. In G.A.E. Weber (Ed.), *The Bureau of Chemistry and Soils: Its History, Activities, and Organization.* (pp. 91-98). Washington, DC: Brookings Institution Institute for Government Research.

Mark, D. (1988). *Network models in Geomorphology*. In M. Anderson (Ed.), *Modelling Geomorphological Systems* (pp. 73–97.). New York: Wiley.

Marsman, B., de Gruijter, J.J. (1986). *Quality of soil maps: a comparison of soil survey methods in a sandy area.* Soil survey papers 15 (9032702173).ISRIC, 15. Wageningen, Netherlands.

Mashimbye, Z.E., de Clercq, W.P., Van Niekerk, A. (2014). An evaluation of digital elevation models (DEMs) for delineating land components. *Geoderma, 213*, 312-319.

Matsika, R. (2007). *Land-cover change: Threats to the grassland biome of South Africa*. (MSc Thesis), University of the Witwatersrand, Johannesburg, South Africa.

Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure, 405*(2), 442-451.

Maynard, J.J., Johnson, M.G. (2014). Scale-dependency of LiDAR-derived terrain attributes in quantitative soil-landscape modelling: Effects of grid resolution vs neighbourhood extent. *Geoderma, 230*, 29-40.

McBratney, A.B., Santos, M.M., Minasny, B. (2003). On digital soil mapping. *Geoderma, 117*(1-2), 3-52.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W. (2002). From pedotransfer functions to soil inference systems. *Geoderma, 109*(1-2), 41-73.

McBratney, A.B., Minasny, B., MacMillan, R., Carre', F. (2011). *Digital Soil Mapping*. In P.M. Huang, Y. Li, M.E. Summer (Eds.), *Handbook of Soil Sciences: Properties and Processes* (pp. 1-44). Boca Raton, FL: CRC Press, Taylor & Francis Group.

McDonnell, J.J., Woods, R. (2004). On the need for catchment classification. *Journal of Hydrology, 299*, 2-3.

McGarigal, K., Tagil, S., Cushman, S.A. (2009). Surface metrics: an alternative to patch metrics for the quantification of landscape structure. *Landscape Ecology, 24*(3), 433-450.

McHugh, M.L. (2011). Multiple comparison analysis testing in ANOVA. *Biochemia medica: Biochemia medica, 21*(3), 203-209.

McKean, J., Roering, J. (2004). Objective landslide detection and surface morphology mapping using high-resolution airborne laser altimetry. *Geomorphology, 57*(3-4), 331-351.

McKenzie, N., Austin, M. (1993). A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. *Geoderma, 57*(4), 329-355.

McKenzie, N., Gessler, P., Ryan, P., O'connell, D. (2000). *The Role of Terrain Analysis in Soil Mapping*. In J. Wilson, J.C. Gallant (Eds.), *Terrain Analysis: Principles and Applications* (Vol. 10). Wiley and Sons, Canada: Chapter.

McKinney, M.L. (2004). Measuring floristic homogenization by non-native plants in North America. *Global Ecology and Biogeography, 13*(1), 47-53.

McMaster, K.J. (2002). Effects of digital elevation model resolution on derived stream network positions. *Water Resources Research, 38*(4), 13-11-13-18.

McMaster, R.B., Monmonier, M. (1989). A conceptual framework for quantitative and qualitative raster-mode generalization. *Unknown Journal*, 390-403.

MEA. (2015). *Current state and trends: findings of the conditions and trends working group*. In R. Hassan, R. Scholes, N. Ash (Eds.), *Ecosystems and Human Well-being.* Washington DC, USA.: Island Press.

Melelli, L., Vergari, F., Liucci, L., Del Monte, M. (2017). Geomorphodiversity index: Quantifying the diversity of landforms and physical landscape. *Science of the Total Environment, 584*, 701-714.

Menezes, M.D.d., Silva, S.H.G., Owens, P.R., Curi, N. (2013). Digital soil mapping approach based on fuzzy logic and field expert knowledge. *Ciência e Agrotecnologia, 37*(4), 287-298.

Menezes, M.D.d., Silva, S.H.G., Mello, C.R.d., Owens, P.R., Curi, N. (2014). Solum depth spatial prediction comparing conventional with knowledge-based digital soil mapping approaches. *Scientia Agricola, 71*(4), 316-323.

Meybeck, M., Green, P., Vörösmarty, C. (2001). A new typology for mountains and other relief classes. *Mountain Research and Development, 21*(1), 34-45.

Meyer, D. (2011). *ASTER Global Digital Elevation Model Version 2 –Summary of Validation Results.*NASA Land Processes Distributed Active Archive Center and the Joint Japan-US ASTER Science Team,

Miller, B.A., Schaetzl, R. (2016). History of soil geography in the context of scale. *Geoderma, 264*, 284-300.

Miller, B.A., Schaetzl, R. (2014). The historical role of base maps in soil geography. *Geoderma, 230*, 329-339.

Milne, G. (1935). Some suggested units of classification and mapping, particularly of East African soils. Soil Research. *Supplements to the Proceedings of the International Society of Soil Science*(3), 183-198.

Milne, G. (1936). Normal erosion as a factor in soil profile development. *Nature, 138*(3491), 548.

Minar, J., Evans, I. (2008). Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology, 95*(3-4), 236-259.

Minasny, B., McBratney, A.B. (1999). A rudimentary mechanistic model for soil production and landscape development. *Geoderma, 90*(1-2), 3-21.

Minasny, B., Hartemink, A.E. (2011). Predicting soil properties in the tropics. *Earth-Science Reviews, 106*(1-2), 52-62.

Minasny, B., McBratney, A.B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma, 264*, 301-311.

Minasny, B., Finke, P., Stockmann, U., Vanwalleghem, T., McBratney, A.B. (2015). Resolving the integral connection between pedogenesis and landscape evolution. *Earth-Science Reviews, 150*, 102-120.

Mitášová, H., Hofierka, J. (1993). Interpolation by regularized spline with tension: II. Application to terrain modeling and surface geometry analysis. *Mathematical Geology, 25*(6), 657-669.

Mitchell, F. (2020). *Annual Operational Plan: Sub-Directorate Natural Resources***.**KwaZulu Natal Department of Agriculture and Rural Development, KZNDARD

Mitchell, M.W. (2011). Bias of the Random Forest out-of-bag (OOB) error for certain input parameters. *Open Journal of Statistics, 1*(03), 205.

Mohamed, M.A. (2020). Classification of Landforms for Digital Soil Mapping in Urban Areas Using LiDAR Data Derived Terrain Attributes: A Case Study from Berlin, Germany. *Land, 9*(9), 319.

Monckton, C.G. (1994). An investigation into the spatial structure of error in digital elevation data. *Innovations in GIS, 1*, 201-211.

Moore, I.D., Grayson, R., Ladson, A. (1991). Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological processes, 5*(1), 3-30.

Moore, I.D., Gessler, P., Nielsen, G., Peterson, G. (1993). Soil attribute prediction using terrain analysis. *Soil Science Society of America Journal, 57*(2), 443-452.

Mora, O.E. (2015). *Morphology-Based Identification of Surface Features to Support Landslide Hazard Detection Using Airborne LiDAR Data*. (PhD Thesis), The Ohio State University, The Ohio State University, USA.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE, 50*(3), 885-900.

Morisette, J.T., Privette, J.L., Justice, C.O. (2002). A framework for the validation of MODIS land products. *Remote sensing of environment, 83*(1-3), 77-96.

Morisette, J.T., Baret, F., Privette, J.L., Myneni, R.B., Nickeson, J., Garrigues, S. (2006). Validation of global moderate-resolution LAI Products: a framework proposed within the CEOS Land Product Validation subgroup. *IEEE Transactions on Geoscience and Remote Sensing, 44*(7), 1804-1817.

Motulsky, H., Searle, P. (1998). *The InStat guide to choosing and interpreting statistical tests***.** Statistics for biologists.GraphPad Software Inc, San Diego, CA.

Mücher, C.A., Klijn, J.A., Wascher, D.M., Schaminée, J.H. (2010). A new European Landscape Classification (LANDMAP): A transparent, flexible and user-oriented methodology to distinguish landscapes. *Ecological Indicators, 10*(1), 87-103.

Mueller, L., Schindler, U., Mirschel, W., Shepherd, T.G., 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.

Mueller, L., Blum, W.E., Schenk, W., Romanenkov, V., Sychev, V.G., Jones, M., Kowarik, I., Ball, B.C., McKenzie, B.M., Gerasimova, M. (2018). About landscapes and their utilisation: status and trends of landscape research. *Новые методы и результаты исследований ландшафтов в Европе, Центральной Азии и Сибири*, 16-25.

Mueller, T., Sassenrath, G.F. (2015). *GIS Applications in Agriculture: Conservation Planning*. Boca Raton: Taylor and Francis, CRC Press.

Mulder, V., De Bruin, S., Schaepman, M.E., Mayr, T. (2011). The use of remote sensing in soil and terrain mapping—A review. *Geoderma, 162*(1-2), 1-19.

Mushkin, A., Gillespie, A.R. (2005). Estimating subpixel surface roughness using remotely sensed stereoscopic data. *Remote Sensing of Environment, 99*(1-2), 75-83.

Nagy, G. (1991). Terrain visibility. *Computers & Graphics, 18*(6), 763-773.

Nanni, M.R., Povh, F.P., Demattê, J.A.M., Oliveira, R.B.d., Chicati, M.L., Cezar, E. (2011). Optimum size in grid soil sampling for variable rate application in site-specific management. *Scientia Agricola, 68*(3), 386-392.

Napieralski, J., Barr, I., Kamp, U., Kervyn, M. (2013). *Remote sensing and GIScience in geomorphological mapping*. In J. Schroder (Ed.), *Treatise in Geomorphology*. San Francisco: Academic Press.

NASA. (2011). ASTER Global Digital Elevation Model V002. *Maintained by the NASA EOSDIS Land Processes Distributed Active Archive Center (LP DAAC) at the USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. 2018,.*

Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology, 10*(3), 282-290.

Nellemann, C., Fry, G. (1995). Quantitative analysis of terrain ruggedness in reindeer winter grounds. *Arctic*, 172-176.

Nelson, A., Reuter, H., Gessler, P. (2009). DEM production methods and sources. *Developments in soil science, 33*, 65-85.

NEMA, (2008). *National Environmental Management Act, 63 of 2008*. In: S.A. Department of Environmental Affairs (Ed.), Gazette No.19519, Pretoria, South Africa Government Printers.

Neteler, H., Mitasova, M. (2007). *Open Source GIS: a GRASS GIS Approach* (3rd ed.). New York: Springer.

NGI, (2018). *Hydrological polygon and river line shapefiles mapped from the 1:50 000 topographical maps*. In: DRDLR: NGI (Ed.), Cape Town.

NHRA, (1999). *National Heritage Resources Act 25 of 1999*. In: S.A. Department of Arts and Culture (Ed.), Government Gazette,506(19974), Pretoria, South Africa Government Printers.

Nikolakopoulos, K.G., Kamaratakis, E.K., Chrysoulakis, N. (2006). SRTM vs ASTER elevation products. Comparison for two regions in Crete, Greece. *International Journal of Remote Sensing, 27*(21), 4819-4838.

Odgers, N.P., Sun, W., McBratney, A.B., Minasny, B., Clifford, D. (2014). Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma, 214*, 91-100.

Oksanen, J., Sarjakoski, T. (2005). Error propagation of DEM-based surface derivatives. *Computers & Geosciences, 31*(8), 1015-1027.

Olaya, V. (2009). Basic land-surface parameters. *Developments in soil science, 33*, 141-169.

Oldeman, L., Van Engelen, V. (1993). A world soils and terrain digital database (SOTER)—An improved assessment of land resources. *Geoderma, 60*(1-4), 309-325.

Oliveira, P.T.S., Rodrigues, D.B.B., Sobrinho, T.A., Panachuki, E., Wendland, E. (2013). Use of SRTM data to calculate the (R) USLE topographic factor. *Acta Scientiarum. Technology, 35*(3), 507-513.

Oliver, M., Webster, R. (2014). A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena, 113*, 56-69.

Ott, R.L., Larson, R., Rexroat, C., Mendenhall, W. (1992). *Statistics: A tool for the social sciences*. Belmont, CA: Duxbury Press.

Padilla, R., Netto, S.L., da Silva, E.A., (2020). *A survey on performance metrics for object-detection algorithms*, 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). IEEE, pp. 237-242.

Panagos, P., Borrelli, P., Meusburger, K. (2015). A new European slope length and steepness factor (LS-Factor) for modelling soil erosion by water. *Geosciences, 5*(2), 117-126.

Panizza, M., (2007). *Workshop Abstracts: Geodiversity, geological heritage and geotourism*, Geomorphorsites, Geoparks and Geotourism, Lesvos, Greece.

Panizza, M. (2009). The geomorphodiversity of the Dolomites (Italy): a key of geoheritage assessment. *Geoheritage, 1*(1), 33-42.

Papadimitriou, F. (2009). Modelling spatial landscape complexity using the Levenshtein algorithm. *Ecological Informatics, 4*(1), 48-55.

Park, S., Rucker, G., Agyare, W., Vlek, P. (2009). Influence of grid cell size and flow routing algorithm on soil-landform modelling. *Journal of the Korean Geographical Society, 44*(2), 122-145.

Parks, K., Mulligan, M. (2010). On the relationship between a resource-based measure of geodiversity and broad-scale biodiversity patterns. *Biodiversity and Conservation, 19*(9), 2751-2766.

Partridge, T., Dollar, E., Moolman, J., Dollar, L. (2010). The geomorphic provinces of South Africa, Lesotho and Swaziland: A physiographic subdivision for earth and environmental scientists. *Transactions of the Royal Society of South Africa, 65*(1), 1-47.

Paterson, G., Turner, D., Wiese, L., Van Zijl, G., Clarke, C., Van Tol, J. (2015). Spatial soil information in South Africa: Situational analysis, limitations and challenges. *South African Journal of Science, 111*(5-6), 1-7.

Paterson, S. (2018). *Soil Spatial Scaling: Modelling variability of soil properties across scales using legacy data*. (PhD Thesis), The University of Sydney, Sydney Australia.

PDALB, (2016). *PDALB*. Department Of Agriculture Forestry and Fisheries, South Africa.

Pellitero, R., Manosso, F.C., Serrano, E. (2015). Mid-and large-scale geodiversity calculation in fuentes carrionas (nw spain) and serra do cadeado (paraná, brazil): methodology and application for land management. *Geografiska Annaler: Series A, Physical Geography, 97*(2), 219-235.

Penížek, V., Borůvka, L. (2006). Soil depth prediction supported by primary terrain attributes a comparison of methods. *Plant, Soil & Environment, 52*(9), 424-430.

Pennock, D., Veldkamp, A. (2006). Advances in landscape-scale soil research. *Geoderma, 133*(1-2), 1-5.

Pennock, D.J., Zebarth, B., De Jong, E. (1987). Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada. *Geoderma, 40*(3-4), 297-315.

Pereira, D.I., Pereira, P., Brilha, J., Santos, L. (2013). Geodiversity assessment of Paraná State (Brazil): an innovative approach. *Environmental Management, 52*(3), 541-552.

Peters, J., De Baets, B., Verhoest, N.E., Samson, R., Degroeve, S., De Becker, P., Huybrechts, W. (2007). Random forests as a tool for ecohydrological distribution modelling. *Ecological Modelling, 207*(2-4), 304-318.

Petrasova, A., Harmon, B., Petras, V., Mitasova, H. (2015). *Tangible modelling with open source GIS*. Switzerland: Springer International

Phillips, J.D. (1988). The role of spatial scale in geomorphic systems. *Geographical Analysis, 20*(4), 308-317.

Piacente, S. (2005). Geosites and geodiversity for a cultural approach to geology. *Il Quaternario, 18*(1), 11-14.

Pierson, F., Mulla, D. (1990). Aggregate stability in the Palouse region of Washington: effect of landscape position. *Soil Science Society of America Journal, 54*(5), 1407-1412.

Piikki, K., Söderström, M., Stenberg, B. (2013). Sensor data fusion for topsoil clay mapping. *Geoderma, 199*, 106-116.

Pike, R.J. (1988). The geometric signature: quantifying landslide-terrain types from digital elevation models. *Mathematical Geology, 20*(5), 491-511.

Pike, R.J. (2000). Geomorphometry-diversity in quantitative surface analysis. *Progress in physical geography, 24*(1), 1-20.

Pilesjö, P., Hasan, A. (2014). A triangular form-based multiple flow algorithm to estimate overland flow distribution and accumulation on a digital elevation model. *Transactions in GIS, 18*(1), 108-124.

Podobnikar, T. (2005). Production of integrated digital terrain model from multiple datasets of different quality. *International Journal of Geographical Information Science, 19*(1), 69-89.

Ponce-Hernandez, R., Marriott, F., Beckett, P. (1986). An improved method for reconstructing a soil profile from analyses of a small number of samples. *Journal of Soil Science, 37*(3), 455-467.

Prasad, A.M., Iverson, L.R., Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems, 9*(2), 181-199.

Prasannakumar, V., Shiny, R., Geetha, N., Vijith, H. (2011). Applicability of SRTM data for landform characterisation and geomorphometry: a comparison with contour-derived parameters. *International Journal of Digital Earth, 4*(5), 387-401.

Pratt, D.J., Gwynne, M., Blackie, J., Bogdan, A., Bredon, R. (1977). Rangeland management and ecology in East Africa. *Environmental Conservation, 5*(4), 316-.

Prima, O.D.A., Echigo, A., Yokoyama, R., Yoshida, T. (2006). Supervised landform classification of Northeast Honshu from DEM-derived thematic maps. *Geomorphology, 78*(3-4), 373-386.

Prosser, C.D., Bridgland, D.R., Brown, E.J., Larwood, J.G. (2011). Geoconservation for science and society: challenges and opportunities. *Proceedings of the Geologists' Association, 122*(3), 337-342.

Przydatek, B., Song, D., Perrig, A., (2003). *Sia: secure information aggregation in sensor networks, SenSys' 03: Proceedings of the 1st international conference on Embedded networked sensor systems*. ACM Press, New York, NY, USA.

Quattrochi, D.A., Goodchild, M.F. (1997). *Scale in remote sensing and GIS*: CRC press.

Quinn, P., Beven, K., Chevallier, P., Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological processes, 5*(1), 59-79.

R Core Team, (2013). *R: A language and environment for statistical computing*. Vienna, Austria.

Raaflaub, L.D., Collins, M. (2006). The effect of error in gridded digital elevation models on the estimation of topographic parameters. *Environmental Modelling & Software, 21*(5), 710-732.

Rabus, B., Eineder, M., Roth, A., Bamler, R. (2003). The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing, 57*(4), 241-262.

Ramamurthy, V. (2018). *Soil-Based Land Use Planning: Impacts on Livelihood Security*. In *Geospatial Technologies in Land Resources Mapping, Monitoring and Management* (pp. 551-567): Springer.

Rassool, C. (2013). Towards critical heritage studies. *Material Religion, 9*(3), 403-404.

Reddy, G.O., Singh, S. (2018). *Geospatial technologies in land resources mapping, monitoring and management*: Springer.

Regmi, N.R., Rasmussen, C. (2018). Predictive mapping of soil-landscape relationships in the arid Southwest United States. *Catena, 165*, 473-486.

Reinoso, J. (2010). A priori horizontal displacement (HD) estimation of hydrological features when versioned DEMs are used. *Journal of Hydrology, 384*(1-2), 130-141.

Reinoso, J.F., León, C., Mataix, J. (2016). Estimating horizontal displacement between DEMs by means of particle image velocimetry techniques. *Remote Sensing, 8*(1), 14.

Reuter, H., Hengl, T., Gessler, P., Soille, P. (2009). Preparation of DEMs for geomorphometric analysis. *Developments in Soil Science, 33*, 87-120.

Reuter, H.I. (2004). *Spatial crop and soil-landscape processes under special consideration of relief information in a loess landscape*. (PhD Thesis), Universität Hannover,

Rexer, M., Hirt, C. (2014). Comparison of free high-resolution digital elevation data sets (ASTER GDEM2, SRTM v2. 1/v4. 1) and validation against accurate heights from the Australian National Gravity Database. *Australian Journal of Earth Sciences, 61*(2), 213-226.

Reza Pahlavan Rad, M., Toomanian, N., Khormali, F., Brungard, C.W., Bayram Komaki, C., Bogaert, P. (2014). Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. *Geoderma, 232*(97), 232.

Rigol-Sanchez, J.P., Stuart, N., Pulido-Bosch, A. (2015). ArcGeomorphometry: a toolbox for geomorphometric characterisation of DEMs in the ArcGIS environment. *Computers & Geosciences, 85*, 155-163.

Riley, S.J., DeGloria, S.D., Elliot, R. (1999). An index that quantifies topographic heterogeneity. *intermountain Journal of Sciences, 5*(1-4), 23-27.

Robinson, N., Regetz, J., Guralnick, R.P. (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of Photogrammetry and Remote Sensing of Environment, 87*, 57-67.

Rodriguez, E., Morris, C., Belz, J., Chapin, E., Martin, J., Daffer, W., Hensley, S. (2005). *An assessment of the SRTM topographic products.* JPL Technical Report.Jet Propulsion Lab, Jet Propul. Lab., Pasadena, CA. JPL

Roecker, S.M. (2013). *Solving for y: digital soil mapping using statistical models and improved models of land surface geometry*. (MSc Thesis), West Virginia University, West Virginia University, VA.

Rossi, V., Salinari, F., Poni, S., Caffi, T., Bettati, T. (2014). Addressing the implementation problem in agricultural decision support systems: the example of vite. net®. *Computers and Electronics in Agriculture, 100*, 88-99.

Ruban, D., A. (2010). Quantification of geodiversity and its loss. *Proceedings of the Geologists' Association, 121*(3), 326-333.

Ruhe, R.V. (1975). *Geomorphology: geomorphic processes and surficial geology*: Houghton Mifflin.

Ruhe, R.V., Walker, P. (1968). Hillslope models and soil formation. I. Open systems. *Int Soc Soil Sci Trans*.

Saadat, H., Bonnell, R., Sharifi, F., Mehuys, G., Namdar, M., Ale-Ebrahim, S. (2008). Landform classification from a digital elevation model and satellite imagery. *Geomorphology, 100*(3-4), 453-464.

Sampson, C.C., Fewtrell, T.J., Duncan, A., Shaad, K., Horritt, M.S., Bates, P.D. (2012). Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models. *Advances in water resources, 41*, 1-17.

SANBI. (2012). *Vegmap Ecosystem Protection Level: GIS metadata report.* Pretoria, South Africa. SANBI

Santos, D.S., Mansur, K.L., Gonçalves, J.B., Junior, E.R.A., Manosso, F.C. (2017). Quantitative assessment of geodiversity and urban growth impacts in Armação dos Búzios, Rio de Janeiro, Brazil. *Applied geography, 85*, 184-195.

SASRI. (1999). *Identification and management of the soils of the South African Sugar Industry.*South African Sugar Research Institute, Mount Edgecombe, Durban, South Africa. SASRI

Sauder, D.C., DeMars, C.E. (2019). An updated recommendation for multiple comparisons. *Advances in Methods and Practices in Psychological Science, 2*(1), 26-44.

Saulnier, G.M., Obled, C., Beven, K. (1997). Analytical compensation between DTM grid resolution and effective values of saturated hydraulic conductivity within the Topmodel framework. *Journal of Hydrological Processes, 11*(9), 1331-1346.

Sayre, R., Frye, C., Karagulle, D., Krauer, J., Breyer, S., Aniello, P., Wright, D.J., Payne, D., Adler, C., Warner, H. (2018). A new high-resolution map of world mountains and an online tool for visualizing and comparing characterizations of global mountain distributions. *Mountain Research and Development, 38*(3), 240-249.

Scarpone, C., Schmidt, M.G., Bulmer, C.E., Knudby, A. (2016). Modelling soil thickness in the critical zone for Southern British Columbia. *Geoderma, 282*, 59-69.

Schoeman, F., Newby, T., Thompson, M., Van den Berg, E.C. (2013). South African national land-cover change map. *South African Journal of Geomatics, 2*(2), 94-105.

Schoeman, J., Macvicar, C. (1978). A method for evaluating and presenting the agricultural potential of land at regional scales. *Agrochemophysica, 10*(3), 25-37.

Schoeneberger, P.J., Wysocki, D.A., Benham, E.C. (2012). *Field book for describing and sampling soils* (Vol. 3). Natural Resources Conservation Service, USDA, Lincoln, NE: Government Printing Office.

Schoorl, J., Sonneveld, M., Veldkamp, A. (2000). Three-dimensional landscape process modelling: the effect of DEM resolution. *Earth surface processes and landforms: The Journal of the British Geomorphological Group, 25*(9), 1025-1034.

Schröder, B., Seppelt, R. (2006). Analysis of pattern–process interactions based on landscape models—overview, general concepts, and methodological issues. *Ecological Modelling, 199*(4), 505-516.

Schulze, R.E. (2007). *Agrohydrological Information Needs, Information Sources and Decision Support.* South African Atlas of Climatology and Agrohydrology.Water Research Commission, 1489/1/06. Pretoria.

Schulze, R.E., Maharaj, M. (2007). *Rainfall Concentration.* South African Atlas of Climatology and Agrohydrology.WRC, 1489/1/06. Pretoria, South Africa.

Schulze, R.E., Chapman, R.D. (2007). *Estimation of Daily Solar Radiation over South Africa.* South African Atlas of Climatology and Agrohydrology.WRC, 1489/1/06. Pretoria, South Africa.

Schulze, R.E., Schütte, S. (2020). Mapping soil organic carbon at a terrain unit resolution across South Africa. *Geoderma, 373*, 114447.

Schulze, R.E., Horan, M.J.C., Lumsden, T.G., Warburton, M.L. (2007). *Recharge to the Groundwater Store and Baseflow.* South African Atlas of Climatology and Agrohydrology.Water Research Commission, Report 1489/1/06. Pretoria, RSA.

Schumann, K., Wittig, R., Thiombiano, A., Becker, U., Hahn, K. (2011). Impact of land-use type and harvesting on population structure of a non-timber forest product-providing tree in a semi-arid savanna, West Africa. *Biological Conservation, 144*(9), 2369-2376.

Scotney, D.M. (1987). *Interim report on the SARCCUS agro-ecological zones project. Report presented to the twentieth ordinary meeting of the Southern African Regional Commission for the Conservation and Utilisation of the Soil.*DAFF, Pretoria.

SCWG. (1991). *Soil Classification: A Taxonomic System for South Africa*. Pretoria: Department of Agricultural Development: Soil and Irrigation Research Unit.

SCWG. (2018). *Soil Classification: A Natural and Anthropogenic System For South Africa*. Pretoria, South Africa: Agricultural Research Council, Institute for Soil, Climate and Water.

Serrano, E., Ruiz-Flaño, P. (2007). Geodiversity: a theoretical and applied concept. *Geographica helvetica, 62*(3), 140-147.

Serrano, E., González Trueba, J.J. (2011). Environmental education and landscape leisure. Geotourist map and geomorphosites in the Picos de Europa National Park. *Geojournal of Tourism and Geosites, 8*(2), 295-308.

Serrano, E., Ruiz-Flaño, P., Arroyo, P. (2009). Geodiversity assessment in a rural landscape: Tiermes-Caracena area (Soria, Spain). *Memorie Descrittive Della Carta Geoligica d'Italia, 87*, 173-180.

Shafique, M., van der Meijde, M., Kerle, N., van der Meer, F. (2011). Impact of DEM source and resolution on topographic seismic amplification. *International Journal of Applied Earth Observation and Geoinformation, 13*(3), 420-427.

Shahid, S.A., Taha, F.K., Abdelfattah, M. (2013). *Developments in Soil Classification, Land Use Planning and Policy Implications*: Springer.

Sharifzadeh, M., Shahabi, C., Knoblock, C. (2005). *Learning approximate thematic maps from labelled geospatial data*. In P. Agouris, A. Croituru (Eds.), *Next-Generation Geospatial Information. From Digital Image Analysis to Spatio Temporal Databases* (Vol. 3). London, UK: Taylor and Francis.

Sharma, A., Tiwari, K. (2014). A comparative appraisal of hydrological behaviour of SRTM DEM at catchment level. *Journal of Hydrology, 519*, 1394-1404.

Sharma, A., Tiwari, K., Bhadoria, P. (2010). Vertical accuracy of digital elevation model from Shuttle Radar Topographic Mission–a case study. *Geocarto International, 25*(4), 257-267.

Sharma, A., Tiwari, K., Bhadoria, P. (2011). Determining the optimum cell size of digital elevation model for hydrologic application. *Journal of earth system science, 120*(4), 573-582.

Sharples, C. (1993). *A methodology for the identification of significant landforms and geological sites for geoconservation purposes.*Forestry Commission Tasmania, Tasmania. F.C. Tasmania

Sharples, C. (1995). Geoconservation in forest management principles and procedures. *Tastforest Hobart, 7*, 37-50.

Shary, P.A. (1995). Land surface in gravity points classification by a complete system of curvatures. *Mathematical Geology, 27*(3), 373-390.

Shary, P.A., Sharaya, L.S., Mitusov, A.V. (2002). Fundamental quantitative methods of land surface analysis. *Geoderma, 107*(1-2), 1-32.

Shepard, M.K., Campbell, B.A., Bulmer, M.H., Farr, T.G., Gaddis, L.R., Plaut, J.J. (2001). The roughness of natural terrain: A planetary and remote sensing perspective. *Journal of Geophysical Research: Planets, 106*(E12), 32777-32795.

Shi, X. (2010). User's Guide of ArcSIE. *SIE LLC, Bow, NH, USA*.

Shi, X., Girod, L., Long, R., DeKett, R., Philippe, J., Burke, T. (2012). A comparison of LiDAR-based DEMs and USGS-sourced DEMs in terrain analysis for knowledge-based digital soil mapping. *Geoderma, 170*, 217-226.

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

Silva, J.P., Pereira, D.I., Aguiar, A.M., Rodrigues, C. (2013). Geodiversity assessment of the Xingu drainage basin. *Journal of Maps, 9*(2), 254-262.

Silva, S.H.G., de Menezes, M.D., Owens, P.R., Curi, N. (2016a). Retrieving pedologist's mental model from existing soil map and comparing data mining tools for refining a larger area map under similar environmental conditions in Southeastern Brazil. *Geoderma, 267*, 65-77.

Silva, S.H.G., De Menezes, M.D., De Mello, C.R., De Góes, H.T.P., Owens, P.R. (2016b). Geomorphometric tool associated with soil types and properties spatial variability at watersheds under tropical conditions. *Scientia Agricola, 73*(4), 363-370.

Simonson, R.W. (1968). *Concept of soil*. In *Advances in Agronomy* (Vol. 20, pp. 1-47): Elsevier.

Singh, R.G. (2009). *Landslide classification, characterization and susceptibility modelling in KwaZulu-Natal*. (MSc Thesis), University of the Witwatersrand, Johannesburg, Gauteng, South Africa.

Skidmore, A.K. (1989). A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographical Information System, 3*(4), 323-334.

Skidmore, A.K., Ryan, P.J., Dawes, W., Short, D., O'LOUGHLIN, E. (1991). Use of an expert system to map forest soils from a geographical information system. *International Journal of Geographical Information System, 5*(4), 431-445.

Small, R.J. (1978). *The study of Landforms: a textbook of geomorphology* (Vol. 2nd edition). London, UK: Cambridge University Press.

Smirnoff, A., Paradis, S., Boivin, R. (2008). Generalizing surficial geological maps for scale change: ArcGIS tools vs cellular automata model. *Computers & Geosciences, 34*(11), 1550-1568.

Smith, B. (2006). *The farming handbook*. KwaZulu Natal, South Africa: University of Natal Press.

Smith, H., Hudson, B.D. (2002). *The American soil survey in the twenty-first century*. Iowa, USA: Iowa State Press.

Smith, M., Pain, C. (2009). Applications of remote sensing in geomorphology. *Progress in physical geography, 33*(4), 568-582.

Smith, M.P., Zhu, A.-X., Burt, J.E., Stiles, C. (2006). The effects of DEM resolution and neighbourhood size on digital soil survey. *Geoderma, 137*(1-2), 58-69.

Smith, M.W. (2014). Roughness in the earth sciences. *Earth-Science Reviews, 136*, 202-225.

Snapir, B., Hobbs, S., Waine, T. (2014). Roughness measurements over an agricultural soil surface with Structure from Motion. *ISPRS Journal of Photogrammetry and Remote Sensing, 96*, 210-223.

Somerville, P.N. (1993). On the conservatism of the Tukey-Kramer multiple comparison procedure. *Statistics probability letters, 16*(5), 343-345.

Sorensen, R., Seibert, J. (2007). Effects of DEM resolution on the calculation of topographical indices: TWI and its components. *Journal of Hydrology, 347*(1-2), 79-89.

Sörensen, R., Zinko, U., Seibert, J. (2006). On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology & Earth System Sciences Discussions, 10*(1), 101-112.

Speight, J. (1977). Landform pattern description from aerial photographs. *Photogrammetria, 32 no 5*, 161-182.

Spijker, V. (2013). *Morphological Landscape-classifications of the Front Range of the Rocky Mountains near Boulder, Colorado*. (MSc Thesis), Wageningen University and Research Centre, Wageningen University and Research Centre, Netherlands.

Srinivasan, R., Engel, B. (1991). Effect of slope prediction methods on slope and erosion estimates. *Applied Engineering in Agriculture, 7*(6), 779-783.

Stage, A.R., Salas, C. (2007). Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *Forest Science, 53*(4), 486-492.

Stalz, J. (2007). Mapping potential soil salinization using rule-based object-oriented image analysis [MSc thesis]. *Stellenbosch: Stellenbosch University*.

STATSSA. (2009). *South African Statistics, 2009***.**Statistics South Africa,

Stein, A., Ettema, C. (2003). An overview of spatial sampling procedures and experimental design of spatial studies for ecosystem comparisons. *Agriculture, Ecosystems & Environment, 94*(1), 31-47.

Stein, A., Riley, J., Halberg, N. (2001). Issues of scale for environmental indicators. *Agriculture, Ecosystems and Environment, 87*(2), 215-232.

Stepinski, T., Bagaria, C., (2009). *A two-stage classification approach for effective geomorphic mapping of planetary surfaces*, Proceedings of Geomorphometry, Zurich, Switzerland, pp. 1-5.

Stone, R.O., Dugundji, J. (1965). A study of microrelief—its mapping, classification, and quantification by means of a Fourier analysis. *Engineering Geology, 1*(2), 89-187.

Stott, T. (2010). Fluvial geomorphology. *Progress in Physical Geography, 34*(2), 221-245.

Strand, G.-H. (2011). Uncertainty in classification and delineation of landscapes: A probabilistic approach to landscape modelling. *Environmental Modelling Software, 26*(9), 1150-1157.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A. (2008). Conditional variable importance for random forests. *BMC bioinformatics, 9*(1), 1-11.

Studley, H., Weber, K. (2011). *Comparison of image resampling techniques for satellite imagery***.** Final Report: Assessing Post-Fire Recovery of Sagebrush-Steppe Rangelands in Southeastern Idaho.Idaho State University: GIS Training and Research Center, Idaho State University, Pocatello, USA.

Stum, A.K., Boettinger, J., White, M., Ramsey, R. (2010). *Random forests applied as a soil spatial predictive model in arid Utah*. In *Digital soil mapping* (pp. 179-190): Springer.

Su, J., Bork, E. (2006). Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. *Photogrammetric Engineering & Remote Sensing, 72*(11), 1265-1274.

Swanwick, C. (2002). *Landscape Character Assessment: Guidance for England and Scotland***.**Department of Landscape University of Sheffield Shefield, UK.

Tachikawa, T., Kaku, M., Iwasaki, A., Gesch, D., Oimoen, M., Zhang, Z., Danielson, J., Krieger, T., Curtis, B., Haase, J., Abrams, M., Crippen, R., Carabajal, C. (2011). *ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results***.** (ASTER GDEM Validation Team Report). Earth Resources Observation and Science (EROS) Center, NASA. NASA

Tagil, S., Jenness, J. (2008). GIS-based automated landform classification and topographic, landcover and geologic attributes of landforms around the Yazoren Polje, Turkey. *Journal of Applied Sciences, 8*(6), 910-921.

Tang, G. (2000). *Research on the accuracy of digital elevation models*. (PhD Thesis), the University of Salzburg, University of Salzburg, Austria.

Tarolli, P., Arrowsmith, J.R., Vivoni, E.R. (2009). Understanding earth surface processes from remotely sensed digital terrain models. *Geomorphology, 113*(1), 1.

Tate, N., Atkinson, P.M. (2001). *Modelling scale in geographical information science*. West Sussex, London: John Wiley & Sons.

Temme, A., Heuvelink, G., Schoorl, J., Claessens, L. (2009). Geostatistical simulation and error propagation in geomorphometry. *Developments in soil science, 33*, 121-140.

TERRASOLID. (2016). Terra Scan. Retrieved from at http://www.terrasolid.com/products.php

TERRATEC. (2016). TerraPos. *Computer Software.* Retrieved from at http://www.terratec.no/terrapos.html

Tesfa, T.K., Tarboton, D.G., Chandler, D.G., McNamara, J.P. (2009). Modelling soil depth from topographic and land cover attributes. *Water Resources Research, 45*(10).

Theobald, D.M. (1989). Accuracy and bias issues in surface representation. *Accuracy of spatial databases*, 99-106.

Thomas, I., Jordan, P., Shine, O., Fenton, O., Mellander, P.-E., Dunlop, P., Murphy, P. (2017). Defining optimal DEM resolutions and point densities for modelling hydrologically sensitive areas in agricultural catchments dominated by microtopography. *International journal of applied earth observation geoinformation, 54*, 38-52.

Thomas, J.A., Joseph, S., Thrivikramji, K., Arunkumar, K. (2014). Sensitivity of digital elevation models: The scenario from two tropical mountain river basins of the Western Ghats, India. *Geoscience Frontiers, 5*(6), 893-909.

Thompson, J.A., Bell, J.C., Butler, C.A. (2001). Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. *Geoderma, 100*(1-2), 67-89.

Thompson, J.A., Roecker, S., Grunwald, S., Owens, P.R. (2012). *Digital soil mapping: Interactions with and applications for hydropedology*. In H. Lin (Ed.), *Hydropedology: Synergistic Integration of Soil Science and Hydrology*: Academic Press, Elsevier.

Thwaites, R.N. (2000). From biodiversity to geodiversity and soil diversity. A spatial understanding of soil in ecological studies of the forest landscape. *Journal of Tropical Forest Science*, 388-405.

Tóth, G., Montanarella, L., Stolbovoy, V., Máté, F., Bódis, K., Jones, A., Panagos, P., Van Liedekerke, M. (2008). *Soils of the European Union.* JRC Scientific and Technical Reports.European Commission Joint Research Centre Institute for Environment and Sustainability Luxembourg.

Tran, T., Raghavan, V., Masumoto, S., Vinayaraj, P., Yonezawa, G. (2014). A geomorphology-based approach for digital elevation model fusion-case study in Danang city, Vietnam. *Earth Surface Dynamics, 2*(2).

Trentin, R., Robaina, L.E.d.S. (2016). Classification Of The Landform Units Supported By Geomorphometric Attributes. *Mercator, 15*(3), 53-66.

Troeh, F.R. (1964). Landform parameters correlated to soil drainage. *Soil Science Society of America Journal, 28*(6), 808-812.

Troeh, F.R., Hobbs, J.A., Donahue, R.L. (1991). *Chapter six: predicting soil loss*. In *Soil and water conservation* (2 ed.). Englewood Cliffs, New Jersey, USA: Prentice-Hall.

Tukiainen, H., Alahuhta, J., Field, R., Ala-Hulkko, T., Lampinen, R., Hjort, J. (2017). Spatial relationship between biodiversity and geodiversity across a gradient of land-use intensity in high-latitude landscapes. *Landscape Ecology, 32*(5), 1049-1063.

Turner, D.P. (2007). *The use of soil and land type information maps in road engineering in South Africa*. Paper presented at the 14th ARC SMGE, Yaounde, Cameroon, Africa.

Turner, R., Panciera, R., Tanase, M.A., Lowell, K., Hacker, J.M., Walker, J.P. (2014). Estimation of soil surface roughness of agricultural soils using airborne LiDAR. *Remote Sensing of Environment, 140*, 107-117.

UN. (2020). *Sustainable Development Goals Report.* SDGs Transform Our World. United Nations, New York, NY.

UNESCO, (2000). *United Nations Educationa, Scientific and Cultural Organisation* https://whc.unesco.org/en/list/985/.

USDA. (2017). *Soil Science Division Staff.: Soil survey manual* (C. Ditzler, K. Scheffe, H.C. Monger Eds.). Government Printing Office, Washington, D.C.

USGS. (2008). GLSDEM: Global Land Cover Facility. Retrieved from at https://lta.cr.usgs.gov/GLS

Uuemaa, E., Mander, Ü., Marja, R. (2013). Trends in the use of landscape spatial metrics as landscape indicators: a review. *Ecological Indicators, 28*, 100-106.

Vajda, A., Venäläinen, A. (2003). The influence of natural conditions on the spatial variation of climate in Lapland, northern Finland. *International Journal of Climatology: A Journal of the Royal Meteorological Society, 23*(9), 1011-1022.

van Asselen, S., Seijmonsbergen, A. (2006). Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser DTM. *Geomorphology, 78*(3-4), 309-320.

Van den Bergh, H., Weepener, H., Metz, M. (2009). *Spatial modelling for semi-detailed soil mapping in KwaZulu-Natal.*ARC-ISCW, Pretoria, South Africa.

Van der Eyk, J.J., MacVicar, C.N., De Villiers, J. (1969). *Soils of the Tugela basin*. Pietermaritzburg, South Africa: PMB Town and Regional Planning Commission.

Van Engelen, V., Dijkshoorn, J. (2013). Global and national soils and terrain digital databases (SOTER). *Report-ISRIC World Soil Information*(2013/04).

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, A., (2012). *Developing a very high resolution DEM of South Africa*, Position IT, pp. 55-60.

Van Niekerk, A. (2014). *Stellenbosch University Digital Elevation Model (SUDEM).*Centre for Geographical Analysis, Stellenbosch, South Africa.

Van Niekerk, A. (2015). *Stellenbosch University Digital Elevation Model (SUDEM)-2015 Edition.*Centre for Geographical Analysis, Stellenbosch University, Stellenbosch, South Africa.

Van Niekerk, A. (2016). *Stellenbosch University Digital Elevation Model (SUDEM): 2016 Edition* Centre For Geographical Analysis.Stellenbosch University, Stellenbosch, Western Cape.

Van Niel, K.P., Laffan, S.W., Lees, B.G. (2004). Effect of error in the DEM on environmental variables for predictive vegetation modelling. *Journal of Vegetation Science, 15*(6), 747-756.

Van Remortel, R.D., Hamilton, M.E., Hickey, R.J. (2001). Estimating the LS factor for RUSLE through iterative slope length processing of digital elevation data within ArcInfo grid. *Cartography, 30*(1), 27-35.

Van Remortel, R.D., Maichle, R.W., Hickey, R.J. (2004). Computing the LS factor for the Revised Universal Soil Loss Equation through array-based slope processing digital elevation data using a C++ executable. *Computers & Geosciences, 30*(9-10), 1043-1053.

Van Westen, C. (2016). Landslide hazard zonation in high-risk areas of Rethymno Prefecture, Crete Island, Greece. *Engineering Geology and The Environment, vol 2 no 1*, 1-11.

Van Zijl, G. (2019). Digital soil mapping approaches to address real-world problems in southern Africa. *Geoderma, 337*, 1301-1308.

Van Zijl, G., Le Roux, P. (2014). Creating a conceptual hydrological soil response map for the Stevenson Hamilton research supersite, Kruger National Park, South Africa. *Water SA, 40*(2), 331-336.

Van Zijl, G., Bouwer, D., Van Tol, J., Le Roux, P. (2014). Functional digital soil mapping: A case study from Namarroi, Mozambique. *Geoderma, 219*, 155-161.

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 Soil and water conservation, 33*(4), 257-264.

Van Zijl, G.M., Le Roux, P.A., Turner, D.P. (2013). Disaggregation of land types using terrain analysis, expert knowledge and GIS methods. *South African Journal of Plant and Soil, 30*(3), 123-129.

Vaysse, K., Lagacherie, P. (2015). Evaluating digital soil mapping approaches for mapping GlobalSoilMap soil properties from legacy data in Languedoc-Roussillon (France). *Geoderma Regional, 4*, 20-30.

Vaze, J., Teng, J., Spencer, G. (2010). Impact of DEM accuracy and resolution on topographic indices. *Environmental Modelling Software, 25*(10), 1086-1098.

Veronesi, F. (2012). *3D Advance mapping of soil properties*. (PhD Thesis), Cranfield University, Cranfield University.

Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M. (1997). Human domination of Earth's ecosystems. *Science, 277*(5325), 494-499.

Vogel, H.-J., Clothier, B., Li, X.-Y., Lin, H. (2013). Hydropedology—A perspective on current research. *Vadose Zone Journal, 12*(4).

Wagner, D., (2004). *Resilient aggregation in sensor networks*, SASN. Citeseer, pp. 78-87.

Walsh, S., Moody, A., Allen, T., Brown, D. (1997). Scale dependence of NDVI and its relationship to mountainous terrain. *Scale in remote sensing and GIS*, 27-55.

Wang, J.F., Li, X.H., Christakos, G., Liao, Y.L., Zhang, T., Gu, X., Zheng, X.Y. (2010). Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *International Journal of Geographical Information Science, 24*(1), 107-127.

Wang, L.J., Sawada, K., Moriguchi, S. (2013). Landslide-susceptibility analysis using light detection and ranging-derived digital elevation models and logistic regression models: a case study in Mizunami City, Japan. *Journal of Applied Remote Sensing, 7*(1), 073561.

Wang, L.J., Tian, B., Koike, K., Hong, B., Ren, P. (2017). Integration of Landscape Metrics and Variograms to Characterize and Quantify the Spatial Heterogeneity Change of Vegetation Induced by the 2008 Wenchuan Earthquake. *ISPRS- International Journal of Geo-Information, 6*, 1-13.

Wang, Y.-J., Qin, C.-Z., Zhu, A.-X. (2019). Review on algorithms of dealing with depressions in grid DEM. *Annals of GIS, 25*(2), 83-97.

Warren, S.D., Hohmann, M.G., Auerswald, K., Mitasova, H. (2004). An evaluation of methods to determine slope using digital elevation data. *Catena, 58*(3), 215-233.

Wascher, D.M. (2005). Landscape character: linking space and function. Final ELCAI Project Report, Landscape Europe. *European Landscape Character Areas e Typology, Cartography and Indicators for the Assessment of Sustainable Landscapes.*, 1-4.

Washtell, J., Carver, S., Arrell, K., (2009). *A viewshed based classification of landscapes using geomorphometrics*, Proc. Geomorphometry Conf., pp. 44-49.

Watkins, R. (2015). *Terrain Metrics and Landscape Characterization from Bathymetric Data: SAGA GIS Methods and Command Sequences***.** Report prepared for the Ecospatial Information Team, Coral Reef Ecosystem Division, Pacific Islands Fisheries Science Center.Coral Reef Ecosystem Division, Pacific Islands Fisheries Science Center, Honolulu, HI.

Webster, R., Oliver, M.A. (1990). *Statistical methods in soil and land resource survey*: Oxford University Press (OUP).

Wechsler, S.J.H., Sciences, E.S. (2007). Uncertainties associated with digital elevation models for hydrologic applications: a review. *11*(4), 1481-1500.

Wechsler, S.P. (2003). Perceptions of digital elevation model uncertainty by DEM users. *URISA, 15*(2), 57-64.

Wechsler, S.P., Kroll, C.N. (2006). Quantifying DEM uncertainty and its effect on topographic parameters. *Photogrammetric Engineering and Remote Sensing of Environment, 72*(9), 1081-1090.

Weeks, R.J., Smith, M., Pak, K., Li, W.H., Gillespie, A., Gustafson, B. (1996). Surface roughness, radar backscatter, and visible and near-infrared reflectance in Death Valley, California. *Journal of Geophysical Research: Planets, 101*(E10), 23077-23090.

Weiler, M., McDonnell, J. (2004). Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology. *Journal of Hydrology, 285*(1-4), 3-18.

Weiss, A., (2001). *Topographic position and landforms analysis*, Poster presentation, ESRI user conference, San Diego, CA.

Wickens, T.D., Keppel, G. (2004). *Design and analysis: A researcher's handbook*: Pearson Prentice-Hall.

Wiese, L., Ros, I., Rozanov, A., Boshoff, A., de Clercq, W., Seifert, T. (2016). An approach to soil carbon accounting and mapping using vertical distribution functions for known soil types. *Geoderma, 263*, 264-273.

Wilding, L., Bouma, J., Goss, D.W. (1994). Impact of spatial variability on interpretive modelling. *Quantitative modelling of soil-forming processes, 39*, 61-75.

Williams, L.J., Abdi, H. (2010). Fisher's least significant difference (LSD) test. *Encyclopedia of research design, 218*, 840-853.

Willmott, C.J., Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*(1), 79-82.

Wilson, C. (1994). *Earth heritage conservation*: Open University.

Wilson, J.P. (2000). Effect of data source, grid resolution, and flow-routing method on computed topographic attributes. *Terrain analysis: principles & applications*, 133-161.

Wilson, J.P. (2012). Digital terrain modelling. *Geomorphology, 137*(1), 107-121.

Wilson, J.P. (2018). *Environmental applications of digital terrain modeling*. Oxford, UK: John Wiley & Sons.

Wilson, J.P., Gallant, J.C. (2000). *Terrain analysis: principles and applications*. New York: John Wiley & Sons.

Wischmeier, W.H. (1965). Predicting rainfall erosion losses from cropland east of the Rocky Mountains. *Agriculture handbook, 282*, 47.

Wischmeier, W.H., Smith, D.D. (1978). Predicting rainfall erosion losses-a guide to conservation planning. *Predicting rainfall erosion losses-a guide to conservation planning.*

Wise, S. (2000). Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological processes, 14*(11-12), 1909-1929.

Wise, S. (2007). Effect of differing DEM creation methods on the results from a hydrological model. *Computers & Geosciences, 33*(10), 1351-1365.

Wise, S. (2011). Cross-validation as a means of investigating DEM interpolation error. *Journal of Computers and Geosciences, 37*(8), 978-991.

Wojtkowski, P. (2003). *Landscape agroecology*. New York, NY: Food Products Press.

Wolock, D.M., Price, C.V. (1994). Effects of digital elevation model map scale and data resolution on a topography-based watershed model. *Water Resources Research, 30*(11), 3041-3052.

Wood, J. (2002). LandSerf: visualisation and analysis of terrain models. Retrieved from http://www.landserf.org/

Wood, W.F. (1960). *A quantitative system for classifying landforms*: Headquarters, Quartermaster Research and Engineering Command, US Army ….

Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology, 19*(2), 125-138.

Wu, S., Li, J., Huang, G. (2005). An evaluation of grid size uncertainty in empirical soil loss modelling with digital elevation models. *Environmental Modeling & Assessment, 10*(1), 33-42.

Wu, S., Li, J., Huang, G. (2008). A study on DEM-derived primary topographic attributes for hydrologic applications: Sensitivity to elevation data resolution. *Applied Geography, 28*(3), 210-223.

www.BlueMarble.com. (2019). GlobalMapper. Retrieved from www.bluemarblegeo.com

XLStat. (2020). How to interpret contradictory results between ANOVA and multiple pairwise comparisons. Retrieved from at https://help.xlstat.com/s/tutorials/

Xue-Rong, D., Yu-Xin, S., Li-Zhong, Y., Liang-Jie, L., Xiao-Qin, H. (2002). Relationship of upslope contribution area and soil water content in TOPMODEL. *Progress in Geography, 21*(2), 103-110.

Yang, L., Meng, X., Zhang, X. (2011). SRTM DEM and its application advances. *International Journal of Remote Sensing, vol. 32, no. 14*, 3875–3896.

Yang, P., Ames, D.P., Fonseca, A., Anderson, D., Shrestha, R., Glenn, N.F., Cao, Y. (2014). What is the effect of LiDAR-derived DEM resolution on large-scale watershed model results? *Environmental modelling and software, 58*, 48-57.

Yemefack, M. (2005). *Modelling and Monitoring Soil and Land-use Dynamics: within shifting agricultural landscape mosaic systems in southern Cameroon.* (PhD Thesis), University of Utrecht, Utrecht, The Netherlands.

Yokoyama, R., Shirasawa, M., Pike, R.J. (2002). Visualizing topography by openness: a new application of image processing to digital elevation models. *Photogrammetric Engineering and Remote Sensing, 68*(3), 257-266.

Zakšek, K., Podobnikar, T., (2005). *An effective DEM generalization with basic GIS operations*, 8th ICA WORKSHOP on Generalisation and Multiple Representation. A Coruña, July.

Zaloumis, N.P. (2013). *South African grassland ecology and its restoration.* (MSc Thesis), University of Cape Town, Cape Town.

Zaragozí, B., Belda, A., Linares, J., Martínez-Pérez, J., Navarro, J., Esparza, J. (2012). A free and open source programming library for landscape metrics calculations. *Environmental Modelling & Software, 31*, 131-140.

Zerizghy, M., Le Roux, P., Hensley, M., Van Rensburg, L. (2013). Prediction of soil distribution on two soilscapes inland type Del 7 east of Bloemfontein, South Africa. *South African Journal of Science, 109 no 11*, 53–58.

Zevenbergen, L.W., Thorne, C.R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes Landforms, 12*(1), 47-56.

Zhang, G.-L., Liu, F., Song, X.-D., Zhao, Y.-G. (2016). Digital soil mapping across Paradigms, scales, and boundaries: A review. *Digital Soil Mapping Across Paradigms, Scales and Boundaries*, 3-10.

Zhang, J.X., Chang, K.T., Wu, J.Q. (2008). Effects of DEM resolution and source on soil erosion modelling: a case study using the WEPP model. *International Journal of Geographical Information Science, 22*(8), 925-942.

Zhang, W., Montgomery, D.R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water resources research, 30*(4), 1019-1028.

Zhang, X., Chen, J., Zhang, G., Tan, M., Ibáñez, J. (2003). Pedodiversity analysis in Hainan Island. *Journal of Geographical Sciences, 13*(2), 181-186.

Zhou, Q., Liu, X.J.C. (2004). Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences, 30*(4), 369-378.

Zhou, Q., Chen, Y. (2011). Generalization of DEM for terrain analysis using a compound method. *ISPRS Journal of Photogrammetry & Remote Sensing, 66*(1), 38-45.

Zhou, Q., Lees, B., Tang, G.-a. (2008). *Advances in digital terrain analysis*: Springer.

Zhu, A.-X. (1997). A similarity model for representing soil spatial information. *Geoderma, 77*(2-4), 217-242.

Zhu, A.-X., Band, L.E., Dutton, B., Nimlos, T.J. (1996). Automated soil inference under fuzzy logic. *Ecological Modelling, 90*(2), 123-145.

Zinck, J.A. (2013). *Geopedology, Elements of geomorphology for soil and geohazard studies*. In J.A. Zinck, G. Metternicht, G. Bocco, H.F. Del Valle (Eds.), *Geopedology: An Integration of Geomorphology and Pedology for Soil and Landscape Studies*. Switzerland: Springer International Publishing

Zunckel, K. (2013). *Investments into ecological infrastructure in the greater Umgeni River catchment.*

Zwolinski, Z., (2018). *Spatial scales of geodiversity and landform taxonomic hierarchy*, Geoheritage and Conservation: Modern Approaches and Applications Towards the 2030 Agenda, Chęciny, Poland.

Zwoliński, Z., (2008). *Designing a map of the geodiversity of landforms in Poland*, IAG and AIGEO International Meeting Environmental Analysis and Geomorphological Mapping for a Sustainable Development, Addis Ababa, Ethiopia, pp. 18-22.

Zwoliński, Z., Najwer, A., Giardino, M. (2018). *Methods for assessing geodiversity*. In E. Reynard, J. Brilha (Eds.), *Geoheritage: Assessment, Protection, and Management* (pp. 27-52): Elsevier.

Appendix A

South African Journal of Geomatics, Vol. 6. No. 2, Geomatics Indaba 2017 Special Edition, August 2017

EVALUATING THE EFFECTS OF GENERALISATION APPROACHES AND DEM RESOLUTION ON THE EXTRACTION OF TERRAIN INDICES IN KWAZULU NATAL, SOUTH AFRICA

Jonathan T. Atkinson[ab], Andrei B. Rozanov, Willem P De Clercq*[c]

[a] Kwazulu Natal Department of Agriculture and Rural Development, Cedara, 3245, KZN
[b]Stellenbosch Department of Soil Science, Matieland, Stellenbosch
[c]Stellenbosch Water Research Institute, Matieland, Stellenbosch,

*Corresponding authors email: wpdc@sun.ac.za

DOI: http://dx.doi.org/10.4314/sajg.v6i2.9

Abstract

Digital elevation model (DEM) data are elemental in deriving primary topographic attributes which serve as input variables to a variety of hydrologic and geomorphologic studies. There is however still varied consensus on the effect of DEM source and resolution on the application of these topographic attributes to landscape characterisation. While elevation data for South Africa are available from several major sources and resolutions: Shuttle Radar Topographic Mission (SRTM), Earth ENV and Stellenbosch University DEM (SUDEM). Limited research has been conducted in a local context comparing the extraction of terrain attributes to high resolution Digital Terrain Data (DTM) such as LiDAR (Light Detection and Ranging) that are becoming increasing available. However, the utility of LiDAR to topographic analyses presents its own challenges in terms of operational-relevant resolution, processing demands and limited spatial coverage. There is a need to quantify the impact that generalisation approaches have on simplifying detailed DEMs and to compare the accuracy and reliability of results between high resolution and coarse resolution data on the extraction of localized topographic variables. In this regional study, we analyse the accuracy on selected local terrain attributes: elevation, slope and topographic wetness index derived from DEMs from varying sources, at different spatial resolutions and using three generalisation algorithms, namely: mean cell aggregation, nearest neighbour and hydrological corrected topo-to-raster. We show that topographic variable extraction is highly dependent on DEM source and generalisation approach and while higher resolution DEMs may represent the "true" surface more accurately, they do not necessarily offer the best results for all extracted variables. Our results highlight the caveats of selecting DEMs not "fit-for-purpose" for topographic analysis and offer a simple yet effective solution for reconciling the selection of DEMs based on neighbourhood size resolution prior to terrain analyses and topographic feature characterization.

Appendix B

Chapter 5
Print ISBN: 978-93-90431-28-1, eBook ISBN: 978-93-90431-36-6

Assessment and Evaluating the Effects of Generalisation Approaches and DEM Resolution on the Extraction of Terrain Indices in Kwazulu Natal, South Africa

Jonathan T. Atkinson[1*] and Willem P. De Clercq[2]

DOI: 10.9734/bpi/ireges/v6

ABSTRACT

Digital elevation model (DEM) data are elemental in deriving primary topographic attributes which serve as input variables to a variety of hydrologic and geomorphologic studies. There is, however, still varied consensus on the effect of DEM source and resolution on the application of these topographic attributes to landscape characterization. At the same time, elevation data for South Africa are available from several major sources and resolutions: Shuttle Radar Topographic Mission (SRTM), EarthEnv and Stellenbosch University DEM (SUDEM). Limited research has been conducted in a local context comparing the extraction of terrain attributes to high-resolution Digital Terrain Data (DTM) such as LiDAR (Light Detection and Ranging) that are becoming increasingly available. However, the utility of LiDAR to topographic analyses presents its challenges in terms of operational-relevant resolution, processing demands and limited spatial coverage. There is a need to quantify the impact that generalization approaches have on simplifying detailed DEMs and to compare the accuracy and reliability of results between high resolution and coarse resolution data on the extraction of localized topographic variables. In this regional study, we analyze the efficiency on selected local terrain attributes: elevation, slope, topographic wetness index and surface roughness derived from DEMs from varying sources, at different spatial resolutions and using three generalization algorithms, namely: mean cell aggregation, nearest neighbour and hydrological corrected topo-to- raster. We show that topographic variable extraction is highly dependent on DEM source and generalization approach. While higher resolution DEMs may represent the "true" surface more accurately, they do not necessarily offer the best results for all extracted variables. Our results highlight the caveats of selecting DEMs not "fit-for-purpose" for topographic analysis and offer a simple yet effective solution for reconciling the selection of DEMs based on neighbourhood size resolution before terrain analyses and topographic feature characterization. Finding the right combination of when to upscale surface data, what DEMS to use and what spatial scale operate at to ensure that surface integrity is most optimal is mostly still context-specific. However, this study demonstrates a robust framework to interpret optimal sensor choice and spatial scale for understanding the geomorphological processes in the landscape for the southern-coastal area of KZN.

Keywords: SRTM; EarthEnv; DEM; SUDEM; ASTER; generalisation.

Appendix C

Geoderma Regional 22 (2020) e00291

Contents lists available at ScienceDirect

Geoderma Regional

journal homepage: www.elsevier.com/locate/geodrs

Multi-resolution soil-landscape characterisation in KwaZulu Natal: Using geomorphons to classify local soilscapes for improved digital geomorphological modelling

Jonathan Atkinson [a,*], Willem de Clercq [b], Andrei Rozanov [a]

[a] Department of Soil Science, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa
[b] Stellenbosch Water Institute, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

ARTICLE INFO

Article history:
Received 31 January 2020
Received in revised form 28 April 2020
Accepted 30 April 2020

Keywords:
Gleysols
Geomorphon
SRTM
Ferralsols

ABSTRACT

Continual advances in quantitative modelling of surface processes, combined with new spatio-temporal and geo-computational algorithms, have revolutionised the auto-classification and mapping of landform components through the automated analysis of high-quality digital elevation models (DEMs). Digital geomorphic mapping (DGM) approaches that can simplify and translate the inclusion of "human knowledge" to automatic terrain classification across a broader spectrum of terrain morphological units as well as a range of spatial scales, therefore, offer great potential for improved topographic and landscape analysis. One such approach is the mapping of landform elements using the concept of the Geomorphon (geomorphological phonotypes). The output of the geomorphon approach is the stratification of the landscape into ten unique but recognisable landform elements: peak, ridge, shoulder, spur, and slope, hollow, foot slope, valley, depression and flat. Equally appealing is the way the model self-adapts to local topography using a line-of-sight principle enabling better matching of landform elements to computational spatial scale. The purpose of this paper is to observe the effects that different pixel resolution (grain size) and digital elevation model source (DEM) would have on the replication of observed geomorphic spatial patterns and representation of terrain selected parameters within the landscape. This paper provides a comprehensive exploratory assessment of digital terrain representation and relief classification using an automated geomorphometric mapping approach, by evaluating three different digital surface models (SUDEM, SRTM, ASTER GDEM2) and different spatial resolution (30 m & 90 m) for an 11,200 ha catchment in KwaZulu-Natal, South Africa. To test the self-adapting ability of the geomorphon approach under regional conditions, we use 4750 gridded terrain samples to quantitatively analyse how the choice of terrain model and scale influence the extraction, generalisation and representation of digitally-derived terrain attributes such as slope, elevation and terrain unit feature extent. We further show how the variation in resulting terrain unit representation is limited by spatial resolution discontinuities of selected elementary soil association distribution, soil texture and soil depth. We also introduce the results of a Similarity Index used to gauge the degree of recall and precision between the different geomorphic landscape features. Finally, the findings of the regional geomorphon-soil relationships are presented in a readily interpretable and qualitative manner, providing a "quasi-landscape signature" for potential localised geomorphons. The application of the study findings may be beneficial to practitioners looking to align or refine modelled terrain classification approaches with expert perception and formalised heuristic approaches.